

Increasing Motorway Operational Capacity by Dynamic Adaptation of Variable Speed Limit Zones Based on Cooperative Learning Agents

Kušić, Krešimir

Doctoral thesis / Doktorski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Transport and Traffic Sciences / Sveučilište u Zagrebu, Fakultet prometnih znanosti**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:119:495083>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-05-09**



Repository / Repozitorij:

[Faculty of Transport and Traffic Sciences -
Institutional Repository](#)





University of Zagreb

FACULTY OF TRANSPORT AND TRAFFIC SCIENCES

Krešimir Kušić

**INCREASING MOTORWAY OPERATIONAL
CAPACITY BY DYNAMIC ADAPTATION OF
VARIABLE SPEED LIMIT ZONES BASED ON
COOPERATIVE LEARNING AGENTS**

DOCTORAL THESIS

Zagreb, 2023



University of Zagreb

FACULTY OF TRANSPORT AND TRAFFIC SCIENCES

Krešimir Kušić

**INCREASING MOTORWAY OPERATIONAL
CAPACITY BY DYNAMIC ADAPTATION OF
VARIABLE SPEED LIMIT ZONES BASED ON
COOPERATIVE LEARNING AGENTS**

DOCTORAL THESIS

Supervisors: Assoc. Prof. Marko Šoštarić & Prof. Edouard Ivanjko

Zagreb, 2023



Sveučilište u Zagrebu
FAKULTET PROMETNIH ZNANOSTI

Krešimir Kušić

**POVEĆANJE OPERATIVNOGA KAPACITETA
AUTOCESTA DINAMIČKOM PRILAGODBOM
ZONA PROMJENJIVOGA OGRANIČENJA BRZINE
ZASNOVANOM NA KOOPERATIVNIM UČEĆIM
AGENTIMA**

DOKTORSKI RAD

Mentori: Izv. prof. dr. sc. tech. Marko Šoštarić i

Prof. dr. sc. tech. Edouard Ivanjko

Zagreb, 2023.

About the first mentor:

Marko Šoštarić is an associate professor at the Faculty of Transport and Traffic Sciences, University of Zagreb. The courses he teaches include Transport Technological Design, Parking and Garages, Management of Urban Transport Systems, Ecology in Transport, Basics of Transport Engineering, and Basics of Transport Infrastructure. Under his mentorship, numerous master's and bachelor's theses have been defended, as well as one doctoral dissertation. In his scientific and professional work, he is engaged in traffic planning and design, regulation and organization of traffic flows, traffic safety, and sustainable mobility. He is a chartered transport engineer and a chartered road safety auditor. As a researcher, team member, or project leader, he has participated in over 200 scientific, research, or professional studies and projects in the field of transportation and traffic. Some of the notable projects he has been involved in are SumBOOST – Sustainable Urban Mobility Boost Smart Toolbox (H2020-EIT), HiReach innovative mobility to cope with transport poverty (H2020), Study on the Provision of Multimodal Traffic Information in the Republic of Croatia, Transport Master Plan of the City of Zagreb and the surrounding area, and the Transportation Part of the Development Strategy of the City of Split. He has published the results of his scientific and research work in approximately fifty scientific and professional papers in journals and proceedings published in Croatia and abroad, including notable works such as Sustainable Urban Mobility Boost Smart Toolbox Upgrade, Data-Driven Methodology for Sustainable Urban Mobility Assessment and Improvement, Model for Estimating Urban Mobility Based on the Records of User Activities in Public Mobile Networks, Sustainable Urban Mobility Plans, and Data Envelopment Analysis for Determining the Efficiency of Variant Solutions for Traffic Flow Organization. Since 2022, he has been the dean of the Faculty of Transport and Traffic Sciences, University of Zagreb.

O prvom mentoru:

Marko Šoštarić je izvanredni profesor na Fakultetu prometnih znanosti Sveučilišta u Zagrebu. Kolegiji na kojima predaje su: Prometno tehnološko projektiranje, Parkiranje i garaže, Upravljanje prometnim sustavima u urbanim sredinama, Ekologija u prometu, Osnove prometnog inženjerstva i Osnove prometne infrastrukture. Pod njegovim mentorstvom obranjeni su brojni diplomski i završni radovi te jedna doktorska disertacija. U svom znanstvenom i stručnom radu, bavi se prometnim planiranjem i projektiranjem, regulacijom i organizacijom promet-

nih tokova, prometnom sigurnošću te održivom mobilnošću. Ovlašteni je inženjer cestovnog prometa te ovlašteni revizor cestovne sigurnosti. Kao istraživač, član tima ili voditelj projekta, sudjelovao je u više od 200 znanstvenih, istraživačkih ili stručnih studija i projekata iz područja prometa i transporta. Neki od značajnijih projekata u kojima je sudjelovao su SumBOOST – Sustainable Urban Mobility Boost Smart Toolbox (H2020-EIT), HiReach innovative mobility to cope with transport poverty (H2020), Studija pružanja multimodalnih prometnih informacija u Republici Hrvatskoj, Master plan prometnog razvitka grada Zagreba i okolnog područja te Prometni dio strategije razvoja Grada Splita. Rezultate svojeg znanstvenog istraživanja i stručnog rada objavio je u otprilike pedeset znanstvenih i stručnih radova, koji su izašli u uglednim časopisima i zbornicima, kako u Hrvatskoj tako i u inozemstvu. Neki od značajnijih su Sustainable Urban Mobility Boost Smart Toolbox Upgrade, Data-Driven Methodology for Sustainable Urban Mobility Assessment and Improvement, Model for Estimating Urban Mobility Based on the Records of User Activities in Public Mobile Networks, Sustainable Urban Mobility Plans i Data Envelopment Analysis for Determining the Efficiency of Variant Solutions for Traffic Flow Organisation. Od 2022. godine obavlja dužnost dekana Fakulteta prometnih Znanosti Sveučilišta u Zagrebu.

About the second mentor:

Edouard Ivanjko is a Professor at the Department for Intelligent Transport Systems, Faculty of Transport Sciences, University of Zagreb. He teaches courses related to computer science, electrical engineering, artificial intelligence, virtual reality and traffic control. He holds membership in the Center of Excellence for Computer Vision and the Scientific Center of Excellence for Data Science and Advanced Cooperative Systems, Research Unit: Data Science. He is a member of the Committee for Quality Management, the Committee for Science and Projects, and the Committee for Rewarding Employees and Students of the Faculty of Transport Sciences, University of Zagreb. His research interests include intelligent transport systems, modeling and simulation of road networks, application of artificial intelligence and computer vision in road traffic control, estimation and prediction of traffic parameters and inclusion of connected autonomous vehicles in the control loop. He serves on the editorial boards of the following scientific journals: Journal of Traffic and Transportation Engineering and International Journal of Intelligent Transportation Systems Research. He is a member of IEEE (Institute of Electrical and Electronics Engineers), IEEE Intelligent Transportation Systems Society, ITS Croatia professional associations, and the company PTV scientific network. He is also an associated member of the Croatian Academy of Engineering (HATZ). He was awarded the Josip Lončar Bronze Plaque Award of the Faculty of Electrical Engineering and Computer Science, University of Zagreb, as the best automation graduate student in 2001, the award of the Faculty of Traffic Sciences, University of Zagreb, for special achievements in teaching in 2017, the company Siemens award for achievements in improvement of education and lectures on automation with the introduction of SIMATIC automation systems in 2018, the "Plaque on Edouard Ivanjko as a sign of gratitude for contributing to the development of the Faculty" by the University St. Kliment Ohridski, Bitola Technical Faculty in Bitola, North Macedonia in 2021, and the Faculty of Transport Sciences, University of Zagreb award, for special achievements in scientific research work in 2022. He has published more than 100 articles in international scientific journals and conferences, chapters in scientific books, extended abstracts and technical reports. He is proficient in both English and German.

O drugom mentoru:

Edouard Ivanjko je redoviti profesor u Zavodu za inteligentne transportne sustave, Fakultet prometnih znanosti, Sveučilišta u Zagrebu. Predaje kolegije vezane za računalstvo, elek-

trotehniku, umjetnu inteligenciju, prividnu stvarnost i upravljanje prometom. Član je Centra izvrsnosti za računalni vid te Znanstvenog centra izvrsnosti za znanost o podacima i napredne kooperativne sustave, Istraživačka jedinica znanost o podacima. Član je Povjerenstva za upravljanje kvalitetom, Odbora za znanost i projekte te Povjerenstva za nagrađivanje djelatnika i studenata Fakulteta prometnih znanosti. Njegov znanstveni interes uključuje inteligentne transportne sustave, modeliranje i simuliranje cestovnih mreža, primjena umjetne inteligencije i računalnog vida u upravljanju cestovnim prometom, procjena i predviđanje prometnih parametara te uključivanje umreženih autonomnih vozila u upravljačku petlju. Član je uredničkih odbora znanstvenih časopisa Journal of Traffic and Transportation Engineering te International Journal of Intelligent Transportation Systems Research. Član je stručnih udruga IEEE (Institute of Electrical and Electronics Engineers), IEEE Intelligent Transportation Systems Society, ITS Hrvatska i znanstvene mreže tvrtke PTV. Također je član suradnik Akademije tehničkih znanosti Hrvatske (HATZ). Nagrađen je nagradom Brončana plaketa Josip Lončar Fakulteta elektrotehnike i računarstva, Sveučilišta u Zagrebu kao najbolji diplomand smjera automatika u 2001. godini, nagradom Fakulteta prometnih znanosti, Sveučilišta u Zagrebu za posebna dostignuća u nastavnom radu u 2017. godini, nagradom tvrtke Siemens za postignuća u poboljšanju obrazovanja i predavanja o automatizaciji uvođenjem SIMATIC automatizacijskih sustava u 2018. godini, priznanjem „Plaketa na Edouard Ivanjko kako znak na blagodarnost za pridonosot vo razvojot na Fakultetot“ Univerziteta „Sv. Kliment Ohridski“ – Bitola Tehnički Fakultet, Bitola, Sjeverna Makedonija u 2021. godini te nagradom Fakulteta prometnih znanosti, Sveučilišta u Zagrebu za posebna dostignuća u znanstveno-istraživačkom radu u 2022. godini. Objavio je više od 100 članaka u međunarodnim znanstvenim časopisima i konferencijama, poglavlja u znanstvenim knjigama, proširenih sažetaka i tehničkih izvještaja. Dobro se služi engleskim i njemačkim jezikom.

Acknowledgements

First and foremost, I want to express my heartfelt gratitude to my supervisor, Professor Edouard Ivanjko. I can't emphasize enough the positive influence he has had on my academic journey. Without him, everything would be different. Scientific achievements in my academic career, including my PhD thesis, would be, I could say, hard to imagine without his unselfish guidance. From our initial collaboration on my bachelor's thesis to the present, I have experienced many challenges, but, as always, hard work pays off. Therefore, if I had to do the same again, I would do it again without hesitation. I also owe a debt of gratitude to my second PhD supervisor, Associate Professor Marko Šoštarić. His expert advice on traffic flow theory has been invaluable, ensuring that my dissertation remains closely aligned with the world of traffic.

A special thanks goes to Professor Ivana Dusparic, my supervisor during my mobilities at Trinity College Dublin. Along with a world of multi-agent learning systems, I also found time to explore Ireland and the warmth of the Irish people, enhancing my experiences with Irish whiskey and Guinness, which I'll cherish forever. Speaking of unforgettable experiences, my time in the charming Swiss town of Sierre, working alongside Professor René Schumann at HES-SO Valais-Wallis, was nothing short of amazing. Amid our exploration of digital twin technology, I also had the pleasure of sampling some delicious raclette cheese, sipping Valais wine, and even trying skiing in Crans-Montana. Looking at my scientific track record, it is evident how much I have learned from Ivana and René.

I am extremely grateful to my family who have provided me with unwavering support throughout my studies. My brother Krunoslav and sister Ivana, as well as my mother Snježana and my father Bruno, have been my pillars of strength during this period, and still are. I also want to extend a special thanks to my Teta Silva and Tetak Joginder, who have been irreplaceable sources of support and guidance throughout my life. I am thankful to all my family members, who have played an integral role in this journey.

Last but certainly not least, a warm and big thanks to my friends, and colleagues who have been there for me every step of the way during my studies.

Abstract

Continuously increasing demand for mobility, especially in road transport, and new transport policies that prioritize optimizing the capacity of existing roads over building new ones, pose major challenges in the field of traffic control towards a competitive and resource-efficient transport system. This thesis is primarily concerned with urban motorways and their characteristic problems, the presence of bottlenecks. Variable speed limit (VSL) as an effective traffic control strategy is used to optimize traffic on motorways. Due to the stochastic nature of traffic congestion, static VSL zones, i.e., the VSL application area remains unchanged, may not always respond accurately to spatiotemporal spreading congestion. Accordingly, a model for dynamic spatiotemporal adaptation of VSL zones that complements the prevailing VSL with static zones is proposed. To address the control problem for the proposed dynamic VSL zones, a novel distributed multi-agent reinforcement learning VSL (DWL-ST-VSL) controller, capable of dynamically adapting the length and position of VSL zones during VSL operation, is suggested. To ensure rigorous evaluation of the controller, a novel paradigm in motorway traffic modeling is proposed and demonstrated by using the continuously synchronized digital twin (DT) model of the Geneva motorway (DT-GM). Thus, DWL-ST-VSL applicability is verified through a virtual field test, which uses DT-GM microscopic simulation of actual traffic received in real-time from the Geneva motorway. Such functional integration presents a novelty in evaluating learning-based traffic controllers using traffic scenarios that more closely reflect the uncertainty in real traffic. Accordingly, DT-GM advances technological capabilities of traffic simulations to a more sophisticated and accurate level by enabling simulation-based control optimization during system run-time. Furthermore, DWL-ST-VSL outperforms all baseline control approaches with static VSL zones. Different VSL zones configurations per traffic scenarios are, thus, learned without requiring a manual setup. This is of practical relevance as it can guide traffic engineers in optimizing VSL zones design in complex motorway environments. Thus, demonstrated functional integration of the DWL-ST-VSL with the DT technology provides the basis for further development of predictive real-time analytics of complex autonomous control processes as a step toward safety-critical decision-making in a nonstationary traffic setting and provides an experimental environment for developing control system prototypes in the context of emerging connected vehicles technologies.

Keywords: traffic control; adaptive speed limit zones; cooperative multi-agent systems; reinforcement learning; smart motorways; digital twins; microscopic simulations

Prošireni sažetak

Povećanje operativnoga kapaciteta autocesta dinamičkom prilagodbom zona promjenjivoga ograničenja brzine zasnovanom na kooperativnim učecim agentima

Uvod

Stalno rastuća potražnja za mobilnošću, posebno u cestovnom prometu, i nove prometne politike koje daju prednost optimizaciji kapaciteta postojećih cesta u odnosu na izgradnju novih, predstavljaju veliki izazov u području upravljanja prometnim tokovima prema konkurentnom i resursno učinkovitom prometnom sustavu. Ovaj doktorski rad prvenstveno se bavi gradskim autocestama i njihovim karakterističnim problemima. Najistaknutiji problem je pojava uskih grla koja uzrokuju prekid prometa. Stoga je glavni cilj aktivnog upravljanja prometom spriječiti takve događaje, odnosno spriječiti aktiviranje uskog grla ako je moguće. Upotrebom promjenjivog ograničenja brzine (eng. Variable Speed Limit, VSL) glavni tok prometa može biti podvrgnut nižim ograničenjima brzine, i na taj način se protok iz područja obuhvaćenog VSL djelovanjem može privremeno ili čak trajno ograničiti. Dakle, postavljanjem odgovarajućih ograničenja brzine volumen prometnog toka koji se kreće u smjeru uskog grla može se prilagoditi tako da ostane unutar vrijednosti ograničenog (narušenog) kapaciteta uskog grla. Na taj način se sprječava preopterećenje uskog grla.

Motivacija

Međutim, postoji nekoliko zahtjeva za dizajn VSL-a koji se moraju odgovarajuće riješiti kako bi se osiguralo učinkovito upravljanje prometnim tokom na autocesti. Na primjer, mora se odabrati odgovarajuća strategija upravljanja na osnovu koje će se izračunavati ograničenja brzine ovisno o stanju prometa. Osim toga, važno je odrediti odgovarajuće lokacije za postavljanje znakova s promjenjivim porukama (eng. Variable Message Signs, VMS) putem kojih se informiraju vozači o ograničenjima brzine. Oni bi trebali biti strateški postavljeni kako bi se postigao željeni učinak upravljanja prometnim tokom na određenim segmentima autoceste, uzimajući u obzir ograničenja kao što su sigurnost, učinkovitost, troškovi itd. Dizajneru VSL sustava je na ta pitanja teško odgovoriti budući da efiksnost sustava ovisi o mnogim vanjskim i unutarnjim varijablama. Na primjer, poznato je da učinkovitost VSL-a ovisi o duljini područja primjene VSL-a i njegovoj udaljenosti od područja uskog grla.

Kako bi se opravdala potreba za razvojem modela dinamičkom prilagodbom zona promjen-

jivoga ograničenja brzine, u ovom doktorskom radu iznosi se nova pretpostavka o tome zašto je pristup VSL-a s promjenjivim zonama vrijedan daljnjeg istraživanja. Konkretno, koriste se teorijske spoznaje iz teorije prometnog toka iz kojih slijedi da se kapacitet autoceste mijenja u ovisnosti o položaju i vremenu. Prvo, na njega utječe geometrija prometnice. Drugo, na njega utječu latentne varijable kao što su vremenski uvjeti ili ponašanje samih vozača, itd. Potonje dovodi do nesigurnosti u kapacitetu definiranom kao funkcija vjerojatnosti nastanka prekida u normalnom odvijanju prometa čija vjerojatnost nastanka raste s povećanjem volumena prometa. Stoga se ne može pretpostaviti da je kapacitet (i pad kapaciteta) uskog grla konstantna varijabla, kao ni kapacitet prometnice prije mjesta zagušenja.

To znači da dodatno akumuliranje vozila (manifestirano povećanjem gustoće prometa) koje stvara VSL zona mora biti pažljivo postavljeno, uzimajući u obzir ograničeni kapacitet neposredno prije područja uskog grla, kako se ne bi stvorio novi prekid prometa uzvodno od uskog grla zbog samog djelovanja VSL-a. Iznesena činjenica implicira da VSL sa statičkom zonom možda neće moći prihvatiti dodatni promet u određenom vremenskom periodu kako bi se rasteretilo ili spriječila aktivacija uskog grla. Štoviše, VSL sa statičkim zonama može stvoriti ili ubrzati stvaranje dodatnih uskih grla uzvodno. Stoga ima smisla dinamički proširiti ili smanjiti, preraspodijeliti ili aktivirati dodatne VSL zone kako bi se zadovoljile trenutne potrebe prostorno i vremenski ograničenog kapaciteta kako bi se odgovarajuće odgovorilo na uska grla, a da se ne naruši cjelokupna dinamika prometa.

Pregled disertacije

Kako bi se potvrdila gore navedena hipoteza, provodi se dubinska analiza koja uključuje temeljito ispitivanje postojećih istraživanja na temu optimalnih lokacija za VSL zone. U drugoj fazi provodi se analiza zasnovana na simulaciji s različitim unaprijed definiranim konfiguracijama VSL zona u odnosu na lokaciju uskog grla kako bi se dobila indikacija korelacije između dizajna VSL zone i performansi prometa. Osim toga, provodi se višekriterijska analiza za određivanje optimalne VSL zone uzimajući u obzir performanse uskog grla zajedno s performansama cijele simulirane mreže autoceste kao kriterij. Na osnovu pruženih dokaza zasnovanih na simulacijama o podoptimalnom radu VSL sustava s fiksnim VSL zonama u okruženju s promjenjivim prometnim zagušenjem opravdava potrebu za razvojem VSL-a s dinamičkim zonama. Koristeći rezultate i spoznaje drugih istraživača, verificirana je potreba za adaptivnim VSL zonama, te je stoga predložen model dinamičke prostorno-vremenske prilagodbe VSL zona koji nadopunjuje do sada prevladavajući VSL sa statičkim zonama.

Kako bi se riješio problem upravljanja za predloženu dinamičku prostorno-vremensku prilagodbu VSL zona, predlaže se novi distribuirani prostorno-vremenski višeagentni VSL (DWL-ST-VSL) regulator koji je sposoban uz prilagodbu ograničenja brzine dinamički prilagođavati

duljinu i položaj VSL zona tijekom rada VSL-a. Kako bi se nadopunila prilagodba ograničenja brzine u trenutnim VSL upravljačkim sustavima, DWL-ST-VSL je modeliran kao algoritam distribuiranog W-učenja (DWL) koji je u osnovi algoritam strojnog učenja zvan potporno učenje (eng. Reinforcement Learning, RL). Stoga DWL kao tehnika strojnog učenja za samooptimizaciju zasnovana na suradničkim učećim agentima omogućuje učenje tj. optimiranje sustava primjenom višestrukih nehomogenih upravljačkih politika. Uz svoje lokalne politike, kroz koncept udaljenih politika predloženih u DWL algoritmu, agenti također uče kako njihovi postupci (odabir akcija) utječu na neposredne susjedne agente i koja politika ili radnja se preferiraju u određenoj situaciji. Međutim, prije bilo kakve implementacije sustava upravljanja prometom (posebice onog koji se zasniva na učenju, a ima sposobnost prilagoditi svoje ponašanje tijekom vremena rada na ne nužno predviđajući način), potrebno je osigurati njegovu rigoroznu evaluaciju. Stoga je potrebno razviti odgovarajući model fizičkog sustava (procesa) koji omogućuje logičku manipulaciju kako bi se odgovorilo na pitanja kako se ponaša sustav upravljanja i kako se sam proces ponaša pod sustavom upravljanja, tj. istražiti međusobni utjecaj jednog na drugog.

U prometnom inženjerstvu model procesa prometnog sustava često se odgovarajuće nadomješta pomoću simulacijskih modela. Pojava široko rasprostranjene informacijske tehnologije, posebice dostupnost podataka o prometu u stvarnom vremenu, daje temelj za dopunu prevladavajućih (izvanmrežnih) pristupa mikroskopskoj simulaciji stvarnim podacima kako bi se stvorio detaljan digitalni prikaz fizičkog prometa u stvarnom vremenu. Međutim, upotreba stvarnih prometnih podataka u analizi autocesta u stvarnom vremenu još nije istražena. Razlog je to što ne postoje prateći modeli za primjenjivost podataka u stvarnom vremenu u kontekstu mikroskopskih simulacija te primjena takvih simulacijskih modela nije još razmatrana i prepoznata u kontekstu simulacija prometa na autocestama. Stoga je u ovom doktorskom radu fokus na mikroskopskoj simulaciji autoceste s integracijom prometnih podataka u stvarnom vremenu tijekom rada sustava. Kao rezultat toga, predložena je i demonstrirana nova paradigma u modeliranju prometa na autocesti korištenjem kontinuirano sinkroniziranog digitalnog blizanca modela ženevske autoceste (DT-GM). Analizirana je primjena mikroskopskog simulatora SUMO u modeliranju, simulaciji i umjeravanju u hodu sinkroniziranih digitalnih replika stvarnog prometa korištenjem tokova detaljnih stvarnih podataka o prometu s brojača prometa na autocestama kao ulaznih podataka za DT-GM. Na taj način se stvarni podatci o prometu izravno integriraju u DT-GM koji se kontinuirano izvodi tj. kontinuirano se umjerava kako se fizički pandan mijenja. Uz minutnu rezoluciju simulirani promet se vrlo približava stvarnoj distribuciji prometa na analiziranom segmentu autoceste. S kontinuiranim ažuriranjem novim dostupnim podacima o prometu i kontinuiranim procesom umjeravanja, DT-GM stvara umjerene simulacijske okvire (gotovo točne replike stvarnog prometa) u stvarnom vremenu tijekom životnog ciklusa simulacije. U skladu s tim, DT-GM se koristi za procjenu ponašanja i performansi DWL-ST-VSL u

fazama dizajna, učenja i evaluacije.

Dakle, primjenjivost DWL-ST-VSL provjerena je na (virtualnom) terenskom testu korištenjem 24-satne DT-GM simulacije stvarnog prometa generiranog u stvarnom vremenu na ženevskoj autocesti u Švicarskoj. Time ovaj pristup predstavlja novost u procjeni regulatora zasnovanih na RL-u korištenjem scenarija koji mnogo bliže odražavaju neizvjesnost u stvarnom prometu u usporedbi s konvencijlnim (offline) simulacijama. Dakle, DT-GM se može precizno umjeriti u stvarnom vremenu korištenjem stvarnih podataka o prometu s osjetila autoceste. Sukladno tome, DT-GM podiže tehnološku dimenziju u mikroskopskoj simulaciji prometa na autocesti na sljedeću razinu omogućavajući simulacijski zasnovanu optimizaciju upravljanja prometom tijekom rada fizičkog sustava koja je prije bila neizvediva. Drugo, pokazalo se da predloženi model dinamičkih prilagodljivih VSL zona nadmašuje sve testirane pristupe upravljanja VSL-om sa statičkim VSL zonama u simulacijskim eksperimentima na sintetičkim modelima te na realističnom scenariju prometa generiranim pomoću DT-GM.

Zaključak

Zaključno, provedeni eksperimenti daju uvid u novi koncept upravljanja VSL-a. Konkretno, primjena DT tehnologija u mikrosimulacijama bi mogla biti okidač za daljnja istraživanja o korištenju naprednih digitalnih tehnologija zasnovanih na učenju za razvoj nove generacije prilagodljivih sustava upravljanja prometom koji ispunjavaju zahtjeve rada u nestacionarnom okruženju u kontekstu nadolazećih novih tehnologija i tehnologija povezanih vozila općenito. Prikazani rezultati su stoga od praktične važnosti jer mogu poslužiti kao vodilja inženjerima prilikom dizajniranja autocesta te nadležnim tijelima kao nit vodilja u implementaciji i opremanju autocesta naprednim upravljačkim sustavima. Integracija regulatora prometa zasnovanih na učećim algoritmima u stvarnom vremenu s tehnologijama digitalnih blizanaca čini temelj za evoluciju i mogućnost prediktivne analitike složenih upravljačkih procesa u stvarnom vremenu kao korak prema sigurnijem donošenju odluka kritičnih za sigurnost prometa u upravljanju prometom općenito.

Ključne riječi: upravljanje prometom; prilagodljive zone ograničenja brzine; kooperativni višeagentni sustavi; potporno učenje; pametne autoceste; digitalni blizanci; mikroskopske simulacije

Contents

1. Introduction	1
2. Variable speed limit control on motorways	10
2.1. Motivation for VSL	10
2.1.1. Relationships among fundamental traffic flow parameters	10
2.1.2. Traffic flow and motorway capacity	12
2.1.3. Creation of congestion on motorways	13
2.1.4. Concept of mainstream traffic flow control	15
2.2. Effect of VSL on motorway traffic flow	15
2.2.1. Speed harmonization effect and traffic safety	16
2.2.2. Impact of VSL on exhaust emissions and energy consumption	16
2.2.3. Impact on motorway throughput	16
2.3. VSL control strategies	21
2.3.1. Classical VSL controllers	21
2.3.2. Learning-based VSL controllers	22
2.4. Aspects of VLS design	24
2.4.1. Spatial aspects of VSL design	24
2.4.2. Temporal aspect of VSL design	27
2.4.3. Additional information on VMS	28
2.5. Concluding remarks	28
3. Digital twin technology on motorways	31
3.1. Current approaches to traffic simulation	32
3.2. Digital twin technology in microscopic traffic simulations	33
3.3. Digital twin motorway concept	38
3.3.1. Physical twin - Geneva motorway	41
3.3.2. Digital twin of Geneva motorway	42
3.3.3. Concept of parallel DTI-GM in TM	43
3.4. Simulation synchronization with run-time motorway traffic	45

3.4.1.	Dynamic simulation calibration using actual traffic data streams	45
3.4.2.	Experimental setup	52
3.4.3.	Results and analysis	55
3.4.4.	Discussion	61
3.5.	Concluding remarks	64
4.	Dynamic spatiotemporal adaptation of variable speed limit zones	66
4.1.	Optimizing VSL zones allocation	66
4.1.1.	Case with feedback VSL controller	67
4.1.2.	Multi-criteria analysis for optimal VSL zones allocation	76
4.2.	Nondeterministic capacity behavior and VSL zones allocation	80
4.3.	Model for dynamic spatiotemporal VSL zones adaptation	83
4.4.	Concluding remarks	84
5.	Reinforcement learning based VSL	86
5.1.	Reinforcement learning	86
5.1.1.	Q-learning algorithm	88
5.1.2.	Function approximation for RL	88
5.2.	Modeling VSL as RL problem	89
5.2.1.	Variable speed limit control as an MDP	89
5.2.2.	Simulation analysis	96
5.3.	Concluding remarks	98
6.	Cooperative multi-agent learning VSL control system	100
6.1.	Multi-agent based RL	100
6.1.1.	W-learning	101
6.1.2.	Distributed W-learning	101
6.2.	Modeling VSL as a W-learning problem	102
6.2.1.	Agents environment formulation	102
6.2.2.	Simulation setup	105
6.2.3.	Simulation model	105
6.2.4.	Results and analysis	107
6.2.5.	Discussion	112
6.3.	Modeling spatiotemporal VSL as distributed W-learning problem	113
6.3.1.	Configuration with two agents	113
6.3.2.	Configuration with four agents	117
6.3.3.	Simulation setup	123
6.3.4.	Simulation results	126

6.3.5. Discussion	135
6.4. Concluding remarks	137
7. Testing VSL using the Geneva motorway digital twin	139
7.1. Traffic and geometry characteristics of DT-DWL-ST-VSL experiment	139
7.2. Modeling DT-GM as DWL-ST-VSL control problem	142
7.2.1. State definition	142
7.2.2. Action space	144
7.2.3. Reward function	145
7.3. Functional integration of DT-GM and DWL-ST-VSL	146
7.4. Experimental setting	147
7.5. Results and analysis	149
7.5.1. Learning performances of DWL-ST-VSL controller	149
7.5.2. Spatiotemporal analysis of dynamic VSL zones adaptation	151
7.5.3. Level of cooperation analysis	154
7.6. Discussion	155
7.7. Concluding remarks	156
8. Conclusion and future work	157
Bibliography	162
List of Figures	174
List of Tables	177
List of Algorithms	178
Nomenclature	179
Biography	182
Životopis	186

Chapter 1

Introduction

Daily commuting in densely populated urban areas is accompanied by repeated traffic congestion, which significantly affects the quality of life in cities and towns. The traditional approach to solving the congestion problem is to expand the capacity of the existing transportation network by building new lanes or even new roads. Especially in metropolitan areas, this can also be done by building urban motorways. The main objective of these motorways is to connect different districts of large built-up areas (megacities) and provide a high level of service (LoS) to travelers. However, as an integral part of the urban road network, motorways are also prone to daily congestion. In addition, the new green and sustainable transport policy no longer promotes the construction of new capacity, but rather a shift to other modes of transport and the optimization of existing capacity with new methods of traffic flow control. Therefore, traffic control approaches from the domain of intelligent transportation systems (ITS) are now being applied to the existing infrastructure to increase the operational capacity of such urban motorways.

The applied traffic control approaches include ramp metering (RM), variable speed limit (VSL), and lane change (LC) control. This thesis focuses on VSL as an efficient traffic control strategy to improve urban motorways' LoS. VSL controls the speed limit in real-time by displaying an appropriate speed limit on variable message signs (VMS) placed across the motorway. Since VSL is a (traffic and weather)-responsive control system, the value of the speed limit is adjusted according to, for example, the current traffic state (traffic load), environmental conditions such as weather, speed regulation in work zones on motorways and other [1]. The main objective of VSL is to improve traffic safety and throughput on motorways through the concept of traffic flow harmonization [2]. In addition, VSL is used for mainstream traffic flow control (MTFC) [3] in case of recurrent bottlenecks which are typical of motorways' merge areas. For example, increased traffic volume at the on-ramp (motorway entrance) can disrupt the mainstream flow and cause congestion. Thus, MTFC applied upstream of the bottleneck can suppress bottleneck

activation or its expansion. To summarize, VSL, thus, has a dual effect on traffic flow: it harmonizes it and increases throughput by reducing the risk of capacity breakdown in areas of potential bottlenecks.

However, despite the fact that VSL has been extensively researched and is being used on motorways [4, 5] (for more examples see [6]), questions remain about its optimal design. Needless to say, speed limit values and their duration are crucial factors for VSL performance. Therefore, over the years, various VSL approaches for computing speed limits have been suggested based on different control system configurations and methodologies, such as rule-based VSL activated by predefined threshold values (e.g., flow, speed or density) [7, 8], the usage of metaheuristics to optimize VSL [9], optimal control [10], machine learning based approaches [11], and model-predictive control [12]. The most prominent VSL design (among classical controllers) uses feedback control [3, 13], where the speed limit is calculated based on current measurements of traffic states, such as traffic volume, density, and traffic flow speed. However, such technologies are typically non-adaptive and, thus, may have limitations while controlling stochastic traffic on urban motorways.

Apparently, in recent years there has been a growing interest in advancing VSL design by incorporating new available technologies with an emphasis on the application of machine learning techniques, especially reinforcement learning (RL). An overview of the existing literature can be found in [11]. RL has a proven track record of solving various complex control problems, including transportation and related control optimization problems with considerable improvements in transportation management efficiency [14, 15, 16, 17]. In particular, RL enables the resolution of complex Markov decision processes (MDPs) and the discovery of a near-optimal solution for discrete-event stochastic systems without requiring an analytical model of the system to be controlled [18]. In addition, RL-based control systems can continuously improve their performance over time by adapting control policies to newly recognized states of the environment. RL, consequently, provides a model-free, adaptive control alternative for controlling complex systems, making it a suitable technique for optimizing traffic control processes on motorways.

The results of the pioneers in developing a single agent-based RL controller for VSL (RL-VSL) prove the potential of this technology in controlling traffic flow on motorways. In the available research, an agent is designed to minimize or maximize the objective function [19, 20] or to pursue multiple objectives [21, 22, 23] regarding the traffic parameters by learning the optimal control policy. However, in the case of large-scale traffic control, the control optimization often seeks to optimize conflicting objectives which are heterogeneous in time and space scales, e.g., simultaneous optimization of RM and VSL on motorways [24], or different levels of priorities in objectives, e.g., giving more weight to safety contrary to throughput [25] or favoring throughput

contrary to traveling speeds [26], etc. In practice, VSL is usually applied on a longer motorway segment. For this reason, the VSL application area should be divided into several consecutive VSL sections. In this way, VSL ensures a smooth (gradual) spatial adjustment of vehicle speeds in traffic flow. However, a single-agent RL approach often encounters an exponential increase in computational cost as the scope of the controlled process is expanded.

In [24], a multi-agent RL (MARL) technique is used to control VSL and RM simultaneously in order to solve motorway congestion. The most sophisticated MARL technique used in research [24] was the maximax MARL approach. Due to the nature of this algorithm, it is likely that the extended applicability containing more agents might be limited due to two-way communication in which the state-action space of each agent increases by a factor equal to the number of actions available to its neighbor. In [25], a multi-agent VSL was tested with two goals. Traffic safety policy reduces the speed difference between adjacent controlled motorway sections, while mobility control policy increases throughput in the bottleneck. Although the authors highlight two conflicting objectives in their analysis, the optimization process was based on the separate use of individual policies i.e., each of them was learned independently (not simultaneously) by a distributed Q-learning algorithm. According to the given objective, three agents for VSL must learn an optimal joint policy. In the used algorithm, the control state is fully observed which is considered more accurate than a partial observation algorithm. On the other hand, it consumes more computational resources. As a result, it does not require additional communication between agents, but it seems ideal for medium-scale control tasks.

Nonetheless, MARL generally has been proven as a technique that makes it possible to use a powerful RL capability for solving complex large-scale control tasks, while still being simple to implement [27]. Thus, it delivers a promising way to further extend RL application to track more complex problems on motorways. However, the potential of MARL in the motorway domain has not yet been realized. In general, there is very little research on the use of MARL technology in VSL control and literally no research on the use of cooperative agent-based techniques. This raises a few practical questions. (i) Is it and to what extent possible to benefit from cooperative MARL techniques in VSL in an uncertain dynamic environment on the motorway? (ii) What kind of control design does such a system allow? For example, does it allow for heterogeneity in agents' policies to address multiple, possibly conflicting objectives with different spatial and temporal scopes simultaneously? (iii) Is it possible to obtain multiple outputs from such a control system at the same time (e.g., speed limit value and VSL zone configuration)?

There are also other factors that need to be considered in the design phase of the VSL system. Research conducted in [28, 29] have analyzed the impact of VSL zone location (VSL application area) on VSL performance regarding bottleneck optimization. An in-depth analysis of their findings will be provided in chapter 2, but it is clear from the results that the efficiency of VSL

depends on the spatial design of VSL zones. Particularly, two practical questions arise: how long should the VSL application area be and where should it be placed? In other words, where should the VSL zone start and how far should the end of the VSL zone be from the bottleneck to achieve good performance of the VSL system? In general, it can be concluded that different lengths and positions of VSL application areas, applied speed limits, and spatial location of congestion creation significantly affect VSL performance. Unfortunately, current studies do not provide a general methodology for selecting or designing a suitable VSL zone. Furthermore, they are partially inconsistent which will be addressed in chapters 2 and 4.

Thus, the designer of a system must choose a VSL zone configuration well suited for the situation at hand, which raises two questions. (i) Can the process of VSL zone setup be automated? (ii) Is it a disadvantage that the prevailing VSL systems use only static VSL zones, and if so, how much benefits can VSL gain from dynamically adapting the length and position of VSL zones (particularly when congestion characteristics and motorway capacity vary widely over space and time) along with speed limits during VSL operation?

In general, when developing and designing new traffic control systems, their simulation-based evaluation is important before they are used to control a real process [28]. It is important to utilize a simulation model that accurately reflects the behavior of the real system. To accomplish this, a precise traffic simulation model must be modeled, which is capable of accurately simulating various traffic scenarios that may occur in the real system. Particularly, each RL optimization process heavily relies on simulations. Therefore, it is important, for RL-VSL (or MARL-VSL in general) to ensure a wide spectrum of training examples so that, controllers once deployed in a real system, can ensure good behavior, namely generalization across all feasible traffic states. Traditional (offline) simulation approaches, however, use historical aggregated traffic data (or artificially created) for generating simulation scenarios representing only a few selected situations that depend primarily on the input parameters of the simulation [30]. In general, such models are inadequate to reflect unexpected events that are likely to occur in reality and cannot be captured by offline simulated results due to their inherent characteristics.

This leads to the following questions. (i) Is it possible to connect microscopic traffic simulation with actual fine-grained traffic data streams received directly from traffic detectors on real urban motorways and, so to speak, make the simulation rely more on actual data and less on domain knowledge in terms of simulation parameters? (ii) Is there a way to calibrate such a simulation dynamically to reflect real-time traffic changes? (iii) Can such a virtual instance (continuously synchronized with the corresponding real urban motorway) be used to adequately capture, predict, and analyze real-time traffic behavior through simulations? (iv) If so, what advantages can such a digital twin (DT) simulation have in VSL design and motorway traffic engineering in general over predominant conventional offline simulations?

In addition to the previously mentioned findings, this thesis presents three main hypotheses that can address the gap in current research and provide answers to the aforementioned questions in both the development and simulation verification of VSL:

- By adjusting the settings of the microscopic road traffic simulation based on the current measurements of traffic variables, it is possible to calibrate the simulation model in real time during its execution;
- Dynamic adaptation of speed limit zones can improve traffic flow control on motorways;
- By applying a cooperative multi-agent machine learning approach, it is possible to simultaneously learn the control law for variable speed limits and for the spatiotemporal allocation of variable speed limit zones.

Therefore, this thesis focuses on the application of DT technology in microscopic traffic simulations and ways in which such a model, calibrated in real time with actual traffic data, can be used for learning and testing learning-based traffic control systems. Next, the focus is on developing a model for dynamic spatiotemporal adaptation of variable speed limit zones. Finally, to control dynamic spatiotemporal adaptation of VSL zones, a cooperative multi-agent learning control system based on distributed reinforcement learning will be developed. Therefore, a research perspective mainly from the fields of traffic simulation (emphasizing DT technology), traffic control, and artificial intelligence is adopted.

The advent of widespread information technology, particularly the availability of real-time traffic data, provides the foundation for supplementing predominated (offline) microscopic simulation approaches with actual data to create a detailed real-time digital representation of physical traffic. However, the use of actual traffic data in real-time motorway analysis has not yet been explored and the applicability of real-time data in the context of microscopic simulations has yet to be recognized. Thus, this thesis focuses on microscopic motorway simulation with real-time data integration during system run-time. As a result, a novel paradigm in motorway traffic modeling is proposed and demonstrated by using the continuously synchronized DT microscopic simulation model of the Geneva motorway (DT-GM). The application of the microscopic simulator Simulation of Urban MObility (SUMO) [31] in modeling and simulating dynamically synchronized digital replicas of real traffic by leveraging fine-grained actual traffic data streams from motorway traffic counters as input to DT-GM is analyzed. Thus, the detailed methodological process of developing DT-GM is presented, highlighting the calibration features of SUMO that enable (dynamic) continuous calibration of running simulation scenarios. By doing so, the actual traffic data are directly fused into the running DT-GM every minute so that DT-GM is continuously calibrated as the physical counterpart changes.

Also, to overcome the aforementioned issues in VSL with static VSL zones, the adaptive design

of VSL zones is desirable, especially in traffic scenarios where congestion characteristics vary widely in space and time. Thus, even if VSL zones are determined optimally, under changing traffic conditions VSL systems with static VSL zones may perform suboptimally. Therefore, in this thesis, an adaptive design of VSL zones that complements the prevailing VSL systems used so far with static zones is proposed. To confirm this hypothesis, an in-depth analysis that includes a thorough examination of existing research on the topic of optimal locations for VSL zones is conducted. In the second phase, an optimization process using a standard feedback VSL controller and various predefined configurations of VSL zones to obtain an indication of the correlation between VSL zone design and traffic performance is conducted. In addition, a multi-criteria analysis for an optimal VSL zone that includes a finer range of VSL zone configurations and considers the performance of both the bottleneck and the entire network as criteria is performed. Finally, based on the results, a model for dynamic spatiotemporal adaptation of VSL zones is proposed. Additionally, to justify the need for adaptive VSL zones (proven by simulation), a new conjecture as to why such an approach is worth further investigation is theorized. In particular, the theoretical continuum assumptions about traffic flow which state that the traffic flow property varies continuously and smoothly in space and time across the motorway [32] are applied. Consequently, motorway capacity is assumed to be constrained as a function of position and time. Moreover, the concept of probabilistic capacity [33] is also discussed in this context. This implies that VSL with a static VSL zone may not be able to accommodate additional traffic from time to time to relieve the bottleneck. Moreover, it can create or speed up additional bottlenecks upstream. Therefore, it makes sense to dynamically expand or shrink, rearrange or activate additional VSL zones to meet the current needs of spatially and temporally constrained motorway capacity in order to adequately respond to uncertainties in bottleneck capacity.

To address the control problem for the proposed dynamic spatiotemporal adaptation of VSL zones, a novel distributed spatial-temporal multi-agent VSL (DWL-ST-VSL) approach capable of dynamically adapting (during the system run-time) the length and position of active VSL zones to complement the adjustment of speed limits in current VSL control systems is proposed. To set up multiple adjustable VSL zones and simultaneously adjust the speed limit values, the DWL-ST-VSL controller uses a distributed W-learning (DWL) algorithm [27]. DWL has the property of collaborative agent-based self-optimization towards the simultaneous deployment of multiple heterogeneous policies. Each agent uses the RL technique to learn local policies, thereby maximizing travel speed and, if necessary, eliminating congestion. In addition to local policies, through the concept of remote policies (originally proposed in [27]), agents learn how their actions affect their immediate neighbors and which policy or action is preferred in a given situation. Thereby, the use of remote policies enables cooperation between agents and can be viewed as a link between them that allows an agent to help its neighbors implement their poli-

cies. Thus, DWL relies only on local learning and interactions. Consequently, no information about the joint state-action space is shared between agents, which means that the complexity of the model does not increase exponentially with the number of agents.

Therefore, this thesis has the following scientific contributions:

- It proposes a method for motorway digital twin creation with the possibility of its calibration in real time for learning and testing of traffic control systems;
- Proposes a model for dynamic spatiotemporal adaptation of variable speed limit zones for increasing operational capacity of motorways;
- Proposes a cooperative multi-agent learning control system for spatiotemporal adaptation of variable speed limit zones based on distributed reinforcement learning.

Generally, it can be hypothesized that the cooperative learning agents may contribute to the ability to dynamically configure VSL zones. The fact that agents can collaborate using remote policies may result in a better response to moving congestion because they can collectively assemble a larger number of feasible VSL application areas. A certain number of configurations can more appropriately respond to the current formation and expansion of the downstream bottleneck. Thus, it can be anticipated that a DWL-ST-VSL system with more agents could use its additional adaptive feature to adjust VSL zones to resolve congestion as much as possible without suppressing the upstream traffic itself. As a result, a further reduction in the overall travel time of the system, and a smoother speed transition by spatially deploying multiple VSL agents can be expected. This aligns closely with the objectives of VSL implementation, which are to achieve harmonized traffic flow. Using an adjustable VSL application area supported by multiple dynamically configurable VSL zones reduces the need for a severe decrease in speed limit values. Agents in upstream zones can prepare vehicles for conditions in downstream VSL zones by slightly decreasing speed limits. This is necessary since speed limits in downstream zones may be lower due to the proximity of the bottleneck. Therefore, this can help to harmonize traffic flow in order to avoid undesirable effects, such as shockwaves caused by sudden deceleration of vehicles. Particularly, this is possible due to the fact that DWL-ST-VSL can implement heterogeneous action sets with a different granularity of speed limits on the observed motorway section. DWL-ST-VSL, thus, enables automatic, systematic learning in setting up the sufficiently accurate VSL zone configuration (selection is learned rather than manually designed) for efficient VSL operation under a fluctuating traffic load and spatiotemporal distribution of congestion. From a technical perspective, the physical VMS may soon be replaced (or enhanced) by advanced technologies related to connected vehicles. Therefore, the proposed approach has the potential to enhance existing VSLs that encompass the current infrastructure design by addressing VSL placement constraints. It can also give forward mo-

mentum in the transition to the use of newly available technologies in the automotive industry. For example, vehicle-to-infrastructure communications, such as an intelligent speed assistance (ISA) system [34] can be used instead of a physical VMS which assumes that vehicles receive speed limit information directly on the driver dashboard [35]. Therefore, this research presumes that the static placement of physical VMSs is no longer an obstacle to the dynamic adaptation of VSL zone configurations in real motorway applications. However, nowadays, DWL-ST-VSL still requires closely spaced VMSs to ensure a virtual VSL zone where controllers can configure various active VSL subzones and, by doing so, allow for a dynamic adaptation of the VSL application area.

To assess the impact of deploying additional agents in the control loop and different cooperation levels of the control process, DWL-ST-VSL is evaluated in configurations with two, three, and four agents. DWL also allows for heterogeneity in agents' policies; cooperating agents in DWL-ST-VSL implement two speed limit sets with different granularity. Thus, an experimental approach is used to verify suggested solutions using simulation experiments. The performed experiment gives data-based evidence about the potential usefulness of DWL-ST-VSL control with adaptive VSL zones when deployed on longer motorway segments, particularly on a real part of the Geneva motorway. Results and analysis provide insights into the modeling of DWL-ST-VSL and the impact of agents' collaboration on system performance. This is a crucial aspect for the development of adaptive controllers in particular, but also for research investigating reliable and more efficient MARL-based traffic controllers.

The exploration of DT technology in transportation focuses on the design and verification of learning-based VSL controllers, as well as its application in real-time traffic management (TM) more broadly. DT-GM can be accurately calibrated in real time using the actual traffic data from the motorway sensors. It, thus, provides the foundation for the further development of real-time predictive analytics to support safety-critical decision-making in TM. Additionally, DT-GM raises the technology dimension in motorway traffic simulation to the next level by enabling simulation-based control optimization during system run-time that was previously unfeasible. Simulation results provide a solid basis for the future real-time analysis of learning-based traffic controllers on an extended Swiss motorway network. Additionally, DWL-ST-VSL is trained and evaluated through a real field test within a virtual safe environment using the proposed DT-GM which reflects actual traffic on the simulated motorway in Switzerland in real time.

Finally, the results obtained by DWL-ST-VSL outperform all baselines: W-learning-based VSL and negative feedback loop-based controllers that are using static VSL zones. Thus, performed experiments yield insights into the new concept of VSL control based on cooperative learning agents where no central component is needed. This could serve as a trigger for further research into the use of advanced learning-based technologies to develop a new generation of adaptive

traffic control systems that meet the requirements of operating in a nonstationary environment. Particularly at the forefront of emerging vehicle-to-everything communication technologies, where vehicles are generally considered to be mobile sensors and actuators themselves. As such, DWL-ST-VSL is envisioned as a new approach to dynamically adjusting speed limits in space and time, anticipating the practical aspect of vehicle speed control that may be found in the leading edge of vehicle-to-everything communication and Connected Vehicles environment in general. The results presented are, thus, of practical relevance as they can guide motorway engineers and authorities in the implementation of advanced VSL control strategies as well as in optimal VSL zone design. Finally, this study provides data-based evidence for the need for adaptive VSL zones and, therefore, may stimulate and trigger further research to systematically evaluate the theoretical findings presented so far on the optimal placement of VSL zones. Additionally, to justify the need for adaptive VSL zones, a new conjecture as to why such an approach is worth further investigation is proposed. To further promote the use of the developed VSL controller and DT-GM, and to serve as a basis for further investigation by the community, the DT-GM model and source code are publicly available.*

The remainder of this thesis is organized as follows: Chapter II discusses in detail the concept of VSL. Chapter III describes the need for good simulation approaches, including the newly created method for producing a digital twin of a motorway. In Chapter IV, the concept of dynamic spatiotemporal VSL zone placement is explained. Then, Chapter V elaborates on applying RL to create a single-agent based VSL. The details of the DWL multi-agent based VSL are discussed in Chapter VI. The description of the final simulation set up and simulation based evaluation of the implemented VSL controllers using the created digital twin of the Geneva motorway is given in Chapter VII. Chapter VIII concludes this thesis and comments on further research on the topic.

*https://github.com/kkusic/Digital_Twins_Based_Traffic_Control

Chapter 2

Variable speed limit control on motorways

VSL as a traffic-responsive control strategy increases the LoS of motorways by adjusting the speed limit on motorway sections according to the prevailing traffic conditions. The speed limit is posted on the VMS located on a certain section of the motorway to inform drivers of the permitted speed. Usually, additional information, for example, warnings about the cause of a speed limit's setting (congestion, slippery pavement, etc.) is also presented to improve driver compliance with the posted speed limit. Accordingly, in this chapter, the motivation for using VSL is provided. Secondly, the concept of mainstream traffic control through VSL is explained, including the impact of VSL on traffic flow dynamics in the context of motorway bottleneck optimization.

2.1 Motivation for VSL

In the following section, the basic relationships between traffic flow parameters and the phenomena of creating congestion on motorways are explained to justify the need for VSL and to better understand the working principle of VSL.

2.1.1 Relationships among fundamental traffic flow parameters

For better understanding and comprehensibility of the later topics, this section presents the basic traffic flow parameters and their relationships using the Highway Capacity Manual 2010 (HCM2010) [36]. Equation (2.1) represents the basic relationship between the three traffic flow parameters, flow rate q [veh/h/lane], speed v [km/h], and density ρ [veh/km/lane], in uninterrupted traffic flow. Although the equation (2.1) algebraically allows a given flow rate to occur in an infinite number of combinations of speed and density, there are additional relationships

that limit the variety of traffic flow conditions that can occur at a location.

$$q = v \cdot \rho \quad (2.1)$$

Figure 2.1 shows a generalized, theoretical representation of these relationships that form the basis for capacity analysis of uninterrupted traffic flow, such as on motorways. The idealized parabola in Figure 2.1 c) representing the flow-density function is placed directly below the speed-density relationship Figure 2.1 a) due to their common horizontal axis, and the speed-flow function in Figure 2.1 b) is adjacent to the speed-density relationship due to their common vertical axis [36]. The shape of these functions is influenced by the prevailing traffic and roadway conditions, as well as by the length of the section under study. Although the diagrams in Figure 2.1 show continuous curves, it is unlikely that the full range of functions will occur at a given location. Real-world data usually exhibit discontinuities, i.e., portions of the curves are not present. The curves in Figure 2.1 illustrate several important points. A zero flow rate occurs under two different conditions. The first is when there are no vehicles on the section - the density is zero and the flow rate is zero. The speed in this case is theoretical and would be chosen by the first driver (presumably with a high value). This speed is represented in the diagrams by the free-flow speed V_f . The other limiting case refers to congestion characterized by the maximum number of vehicles per observed motorway section (traffic density ρ_{jam}), where speed and traffic flow approach zero due to pronounced congestion, i.e., traffic comes to a standstill. Between these two extreme points, the dynamics of the traffic flow cause a maximization effect. When the traffic flow increases from zero, the density also increases because there are more vehicles on the road. In this case, the speed decreases due to the interaction of vehicles. When density and traffic flow rate are low and medium, the decrease is negligible. When density increases, the generalized curves indicate that speed decreases significantly before capacity is reached. Capacity is reached when the product of density and speed gives the maximum flow rate. This condition is represented by the speed at capacity V_{cap} (often called critical speed V_{cr}), the density at capacity ρ_{cap} (sometimes referred to as critical density ρ_{cr}), and the maximum flow rate q_{max} . Based on Equation (2.1), the slope of any ray originating from the speed-flow curve represents the reciprocal of the density. Similarly, a ray in the flow-density curve represents speed. Figure 2.1 shows example of the average speed at free flow and at capacity, as well as the optimum and the congestion density. If one of the relationships shown in Figure 2.1 is known, the other two are uniquely defined. Figure 2.1 shows that any deviation of flow rate from capacity can occur under two conditions: low density and high speed, and high density and low speed. The side of the curves with high density and low speed represents an oversaturated or congested traffic flow. In such conditions, sudden changes in traffic, i.e., in speed, density, and traffic flow rate may occur. In contrast, the undersaturated traffic is characterized by the

low density and high speed part of the curves [36].

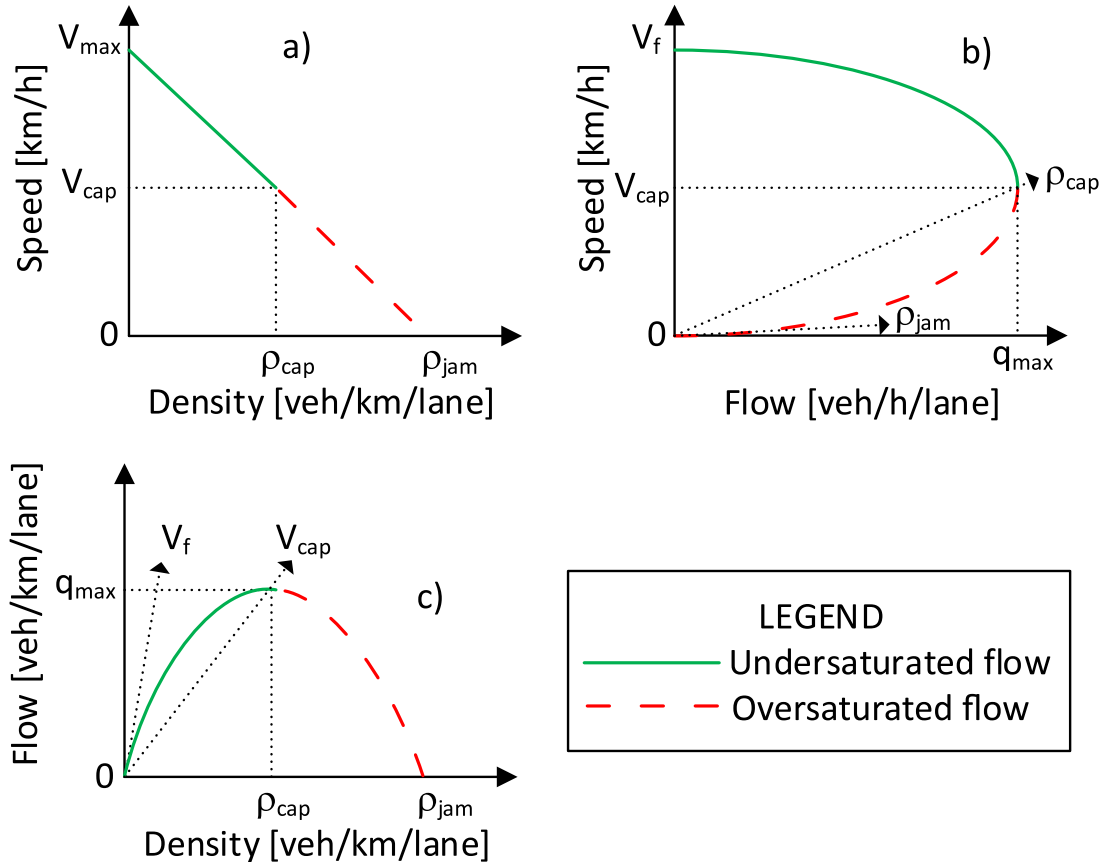


Figure 2.1: Fundamental diagrams of functional dependencies between flow rate, traffic density, and speed [36]

2.1.2 Traffic flow and motorway capacity

According to HCM2010 [36], the capacity of a motorway segment is the maximum sustainable hourly flow rate at which vehicles can be expected to pass through a point or uniform section of a road or lane during a specified period of time under a prevailing road, environmental, traffic, and control conditions. This means that there is no interference from downstream traffic, such as congestion backing up to the analysis point. Any change in prevailing conditions will alter the capacity of a motorway section. A reasonable expectation is the basis for defining capacity. That is, the stated capacity for a given motorway segment is a flow rate that can be achieved repeatedly during peak periods of sufficient demand. Therefore, capacity is not the absolute maximum flow rate observed on such a motorway section. The absolute maximum flow rate may vary depending on the day and place.

An important characteristic of the flow capacity on motorways is its spatiotemporal variation (see [37]). Flow capacities should not be considered only as infrastructure-dependent, constant values. Instead, flow capacities on motorways can vary depending on prevailing environmental conditions such as visibility, lighting, pavement conditions, and driver and/or vehicle type

characteristics, e.g., drivers who are familiar with the motorway system in the area and those less familiar, the proportion of heavy vehicles and passenger cars in traffic stream, etc. However, even under similar environmental conditions, the traffic flow value at which the first traffic breakdown (transition from uncongested to congested state) occurs can vary up to 10% from day to day, as found in [33]. It has been shown that breakdown does not necessarily occur at maximum flow but it can occur at a flow lower or higher than that which is traditionally assumed as capacity. The same study proposed the so-called probabilistic capacity model, which states that traffic breakdown is not a deterministic event, but that its probability increases along with the flow volume. Thus, the breakdown is defined not only by the numerical value of the flow rate but also by the probability of the motorway traffic to break down at that specific value. Accordingly, motorway capacity has nondeterministic properties.

For the purposes of this thesis, the term *operational capacity* is introduced to emphasize that the capacity on the motorway section where VSL is used is determined not only by the above-mentioned factors but also directly by VSL operation. Further on in this context, the text will elaborate on how the operational capacity can be improved under various traffic conditions. This will be achieved by selecting appropriate speed limits and the strategic placement of VSL zones, utilizing the proposed DWL-ST-VSL concept.

2.1.3 Creation of congestion on motorways

Bottlenecks emerge on motorway sections that have present changes in geometry, such as on-ramps, off-ramps, lane drops, grades, curvatures, tunnels, etc. Bottlenecks can also be caused by regulatory measures such as fixed speed limits or unpredictable events such as traffic accidents. As shown in [38], a bottleneck is a location on the motorway where the flow capacity q_{cap}^{up} upstream is higher than the flow capacity q_{cap}^{down} downstream (see Figure 2.2). The nominal bottleneck capacity q_{cap}^{down} is the maximum traffic flow rate that can be maintained if the traffic flow q_{in} arriving from upstream happens to be equal to (or controlled by) q_{cap}^{down} . On the other hand, if the traffic flow q_{in} is higher than q_{cap}^{down} , the bottleneck is activated and traffic breakdown (congestion) occurs. The congestion head is located at the bottleneck, while the congestion tail is moving upstream, as long as the incoming upstream traffic flow q_{in} is higher than q_{cap}^{down} [38].

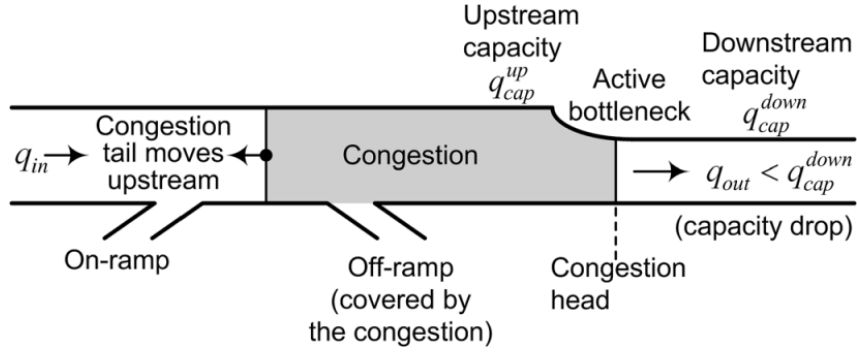


Figure 2.2: VSL concept for mainstream traffic control on motorway [37]

The congestion forming at an active bottleneck has damaging effects on the motorway capacity and throughput. Bottleneck activation leads to speed reduction upstream of the bottleneck location. As a consequence, vehicles need to accelerate (after leaving the congested area) downstream of the bottleneck. This is deemed to lead to a capacity drop phenomenon [39] measured at the congestion head (Figure 2.2). An active bottleneck outflow can be, thus, reduced up to 20% compared to the nominal downstream capacity q_{cap}^{down} [38]. This is illustrated in the empirical flow-density fundamental diagram in Figure 2.3, where two different flows can be observed for the same so-called critical density (ρ_{cr}). The maximum value (horizontal line) represents the capacity flow, while the reduced flow represents the capacity drop.

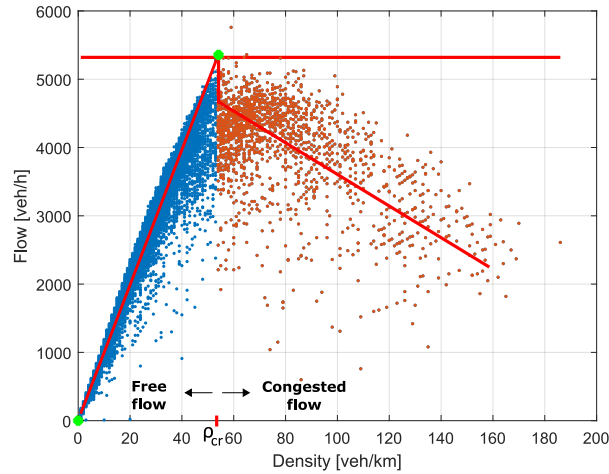


Figure 2.3: Flow-density fundamental diagram with capacity drop at critical density ρ_{cr} [40]

If the increasing congestion spreads upstream and affects, for example, an off-ramp (see Figure 2.2), vehicles trying to exit the motorway before the bottleneck will also face delays. This is known as the blocking of off-ramps effect [38]. Of course, if an upstream on-ramp is affected instead of the off-ramp, the limited nominal storage capacity of an on-ramp will cause a spillback effect onto the local road network.

2.1.4 Concept of mainstream traffic flow control

The task of MTFC at a selected location (usually upstream of the bottleneck) is to adopt the speed limit values specified by an appropriate control strategy to ensure optimal traffic conditions in terms of maximum sustainable throughput (capacity flow) for any traffic demand that occurs. A local aspect of this idea is described in [38] and shown in Figure 2.4. The bottleneck of Figure 2.4 would not be activated (no need for MTFC) as long as $q_{in} < q_{cap}^{down}$, in which case $q_{out} \approx q_{in}$. When q_{in} becomes larger than the capacity of the bottleneck q_{cap}^{down} , the bottleneck would be activated without control, as in Figure 2.2, and q_{out} would be reduced due to the capacity drop. On the other hand, MTFC can implement a controlled outflow q_c equal to the capacity of the bottleneck (or correspondingly less if the bottleneck is due to a merging on-ramp). Clearly, the mainstream congestion cannot be avoided by MTFC because $q_{in} > q_{cap}^{down}$, but it can be diminished. Otherwise, if $q_{in} < q_{cap}^{down}$, MTFC is not needed.

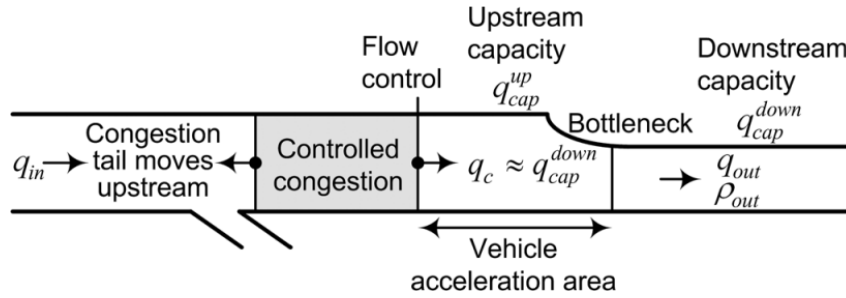


Figure 2.4: VSL concept for mainstream traffic control on motorway [37]

The congestion outflow is higher in the MTFC case than in the no-control case because the capacity drop is avoided. For the same reason (higher outflow with MTFC), the resulting congestion in the MTFC case has a higher internal speed and is shorter in space and time than in the no-control case. Thus, MTFC results in less blockage of upstream traffic, potentially mitigating the negative impact on upstream infrastructure capacity.

2.2 Effect of VSL on motorway traffic flow

VSL is one of the ITS traffic control services that can be used to increase motorway efficiency and LoS in areas with frequently recurring bottlenecks [28]. The effects of VSL on traffic flow were studied in [4, 6, 41]. VSL control measures were first used to improve motorway traffic safety by harmonizing traffic [2, 42, 43]. Additionally, VSL was also examined in the context of improving motorway throughput.

2.2.1 Speed harmonization effect and traffic safety

The use of the VSL strategy to harmonize traffic flow is based on the assumption that lower speed limits reduce spatial variations in speed differences between vehicles and differences in average speed between lanes (homogenization of speed) [4]. Accordingly, in such a traffic flow, vehicles travel more uniformly, i.e., there is less variation in distance and speed between them. The temporal distance (known as *headway*) is generally shorter, which means higher throughput. Therefore, the speed limits are set around the critical speed V_{cr} , which provides ideal conditions for traffic flow regarding throughput, where the maximum sustainable throughput, equal to the motorway section's capacity, is guaranteed for given prevailing conditions. As shown in [44], the speed limit is among multiple dependent factors that impact the level of vehicle crash risk on motorways. Mainly, reduced traffic flow speed variance solves both the road safety level and the risk of capacity drop [2].

2.2.2 Impact of VSL on exhaust emissions and energy consumption

Reducing the speed limit by applying VSL has the effect of decreasing the average speed of vehicles in traffic flow and, consequently, decreasing the resistance to the movement of the vehicles (the most important is air resistance, which is proportional to the square of the speed). This reduces the fuel consumption (energy in general) required to overcome the resistance of the vehicle's motion, and, thus, reduces harmful exhaust emissions [45]. For example, in paper [46], analyzes have shown that it is possible to reduce emissions by 5–7% by applying VSL. If the potential impact of preventing traffic congestion through the application of VSL and, thus, avoiding stop-and-go actions is taken into account, the savings could be much higher. For example, it is shown in [47] that by the use of VSL, more stable traffic speeds during congested periods can be achieved (avoiding stop-and-go oscillations), thus, significant emission and fuel consumption reductions could be achieved, noting that the reduction potential per vehicle kilometer during congested periods is at least 40% higher than in free-flow traffic.

2.2.3 Impact on motorway throughput

The second useful effect of the VSL application is seen in the possibility of regulating the flow rate by the use of VSL. Thus, the goal of VSL is to reduce the inflow into the downstream bottleneck to eliminate or prevent bottleneck activation, thereby eliminating the determinantal capacity drop effect.

Impact on undersaturated flow

As shown in [4] and originally justified in [48], VSL application in case of traffic characterized as free-flow (undercritical density $\rho < \rho_{cr}$), mainly serves the same flow at lower speed and higher density. However, some countries have good experience with lowering the speed limit to stabilize (harmonize) traffic flow. By doing so, the capacity can increase up to 10% if the speed limit is reduced from free-flow speed to approximately 80 [km/h]. In Germany and England, for example, speed is reduced by about 20 [km/h] at a saturation value of 0.7 to harmonize traffic flow before traffic density increases too much [2]. When VSL is applied to free-flow traffic (characteristics of traffic far from saturation), vehicles take more time to travel a given distance at a lower speed, thus, travel time is generally increased. Nevertheless, something interesting happens with the traffic flow rate. VSL temporarily decreases the mainstream flow measured at the end of the VSL application area. The temporary decrease in traffic flow (see Figure 2.5c) is due to the fact that the traffic density under the VSL application is higher than in the original no-VSL case (Figure 2.5a). Therefore, the traffic flow is temporarily reduced *during the transition* to reach the higher traffic density of the VSL condition. Once the higher density is reached, the effect of the temporary reduction in traffic flow disappears (Figure 2.5d). To clarify slightly, the traffic density in Figure 2.5 is represented by dots, where one dot represents one vehicle.

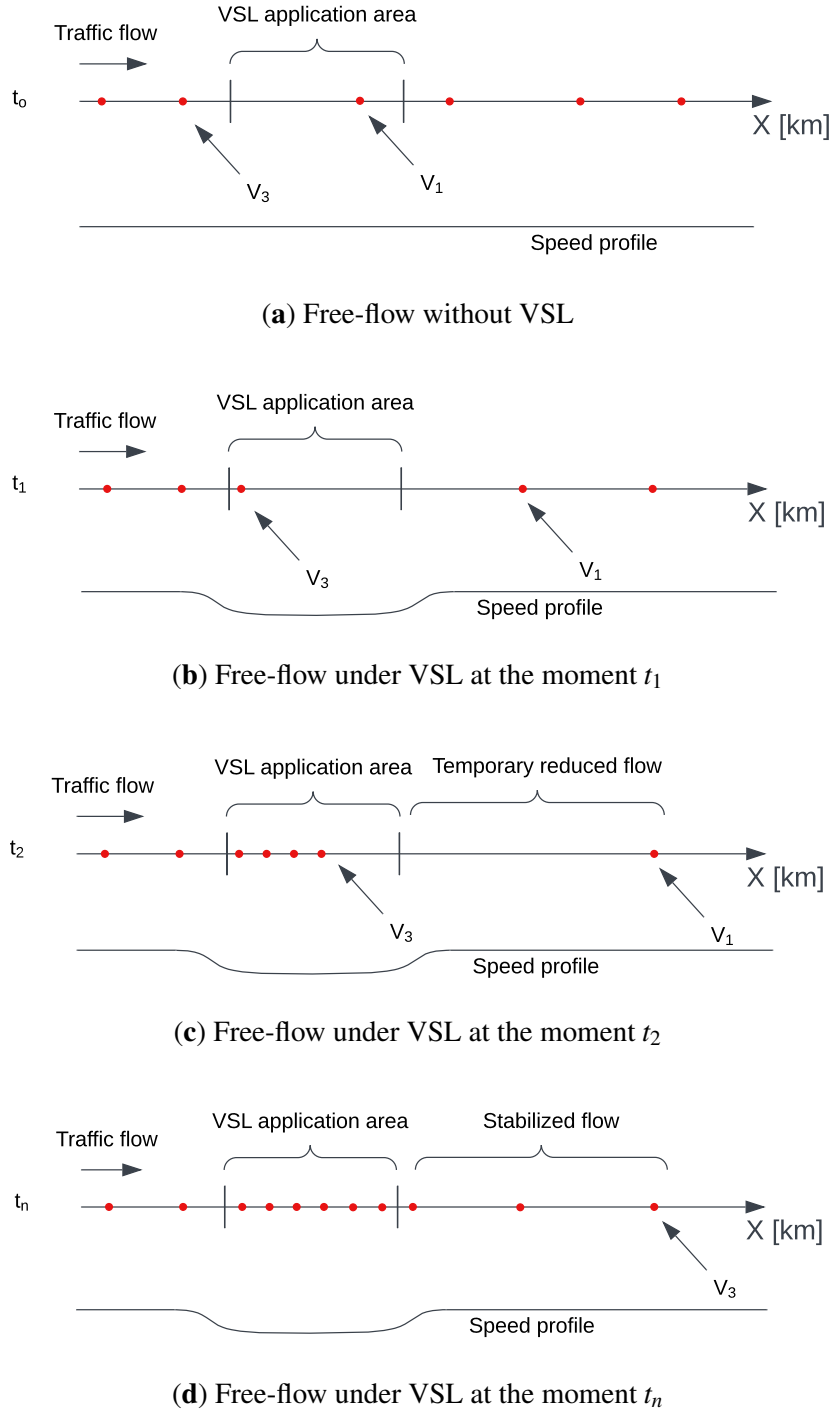


Figure 2.5: Temporary reduction of traffic flow due to the transition of density (from a low value to a higher one) in the VSL application area

Impact on oversaturated flow

When the traffic flow density reaches values above the critical density (overcritical density $\rho > \rho_{cr}$), the traffic flow becomes unstable (the right side of the fundamental diagram, Figure 2.3). In the case of unstable traffic flow, the interactions between vehicles are more pronounced. The resulting disruption of the traffic flow (e.g., due to the braking of one vehicle in

a queue) quickly spreads to other vehicles nearby (moving traffic congestion known as a shock wave [49]), triggering a chain reaction that can lead to a complete traffic breakdown on the motorway. Complementing the previous findings, more detailed studies on the effects of VSL control systems on unstable traffic flow are summarized in the works [4, 50].

In the simplified triangular version of the fundamental diagram (Figure 2.6), the speed limits are associated with the slope of the line starting from the origin. For a given slope corresponding to the speed limit ($\tan(\theta) = v$), the point of intersection of the line with the triangle is defined. This point represents the theoretical flow rate expected at a given speed limit. From [51] it appears that the critical density threshold ρ_{cr} in the fundamental diagram shifts to a higher value ($\rho_{cr} \rightarrow \rho_2$) under the effect of the new speed limit v_2 , where $v_2 < v_{cr}$.

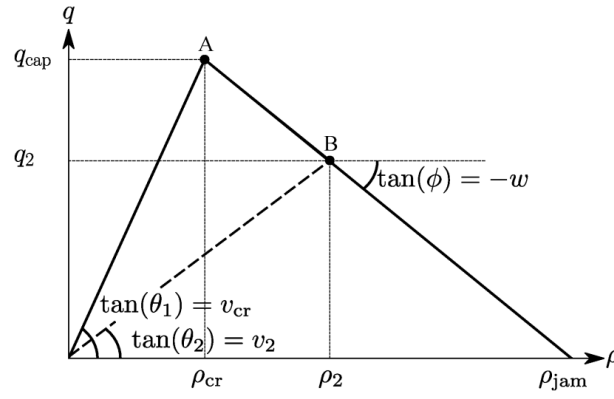


Figure 2.6: Impact of a new imposed speed limit v_2 on simplified flow-density triangular fundamental diagram [28]

By shifting the critical density point, it is possible to accommodate more vehicles on the same length of the motorway section without causing the traffic flow to become unstable. This effect has been confirmed in research [4] and is attributed to the aforementioned effect of homogenization of speed. Accordingly, the new flow rate q_2 is set lower than the capacity flow rate q_{cap} . It is expected that vehicles entering the VSL application area will also encounter the prevailing VSL operating conditions, i.e., they will adjust their speed to heavier traffic, resulting in slight changes in traffic flow known as traffic waves. Such a wave propagates upstream at speed w relative to the road (this is the reason for the negative sign in front of w) and has a magnitude equal to the slope $\tan(\phi)$.

Regarding the capacity increase of the motorway section, the research results are insufficient, as the capacity is slightly increased at certain locations of European motorways. At other locations, no capacity increase was observed, regardless of the set amount of VSL.

But whether capacity is slightly increased or not, it has been shown that sufficiently low values of the speed limit lead to a reduction in capacity in the fundamental diagram (Figure 2.6). Thus, if the traffic volume upstream of the VSL q_{in} (incoming traffic flow in the VSL application area)

is greater than the capacity created by the VSL (operating area of the VSL), then the VSL area becomes an active *controllable* bottleneck whose outflow q_{VSL} is defined by the amount of the VSL, i.e., the capacity of the artificially created bottleneck. Accordingly, different speed limits lead to different reductions in traffic flow. The effects of speed limits on traffic flow are shown in Figure 2.7, where the speed limit value b is given as the ratio between the speed limit and the nominal speed.

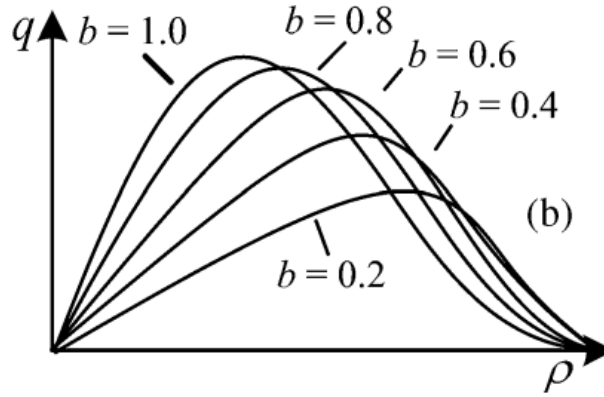


Figure 2.7: Fundamental diagram for different VSL rates [38]

Thus, the incoming flow q_{in} , whose volume is larger than the capacity created by VSL, can be permanently reduced to q_{VSL} . Such a reduced flow remains permanent even after the transition period illustrated in Figure 2.5. An artificially created *controllable* bottleneck on the motorway section upstream of the actual bottleneck (congestion) enables the preventive effect of VSL regulation by limiting the flow rate into the congested area to the approximate value of the bottleneck's capacity ($q_{VSL} \approx q_{cap}^{down}$). In this way, an attempt is made to relieve the overloaded region and, thus, prevent an additional capacity drop at the bottleneck area. With such a main-stream flow control, maximum throughput at the bottleneck is achieved given the current traffic condition (Figure 2.8). The control logic of the VSL controller for bottleneck optimization utilizes this fact, and research related to the application of different strategies to control the traffic exiting the VSL application area (q_{VSL}) aligns with this objective, with the goal of achieving the highest possible operational throughput on critical motorway sections. It is also worth noting that vehicles leaving the VSL application area at relatively high speeds (e.g., when the speed limit cannot fall below 60 km/h) may accelerate and exceed v_{cr} if the distance between the end of VSL application area and the bottleneck is too long. Therefore, to allow efficient traffic flow through the bottleneck, it is also advisable to set an appropriate speed limit on VMS3 for vehicles entering the acceleration area (see Figure 2.8). More on this can be found, for example, in [38]. On the other hand, one must ensure the appropriate distance between the end of the VSL application area and the bottleneck so that vehicles can enter the bottleneck at v_{cr} to ensure maximum throughput throughout the bottleneck.

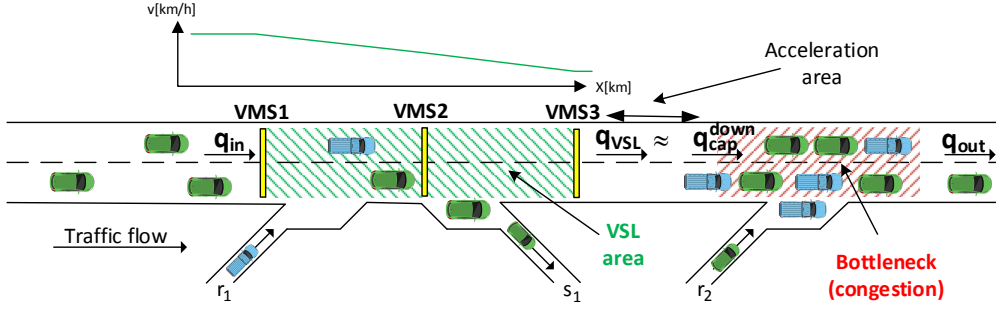


Figure 2.8: Application of VSL to increase throughput in areas with bottlenecks

In general, the previously described effects of VSL on free flow (temporary reduction of vehicle flow) or congested traffic flow (permanent reduction of vehicle flow) cannot be related to a specific value of the speed limit, since the occurrence of the mentioned effects is also influenced by the amount of traffic flow (q_{in}) entering the controllable motorway area. A temporary reduction in traffic flow is specific to higher speed limits and lower traffic loads, and the occurrence of a permanent reduction in traffic flow is specific to lower speed limits with higher traffic loads [38].

2.3 VSL control strategies

To ensure an efficient MTFC, a traffic-responsive control logic for computing speed limit values must be defined. Over the years, various VSL control approaches have been proposed based on different system configurations and methodologies, e.g., optimal control, model predictive control, feedback control, shock wave theory [28, 52], and researchers have recently focused on the development of learning-based VSL controllers [11]. Therefore, classical control approaches based on control theory are further reviewed and recent work on the application of intelligent learning-based agents, focusing on RL techniques, is presented.

2.3.1 Classical VSL controllers

The most prominent VSL design (among classical controllers) uses a negative feedback loop. Feedback-based VSL controllers base their speed limit changes as reactive responses (corrective behavior) to the deviation of the controlled process variable (e.g., traffic density) from the reference (e.g., predefined desired density value in the bottleneck) [38] (see Figure 2.9). Feedback-based VSL can be extended by model predictive control and can work in a coordinated fashion in the case of multiple consecutive VSL application zones to address the shortage of delayed responses. However, model predictive control generally does not guarantee the stability of the control loop and is much more computationally intensive [13] since it requires an

accurate model (e.g., either in the form of equations or simulation) of the controlled process.

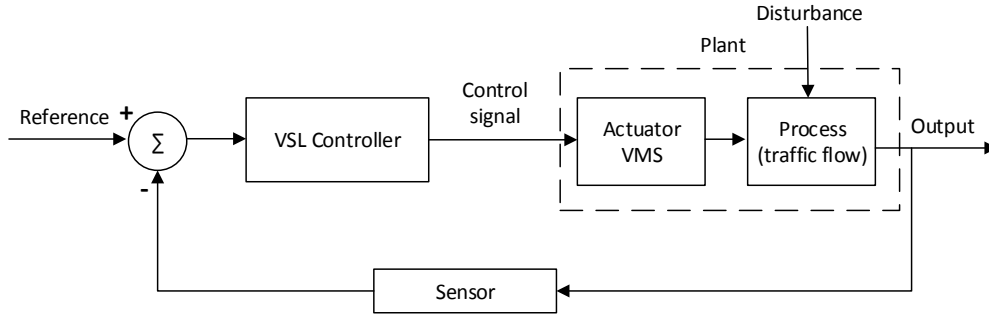


Figure 2.9: Feedback VSL control scheme

A local feedback-based VSL controller (FB-VSL) is, thus, a simple yet efficient MTFC strategy that is directly suitable for possible implementation in practice [38]. Because it is traffic-responsive, it is robust to changes in actual traffic conditions. However, such controllers are linearized and tuned to a specific operating range, and, thus, are not adaptive. Due to the non-linearity of the system, e.g., if traffic patterns and traffic load change significantly, the state of the system may exceed the operating state space of FB-VSL, and, consequently, the controller may not be able to achieve the desired system state in a timely manner. Therefore, VSL could operate suboptimally [20]. However, this can be partially overcome by linearizing the model around several operating points. A linear controller can be developed for each of these operating points, and control gains can be tuned from one region to the next (an approach known as gain scheduling) [13, 28].

2.3.2 Learning-based VSL controllers

Over the last few years, there has been a renewed interest in improving VSL optimization through control concepts based on RL [19, 20, 21, 24, 53]. In [20], it is shown that RL-VSL can yield better results when applied to system travel time optimization in the case of recurrent motorway congestion as compared to a two-loop feedback cascade VSL control structure. The results reported that the feedback-based VSL controller could lead to a delayed response to the fluctuating traffic load when controlling the bottleneck density. On the contrary, the RL-VSL can learn traffic patterns that trigger the activation of a bottleneck through the learning process. Hence, in some cases, RL-VSL can anticipate bottleneck activation and respond proactively.

In [21], the control policy of RL-VSL was further improved by enriching the agent's state variables with predictive information about the expected traffic state by forecasting the speed and density of the controlled simulated motorway segment with the help of running parallel

simulations. RL can be integrated with function approximation techniques (linear or nonlinear). Approximations address the large dimensionality problem of storing state-action values in the computer's memory [18] and enable the computer to work with continuous state/action variables, which, in the end, plagues many real systems with a large solution space, such as RL-VSL [24, 53]. Nonlinear function approximation techniques may improve control if the underlying controlled process is nonlinear and nonstationary, as is the case with motorway traffic flow control [21, 54].

In [25], a multi-agent VSL with two objectives was tested. Flow control policy aims to increase throughput in the bottleneck, while traffic safety policy aims to reduce the speed difference between adjacent controlled motorway segments. Each policy was learned and evaluated separately. According to the defined objective, VSL agents have to learn an optimal joint strategy (policy) using distributed Q-learning. The results indicated an improvement in vehicle stops and total travel time compared to the no-control case. Similarly, in [22], a Q-learning based coordinated hard shoulder control strategy and VSL were introduced.

Accordingly, in chapter 6 of this thesis, the RL-VSL approach introduced in chapter 5 is extended in a multi-agent structure. Using the W-learning (WL) algorithm [55], two RL-VSL agents learn to jointly control two motorway segments in front of a congestion area. WL gave better results in tested traffic scenarios, including dynamic and static traffic loads, and proved suitable as a multi-policy optimization technique in VSL when used for noncooperating agents.

Several manually configured WL-VSL configurations were also analyzed, including different VSL zone lengths and their distances relative to the bottleneck area. The results confirmed that changes in VSL zone configurations affect the traffic flow control process differently. These results are consistent with the findings of other researchers regarding the optimal location and length of the VSL application area [28, 29].

Thereby, one can hypothesize whether VSL performance under such conditions could be improved by having the VSL controller dynamically adjust the length and location of the VSL zone (adjustable VSL application area, *similar* to the concept of the dynamic control cycle in VSL suggested in [56]) in response to changing congestion rather than using static VSL zones. In chapters 6 and 7, this hypothesis is confirmed experimentally for a two, three, and four-agent system configuration using realistic microscopic traffic simulations generated by DT-GM. Especially for spatially and temporally varying traffic congestion, dynamic VSL zone allocation proved to be advantageous over static VSL zones (fixed length and location). The appropriate adaptive VSL zone configurations were learned using DWL-ST-VSL without the need for a manual setup. Furthermore, the results demonstrate the necessity of a collaborative multi-agent approach in VSL systems that can implement heterogeneous policies and action spaces, which enables the implementation of the concept of adaptive VSL zones on a long motorway segment.

2.4 Aspects of VLS design

The value of the speed limit and the proper placement of the VMS prior to the occurrence of congestion, i.e., the location of the VSL zone, are key factors in VSL performance. Therefore, continuing with this theme, an overview of the current findings on the design of VSL zones is provided.

2.4.1 Spatial aspects of VSL design

Theoretical aspect of VSL zone positioning

Pioneers in defining important theoretical assumptions regarding VSL zone setup are the following works [37, 57]. These researchers introduced the concept of VSL application area and vehicle acceleration area. The former, of course, refers to the area where a speed limit is imposed, while the acceleration area is defined as the distance between the end of the VSL application area and the bottleneck. This results in two kinds of degrees of freedom with respect to the placement of the VSL zone. In principle, the use of an acceleration zone makes sense when MTFC is used as a control measure to reduce traffic flow, which means lower speed limits of up to 40 [km/h] or even less. Thus, after VSL is applied, it makes sense to provide additional space for vehicles to accelerate back to v_{cr} before entering the bottleneck for which capacity flow is reached [38].

Recent research [28, 29, 58] has reignited the question of the VSL zone's optimal location. In [28], a simulation approach is used to determine the optimal location and length of the VSL application area with respect to its distance from the bottleneck. The stepwise variation of the length of the VSL application area and the acceleration area is used to show the dependence between the lengths and the system travel time, measured in total time spent (TTS). TTS is the number of vehicle hours, representing the amount of time vehicles spend on the observed motorway [21]. The analysis shows the existence of a mostly linear correlation in the variation of VSL application area, while a non-linear correlation is present in the variation of the acceleration area and measured TTS (see Figure 2.10).

Despite the fact that optimal points were found for both distances, they are determined separately. It is more likely that, for example, a multi-criteria analysis involving both parameters simultaneously could be useful to find the joint sweet point.

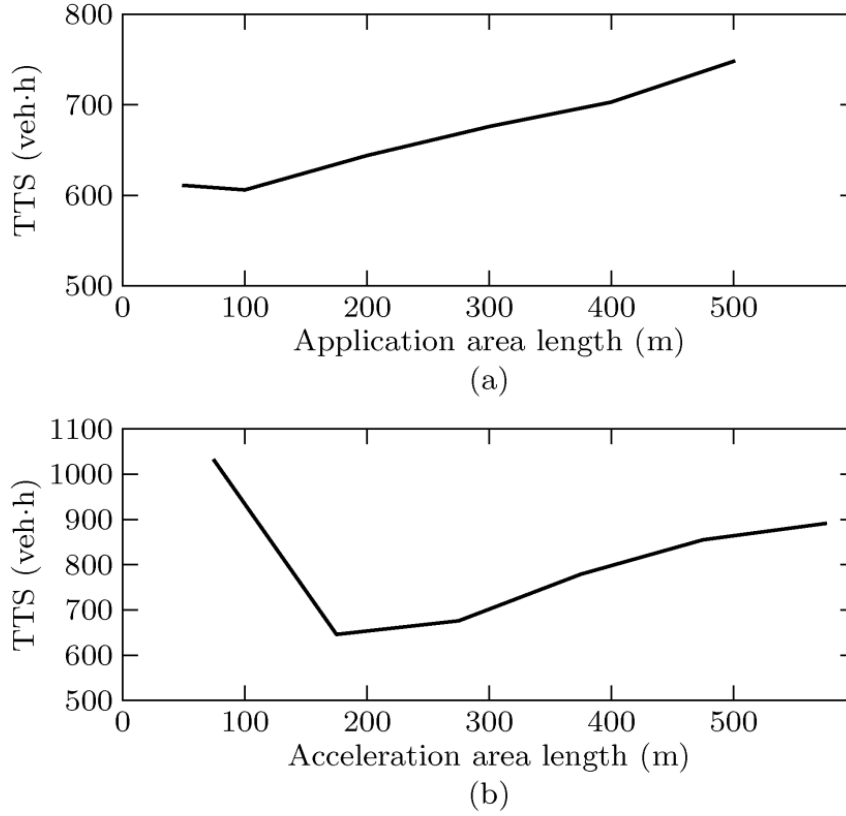


Figure 2.10: Variation of TTS with the length of (a) application and (b) acceleration areas [28]

In [58], the effect of MTFC locations (through the use of green-red light signals) is proposed to achieve efficient merging of vehicles entering the work zone bottleneck (lane reduction). In particular, the study [58] quantifies travel time savings (by avoiding the capacity drop) as a function of the MTFC location. The key findings state that traffic lights located too close to the bottleneck do not allow vehicles to accelerate sufficiently and, consequently, travel delay increases. Therefore, traffic lights should be placed far upstream of the work zone to allow vehicles to accelerate and pass through the merge area at higher speeds. It is also shown that, for a given range of distances, there is no advantage to moving the proposed MTFC further upstream from the bottleneck. Thus, the presented results are similar to the results published in [28], which indicates the need for an acceleration area.

The recent results of [29] provide new insights into the optimal placement of the VSL application area compared to previous findings and are confirmed analytically. It is shown that the general assumption that the lower the speed limits, the larger the distance between the VSL application area and the bottleneck (to enable vehicles to reach v_{cr} before entering a bottleneck) is not always the case. Instead, the results indicate that at a higher value of the speed limit, the distance between the VSL zone and the bottleneck should be larger. Although the analytical study presented is quite comprehensive and complete, it was conducted using the Lighthill-Whitham-Richards (LWR) traffic model with bounded acceleration (BA-LWR) and, therefore,

may lack accurate modeling of specific traffic phenomena due to LWR inherent characteristics. The LWR model is based on the first-order fluid approximation of traffic flow dynamics proposed by Lighthill and Whitham [32] and Richards [59]. LWR in general provides a rough description of traffic behavior for a single one-way road using three variables that vary in time and space: flow, density, and speed. Other shortcomings of the LWR theory are that it does not properly describe platoon diffusion, a phenomenon that occurs over long distances in case of low traffic volume. Conversely, it cannot explain the instability of heavy traffic, which exhibits oscillatory phenomena observable at the minute resolution scale. Moreover, since the theory behind it is based on fluid and gas theory, LWR does not account for the anisotropic property of vehicles [60]. In fluid mechanics, for example, a particle (or molecule) feels all the forces acting on it equally, regardless of their direction (the pressure exerted by neighboring particles). On the other hand, in road traffic, drivers mostly take into account the traffic in front of them (less lateral traffic and traffic behind them). Accordingly, the physical property of vehicle dynamics has different values, i.e. it senses front, lateral and rear traffic differently, which LWR does not take into account. For numerical verification, the authors also used traffic data at a lane drop bottleneck on a Toronto motorway from 1988. These data appear to be quite old, as the dynamics of today's vehicles can vary widely compared to those from the 80s, which can affect flow characteristics near the bottleneck, where deceleration and acceleration rates are critical variables in the used BA-LWR model. Therefore, some conclusions described in [29] may involve the incomplete effect of VSL on traffic behavior, and it would therefore be appropriate to investigate the presented finding in more detail using the microscopic traffic models with a better representation of vehicle flow dynamics to strengthen the levels of confidence in the results and conclusions.

Practical implication in modeling spatial VSL

In [61], the authors point out the problem of the optimal VSL zone design for the optimization of the bottleneck. Therefore, they propose three VSL zones: the critical VSL zone for regulating the discharge section flow to match the bottleneck's capacity, the VSL zone for the potentially congested area (mainstream storage in case congestion propagates upstream), and the VSL zone upstream of the congestion tail. The analysis performed in [62] suggested a VSL control model that is able to determine whether the specific location on the motorway section is congested or not based on predefined thresholds (density, speed, and acceleration), and using this information the starting point of VSL zone is determined. In [63], the bilevel programming model is used to find the most appropriate speed limits and corresponding locations of VMSs in VSL control. The first objective of the bilevel programming model was to optimize the number of VMSs and speed limits by creating a model to minimize the overall accident rate in areas of the motorway work zone. The second objective was modeled to optimize VMSs locations by solving the

improved maximum information benefit model. The results presented confirm that appropriate speed limits and proper VMSs placement can reduce the average queue length, total delay, and total frequency of vehicle stops in motorway work zones which ultimately reduces the vehicle crash risk.

Although the results of the above-mentioned analyses point to a possible direction for addressing optimal VSL zone placement, in general, the results and findings indicate that there is no absolute guideline for where the VSL zone should be placed for optimal performance. Instead, it appears that the near-optimal placement of VSL zones depends on the location and intensity of congestion and the speed limit values range used for MTFC.

Given that the congestion characteristic varies in time and space due to stochastic traffic behavior, one can try to experimentally confirm the usefulness of the DWL-ST-VSL concept of dynamic VSL zone allocation for speed limit control. It is also worth demonstrating that the DWL-ST-VSL agents and the motorway system could benefit from their cooperation in selecting appropriate actions, not only for their own policies but also for the policies of the surrounding agents they affect.

Therefore, in chapters 6 and 7, a simulation proof is provided of the extended concept of DWL-ST-VSL and its applicability to speed limit control on a longer motorway segment along with adaptive VSL zones, which is more in line with what is required in the real world in order to achieve harmonized traffic flow control. The analysis gives detailed insight into the steps of modeling DWL-ST-VSL and provides some interesting information on the pros and cons of the proposed learning-based multi-agent VSL controller. These are the primary research motivators for implementing the DWL-ST-VSL strategy that learns appropriate speed limits and spatiotemporal VSL zone adaptation in an automated manner using the DWL algorithm. The proposed DWL-ST-VSL is trained and evaluated on the synthetic simulation model as well as on the proposed DT-GM that is discussed in more detail in chapter 3.

2.4.2 Temporal aspect of VSL design

In addition to appropriate speed limits, it is important to ensure the timely application of speed limits in response to a change in traffic conditions to achieve MTFC or traffic harmonization. This is ensured by the timely detection of traffic changes, or even their anticipation, and appropriate VSL control responses. However, most VSL control approaches use fixed (best found) control cycles for computing decisions. Considering that the duration of traffic congestion may vary, such an approach may not be the best.

Accordingly, a dynamic control cycle was proposed in [56] to calculate the optimal duration of control cycles in VSL. It has been shown that the dynamic control cycles perform better

than the fixed cycles. With the proposed strategy, the duration of each control cycle can be adjusted considering the current traffic condition and speed limits, so that the VSL can respond appropriately to time-varying traffic conditions, which makes the performance of the VSL more robust in a nonstationary environment such as motorways.

2.4.3 Additional information on VMS

Even in the case of a perfect VSL design, its efficiency is affected by the compliance rate of drivers. To ensure a higher compliance rate, it is worthwhile to give additional information next to the speed limits displayed on VMS, in particular, a brief explanation of why a certain speed limit is applied. Otherwise, drivers may not take the speed limits seriously. Research in [64] suggests that understanding time is low when there are fewer units of information on a VMS. However, too little information may be unclear or ambiguous, while too much information may be difficult to understand and cause further driver slowdown. Moreover, it has been shown that drivers easily understand and therefore comply with VSL recommended speed limits in bad weather, but it is more difficult to understand speed limits when they are used solely to increase throughput. For example, in Croatia, when VSL is active in sunny weather without incidents, most drivers think that the VSL system is malfunctioning [65], and, thus, drivers do not obey the recommended speed limits.

Based on [66], one of the main variables affecting the probability of speed compliance at a given speed limit is the driver's understanding of the VSL system, with education being the second factor that has a positive effect on speed compliance. Driving schools need to provide quality driver education that also takes into account all traffic situations and conditions on motorways where VSL is used, as well as the need to comply with recommended speed limits to ensure a more efficient VSL system. Thus, education is also connected with the ability to understand how VSL work. For example, if road operators were to invest in testing and marketing VSL systems, it could have a statistically significant impact on speed limit compliance, and, thus, VSL performance.

2.5 Concluding remarks

Due to the non-linear nature of the traffic flow, this chapter summarizes some important findings which draw several questions that will be connected with the research described in chapter 4. First, the basic concept of the VSL application is to homogenize vehicle speeds in traffic flow, thereby VSL:

- increases road safety operation level;

- reduces the risk of capacity drop (particularly in the bottleneck);
- can increase the operational capacity when the speed limit is reduced from free-flow speed (nominal speed limit) in uncongested traffic (saturation value ≈ 0.7).

Second, in optimizing the bottlenecks, VSL aims to eliminate or prevent bottleneck activation, thereby eliminating the negative capacity drop effect, thus, ensuring a higher throughput (up to 20 %). This is achieved through a fundamental principle of VSL to reduce the flow rate from the VSL application in order to relieve the bottleneck. Basically, two kinds of flow reduction can be achieved by VSL:

- a temporary reduction in traffic flow is specific to higher speed limits and lower traffic loads;
- a permanent reduction in traffic flow is specific to lower speed limits deployed in higher traffic loads.

To ensure adequate traffic-responsive MTFC, a VSL controller is needed. It turns out that among the controller logic, some other parameters are also important for adequate VSL system design and driver compliance rate:

- the traffic-responsive VSL controller logic;
- the duration of the control cycle and the speed limit values;
- the lengths of both the VSL application area and the acceleration area in relation to the bottleneck's location;
- optional: speed limit for acceleration area;
- the additional units of information on VMS.

Finally, an important fact regarding motorway operation is the uncertain flow capacity, which states that traffic breakdown is not a deterministic event, but that its probability increases along with the flow rate (i.e., the flow capacity is defined by the numerical value of the flow rate but also by the probability of traffic breakdown). Thus, capacity is a function of several variables:

- geometry of the motorway;
- environmental variables like visibility, weather conditions, etc.;
- driver's behavior characteristics and vehicle attributes;
- as a function of traffic load itself.

Additional questions might be of interest for VSL design comprising traffic flow and motorway capacity characteristics:

- Is there a relationship between the optimal deployment of VSL zones that comprises both, the probabilistic spatiotemporal constraint on motorway capacity $q_{cap}(x, t)$ upstream of the bottleneck, and also the bottleneck capacity $q_{cap}^{down}(t)$ (assuming fixed bottleneck location)?
- Is it possible to prolong the effect of temporary flow reduction in undersaturated flow conditions, e.g., by dynamically reallocating VSL zones?

To the best of the author's knowledge, for example, the probabilistic spatiotemporal constraint on motorway capacity has not been considered so far in the context of optimal VSL zone placement and may contribute to a better understanding of the current results as well as provide new insight into the need for dynamic expansion (or contraction) and reallocation of VSL zones during VSL operation for a more efficient VSL. Therefore, this and particularly the phenomenon of temporary flow reduction, along with the fact that motorway capacity varies in position and time, are further elaborated in chapter 4 in the context of deriving the dynamic adaptive VSL zones model.

Chapter 3

Digital twin technology on motorways

As pointed out in the first chapter, one of the goals of this thesis is to define a simulation framework for learning traffic controllers that reproduces accurately actual traffic dynamics (in near real time) for the controller's training and evaluation process. For this reason, in this chapter, the concept of digital twin technology for microscopic simulation modeling of motorway traffic is proposed.

For complex stochastic systems, the corresponding dynamic model may be overly challenging to obtain in a closed form. In such cases, acquiring an exact mathematical model requires simplified assumptions about the system and, therefore, results in an approximate closed-form model. The simplifications often make them too simplistic for real-world usage. One way to create useful models for large and complex systems is to utilize a computer program (a simulation model) that mimics the behavior of the system. Such a model achieves its goal of predicting and simulating system behavior by generating pseudo-random numbers (adequately approximated random by distributions) for the governing random variables of the system [18].

The use of simulation models in traffic processes is various. Traffic simulations are often implemented because conducting experiments on a real system is not feasible due to cost, safety impossibility, or other constraints. This is especially the case when experiments on the real system are dangerous or may result in significant degradation of system performance. Therefore, when developing new traffic control systems, their simulation-based evaluation is important before they are used to control a real process [28]. It is crucial to use a simulation model that accurately reflects the behavior of the real system. Thus, it is an advantage if the simulation model can be logically manipulated to answer questions about how the physical system will behave under the use of control systems and what benefits it will provide. To accomplish this, an accurate traffic simulation must be modeled that can precisely simulate various traffic scenarios that may occur in the real system. Especially those that are meant to be managed

by a controller. These scenarios, which can be associated with an operational domain of the controller, are then used to train/tune and evaluate the control system. In traffic processes, scenarios of interest are, for example, those related to rush hours, when traffic volume exceeds available road capacity and congestion occurs. In such a traffic regime, additional traffic control strategies are required to maintain a stable traffic flow. However, for adaptive learning-based controllers, other scenarios (one that might occur before and after rush hours) must also be presented during the training phase. Thus, it is important to provide a wide spectrum of training examples so that learning controllers can ensure good generalization over the remaining traffic states. Conventional (offline) simulation approaches, however, use historical aggregated traffic data (or artificially created data) for creating simulation scenarios representing only a few selected scenarios. In general, such models are inadequate to reflect unexpected events that are likely to occur in reality and cannot be captured by offline simulated results due to their inherent stochastic characteristics.

Therefore, the use of actual, fine-grained traffic data received directly from traffic detectors as input to the running simulation enables the capture, prediction, and analysis of such micro-level traffic events that can only be adequately captured with a fine-resolution sampling rate. This is of utmost importance for real-time traffic control when learning-based traffic controllers are meant to be used. Accordingly, a new approach to creating relevant DT-based microscopic simulation models synchronized in real-time with their physical motorway counterparts is needed to evaluate adaptive control systems, which is elaborated in the following paragraphs.

3.1 Current approaches to traffic simulation

Created control strategies and their impact on traffic can be analyzed in various traffic simulators with numerous simulation models at the microscopic, mesoscopic, and macroscopic levels [30]. Despite their widespread use, up to this day, traffic simulations have mainly been used for (offline) verification, validation, and optimization of proposed traffic control solutions and transport system design in the early planning phase. Such an approach enables traffic safety analysis and control strategies design in their early planning phase. However, it is necessary to gain confidence in the control system's performance in unforeseen (unexpected) traffic situations as well. Having a control strategy that performs well in all relevant traffic states (generalization) is often more important than superior performance in some states. Especially in the case of learning-based controllers. But in general, this is hard to achieve in a stochastic dynamic system (like a motorway with its corresponding complex, non-linear traffic flow dynamics) in which the process state changes randomly. Traffic control systems must, therefore, adapt to a traffic demand that varies significantly in time and space. Real-time tuning of such systems becomes infeasible for a human operator, and prevailing offline simulation approaches

appear to be inadequate, making new technological support, particularly in the domain of traffic simulations, essential.

However, it has been demonstrated that microscopic simulation models provide a high level of detail in terms of interactions between vehicles and network modeling, making them, however, more computationally intensive [67]. Nevertheless, as such, they can be used for in-depth analysis of smaller traffic networks or isolated sub-networks such as motorways [68], to analyze advanced traffic control approaches [69, 70, 71], or to, for example, the energy consumption of vehicles [72]. In addition, simulator frameworks can be used as a basis for testing new traffic models, such as a lane-free traffic model for Connected Automated Vehicles (CAVs) built and tested on the simulation infrastructure SUMO [73]. Or for a more general simulation analysis of different CAVs penetration levels and their impact on the traffic [68]. Moreover, calibrated microscopic simulation allows for very realistic traffic simulations that closely resemble reality. This makes such models attractive because, given current traffic conditions and appropriate simulation parameters, they can be used for traffic forecasting.

Therefore, the prevailing microscopic simulation models appear to be an effective tool that can be used at any stage of transportation system design, planning, or testing of control solutions [28, 74]. Consequently, offline simulation approaches (based on historical or artificial data) are useful in transportation planning when analytical changes in traffic flow patterns are related to long-term changes in infrastructure, travel demand, mode choice distribution, etc. Although the prevailing *offline* microscopic simulations are widely used in all areas of traffic engineering, they lack the ability to be used in real-time decision-making processes and real-time analytics in TM because they do not account for actual traffic.

3.2 Digital twin technology in microscopic traffic simulations

In this section, an introduction to the current state of research on the application of DT technology in transportation is provided. In particular, using microscopic traffic simulations in this context.

Related Work

In a microscopic traffic simulation model, actual vehicle and traffic dynamics can be simulated and visualized in real-time. This feature makes it possible to introduce the concept of DT into traffic modeling. DT represents a virtual instance of a physical system that is continuously updated with performance, maintenance, and health status data throughout the lifecycle of the physical system [75, 76]. In essence, Big Data, the floating car data, wireless sensor networks, GPS tracking, traffic counters, etc., are rapidly evolving and offer a large potential

for the development and use of DT technology in the transportation sector [77]. However, the time component of the data plays an important role in the creation of run-time DT, i.e., data need to be available in *near real time* to be integrated in a timely fashion into the simulation model. Thus, it is important to ensure real-time data transmission for run-time synchronization of DT with real traffic.

Since real-time transmission of high resolution traffic data is still in its infancy, it is the bottleneck in the development of *continuously synchronized* DTs and their wider use in traffic analysis. However, several research studies have used microscopic traffic simulations to model the accurate digital replicas of physical traffic systems with some delay (remarkable latency) and low resolution of the traffic data used. Consequently, several articles [78, 79, 80] have used DT terminology in the context of traffic simulation modeling using a more (*traditional*) offline simulation approach. They emphasized the use of real, yet historical, traffic data regarding the notable latency in modeling realistic traffic models. In this context, the aggregated historical traffic data were down-scaled to finer data resolution (depending on the purpose of the analysis) by using an estimated probability distribution of random traffic variables like flows, trips, vehicle arrivals, speeds, etc.

However, the property of generality provided by offline simulations may lack the ability to adequately represent events that may occur in daily traffic on a finer time scale, as some traffic events may not be captured by aggregate traffic data that are often used in offline simulations. Thus, using actual run-time traffic data in simulations allows for capturing and predicting unexpected traffic events in real-time. In addition, real-time traffic data combined with simulations enable powerful mathematical simulation models to visualize, analyze, interpret, predict, and optimize traffic. Such real-time predictive analytics is essential for real-time TM and proactive decision-making. Therefore, it may be useful to continuously use fine-grained real-time traffic data (if available) as input to the simulation so that the simulation model can evolve (calibrate itself online) as motorway traffic evolves. In this way, the physical transport system with its corresponding traffic sensors can be used in combination with the real run-time simulation model to provide a more accurate digital replica of real traffic and to see if it is possible to predict traffic failures clearly enough in advance, so that preventive traffic control measures can be taken. This gives the DT simulations a foundation as they can be used as additional support in real-time TM while raising questions such as whether such data could improve the simulations themselves, what benefits such *live* simulation models might have, and to what extent can they be compared to conventional offline simulations.

Inherently, the integration of DT technologies and the real world motorway system offers the ability to augment the existing motorway performance monitoring systems with the digital virtual instance so that the created DT simulation can be updated in near real-time as the physical

motorway equivalents change. As such, DT could play an important role in TM by allowing the virtual model to interact bidirectionally with the physical entity in real-time [76]. This is conceptualized in section 3.3 using the example of DT application for monitoring and controlling VSL on motorways.

From the perspective of a microscopic simulation-based DT traffic model (running in real time and corresponding to real traffic), the evolution of spatial and temporal traffic characteristics can be predicted from a respective moment forward, taking into account the current state of traffic. Initializing from that particular traffic state, several parallel simulations from the running DT simulation can provide a basis for traffic forecasts. Thus, each of these simulations can serve as a test environment to evaluate different strategies and finally identify potential traffic problems before they occur. In this way, failures can be avoided, and future actions can be planned while reducing uncertainties about the system's response [81]. However, neither the development of the DT for motorway traffic nor the strategic advantage of the conceptual application of DT with TM on motorways has yet been fully realized. The following paragraphs summarize the findings of other researchers related to microscopic simulation modeling and concepts, pointing to the definition of DT concepts in transportation in general.

In [78], an application of SUMO microscopic simulation was presented in the application of ITS Austria West. Aggregated vehicle loop sensors and floating car data were used together with an origin-destination (OD) matrix based on historical traffic data from the road network of Upper Austria to generate routes that were used in a microscopic simulation to generate traffic demand. The simulation model was used for five-minute traffic simulations (short-term forecasts) to calculate the new traffic state from a given current state. The traffic condition at the end of each period is integrated with the newly obtained traffic data. These integrated data were used to create snapshots of the calculated LoS information used by the services from ITS. Finally, the simulation scenario is adjusted using the aggregated real-time network sensor data along with the route distribution based on the estimated OD. The updated scenario is then used as the basis for the next short-term estimate of traffic conditions.

In view of the next generation of ITS, a concept for a virtual vehicle (VV) model was presented in [82]. The goal of this concept is to use vehicle and traffic information via edge clouds and specific analytics (deep machine learning algorithms) to predict driver intention. Information about the driver is taken into account, such as driving preferences, which lane the driver or automated vehicle is likely to choose, and route plans. The intent is then used by the VV to compute interactions with other vehicles within the DT model, with the goal of finding the best routing through the network to reduce congestion. Yet, this is a more conceptual approach at the moment and will be more favorable in the vehicle-to-everything communication environment in the future [83].

The article [80] presents a proof-of-concept application of DT for Adaptive Traffic Signal Control (ATSC). The control algorithm aims to distribute the waiting time over a network of signalized intersections instead of loading a single intersection with increased traffic demand. Although the use of real-time data was emphasized, the data itself were generated by the simulation process and used as input to the ATSC. Thus, the simulation scenario itself does not use external inputs from the physical roadway network to calculate the optimal traffic signals, and the initial traffic demand was used during the simulation process (no online change in loading). Online changes were only applied to the traffic signals during the simulation within the simulated traffic scenario via the Traffic Control Interface (TraCI) with respect to the applied ATCS algorithm.

In [84], the simulation process is extended by connecting the traffic simulator SUMO with the game engine Unity, which provides virtual insights into the simulated scenario through 3D virtualization. The synchronized platform enables SUMO to control legacy vehicles, while Unity controls CAVs with the proposed algorithm to achieve optimal coordination of CAVs in a mixed traffic flow of road vehicles. As the trends in the automotive industry are focused on autonomous driving, a DT for security and safety validation of autonomous driving was presented in [85]. A method is proposed to address a selection of exposed vulnerabilities using virtualization through DT of complex systems. In [86], the integrated conceptual framework of the Unity game engine and SUMO is presented for the virtualization of CAVs, and in [87], it is used to demonstrate the mobility-DT framework for personalized adaptive cruise control (P-ACC) in Connected Vehicles (CVs) environment. Combined with Artificial Intelligence (AI) based data-driven cloud-edge computing in Amazon Web Services (AWS) IoT Core, the conceptual framework consists of three building blocks in the physical world: human, vehicle, and traffic, and the associated digital spaces of human DT with user management and driver type classification, vehicle DT with cloud-based driver assistance systems, and traffic DT with traffic monitoring and variable speed limit advisors. Finally, a case study is presented with the application of Mobility-DT to P-ACC for vehicles. However, the approaches are quite conceptual, at least in part, because they do not cover the integration of real world (run-time) data and are, therefore, only partially a DT. In [88], the first field test of DT for cooperative ramp metering is presented. Three vehicles were equipped with on-board devices that send data to the cloud server where the DT of vehicle locations is created. Based on this, a cloud server calculated the target speed for the vehicles in real time to perform the cooperative merging from the on-ramp to the motorway.

The authors of the article [81] present a DT-Smart City for citizen feedback and urban planning. The structure of the DT model is based on different layers (including mobility) aiming to provide an online urban planning view of skylines and green spaces in Dublin so that users can

interact through the 3D virtual model and provide feedback on planned changes. In [79], a new large-scale traffic microscopic simulation model for the Barcelona urban area was presented. The 24-hour simulation scenario is based on fine-grained empirical Big Data that includes mobility data from cell phone records with traditional annual mobility surveys. The objective of this research was to develop an operational and effective approach to create a large-scale digital (microscopic simulation) replica of Barcelona traffic, based on empirical mobility data in SUMO. Finally, hourly trips were generated using OD matrices and used to calibrate the simulation model.

Identification of research gap

While reviewing related work (summary is given in Table 3.1), it was found that the results presented in existing studies can be distinguished depending on the extent of virtualization, the temporal resolution of DT, and the data source (real or artificial). Depending on the data resolution, the quality of the collected data used as input to the simulation model may be overestimated, and these errors are reflected in the aggregate results. Thus, leaving the degree of confidence in the results and conclusions incompletely known [89]. The lack of access to reliable, high-resolution, real-time traffic data during the DT lifecycle appears to be the most limiting factor so far, ultimately determining the update frequency of the simulated DT model and, thus, the system error. None of the above approaches were actually able to present a functioning integration of actual run-time traffic data into a running microscopic simulation, and, thus, the potential of implementing a DT for TM was a kind of open promise.

Moreover, there is virtually no research on the methodological development (and application) of microscopic simulation-based DT of a motorway system. Therefore, the work presented in this thesis extends the body of knowledge by proposing a run-time synchronized DT of the Geneva motorway, along with a method for developing such DTs. Since there is access to real-time sensor data streams from traffic counters on the Geneva motorway via Open Data Platform Mobility Switzerland (ODPMS) [92], for the first time, the methodological steps necessary to create a motorway DT can be proposed, thus, virtualizing the traffic dynamics of the selected motorway in real-time. Unlike existing DT models in transportation in general, the DT model presented in this thesis can be considered as a virtual, microscopic simulation-based instance of the physical motorway system that is continuously updated as real traffic changes over time. Such a run-time synchronized DT-GM model provides the foundation for a wealth of new ideas and concepts in TM and opens up a previously unfeasible avenue for advanced mobility research in general. Therefore, some of them are discussed in the context of safety-critical decision processes on motorways.

Table 3.1: DT IN THE TRANSPORT SYSTEMS AND URBAN MOBILITY

Ref./Year	Physical model/Data source	Virtual model/Calibration	Contribution
[78] (2013)	Traffic counters, floating car data, historical OD matrix	Scenario generator module - periodic calibration every 5 [min] (SUMO)	Visualization of the traffic situation
[90] (2016)	Offline simulation of CVs and Big Data	OMNET++, SUMO, Veins coupled with Cassandra based Big Data cluster, not calibrated	Identifying congestion and rerouting vehicles accordingly
[82] (2018)	TAPAS (computed mobility plans for an area population)	Offline (SUMO)	Concept of coordination of learning VVs in virtual transport network
[80] (2021)	Synthetic network and data	Interaction with traffic signals in SUMO via TraCI, not calibrated	DT concept to allocate signal phase and timing of ATSC
[86] (2021)	DT simulation of CAVs in integrated Unity-SUMO	Conceptual framework, not calibrated	A case study of P-ACC
[88] (2021)	Vehicle to Cloud communication (V2C)	On-board devices send data to the cloud (synchronized in real time)	DT advises CVs for cooperative RM (real experiment with three vehicles)
[84] (2021)	Unity-SUMO Co-Simulation	Synchronized SUMO-Unity platform	Merging coordination of CAVs in mixed traffic
[81] (2021)	Dublinked (Open data store)	3D DT smart city with mobility layer of pedestrian simulation (Unity)	Virtual feedback of citizens on urban planning and policy decision
[91] (2022)	Real data (coarse granularity) from IoT and static geo-data for building OD matrices	An hourly OD flow estimation of each mode of transportation	Simplified field test for a modal split estimate within an interval of two hours in a small district
[87] (2022)	Adopted AWS IoT Core to design AI based data-driven cloud-edge-device framework for CVs	Mobility-DT framework with human DT, vehicle DT, and traffic DT	P-ACC
[79] (2022)	Hourly scaled OD flows of cell phone data based on yearly mobility surveys	Large-scale 24-hour simulation in SUMO for the Barcelona urban area	Effective approach to creating a large-scale digital replica of traffic
This thesis (2023)	Real-time data streams (minute resolution) from sensors on the Geneva motorway (via ODPMS)	Continuously calibrated simulation with actual run-time traffic (synchronized SUMO-ODMPS)	DT of physical Geneva motorway and concept of DT-GM application in real-time TM

3.3 Digital twin motorway concept

Motorway systems have become increasingly smart in recent decades. This evolution is mainly based on the foundation of data collected by roadside sensors and the vehicles themselves, and the application of advanced traffic control approaches. These data enable services like real-time monitoring of the transportation system, and in general, the digitization of the transportation system, providing many benefits to travelers and operations centers [93]. With this information, operators (or autonomous road transportation support [94]) can deploy ITS services and apply TM approaches (control strategies) to improve traffic flows [95]. However, the impact of the control strategies on the motorway system is not apparent until they are deployed in real-world motorway applications. Due to the high speeds on motorways, inappropriate TM strategies placed wrongly in space and time can pose a serious threat to motorway safety and operations [70].

Meanwhile, the enormous potential of information technologies in road transportation enables real-time fusion of the physical road system and the digital world (simulations) in traffic modeling and analysis [79, 80, 88]. In general, such integration is referred to as the DT concept, which considers a fusion of the virtual (digital) and the physical world systems or processes [81]. As such, the applications of DT offer a promising way to address the growing complexity of transportation systems through the concept of monitoring their entire life cycle. *Continuously synchronized* DT model with its physical counterpart enables real-time bidirectional interaction between the physical system and corresponding digital replica in the early stages of design, construction, reconstruction, traffic planning or control, and various other analysis purposes. Accordingly, DT technology has shown considerable application potential in several areas and is attracting considerable interest from industry and academia [76].

However, the strategic advantage of this integration has yet to be exploited in the motorway domain. Motivated by the aforementioned predominance of conventional (offline) simulation approaches, it was decided to explore the possibilities of integrating the SUMO and newly available real-time motorway traffic data from ODPMS, which allows one to pay attention to the simulation application during motorway system run-time.

Thus, in this section, a run-time synchronized DT-GM is proposed. Compared to the existing literature on DT in road traffic applications in general, the novelty of this approach is the use of fine-grained traffic data streams coming directly (in real time) from motorway traffic counters. The traffic data are recorded by the traffic counters on the motorways and sent to the Swiss Federal Roads Office (FEDRO) servers in real-time. These actual traffic data (in minute resolution) are now available in real time via ODPMS, the customer information platform for public transport and individual mobility in Switzerland. Thus, the number of passing vehicles, their speed, and the vehicle category recorded by detectors are continuously fed into the running simulation via SUMO's TraCI. TraCI provides access to a running road traffic simulation and allows for retrieving values from simulated objects and manipulating their behavior during simulation run-time [96]. To the best of the author's knowledge, there is no research on simulating dynamic motorway flows in SUMO using external flow inputs, so one can consider this research a novelty in this field. In particular, the capabilities of SUMO *calibrators* objects to make the traffic demand within the running simulation scenario adjustable (controllable) via TraCI are leveraged, and the importance of these objects in creating the proposed synchronized DT-GM is highlighted. In addition, based on calibrators, the Dynamic Flow Calibrator (DFC) mechanism that continuously calibrates traffic flows and re-routes vehicles in the running simulation is proposed. By the use of DFC the proposed *continuously synchronized* microscopic simulation-based motorway DT is updated (in a spatiotemporal manner) as the physical motorway traffic changes in the areas of installed traffic counters. This brings the created DT-GM

closer to the actual dynamics of the Geneva motorway, as it continuously adapts itself in real time to the changes in the physical motorway dynamics. Fortunately, access to fine-grained real-time traffic data allows DT-GM to sample three types of random variables (traffic volume, speed, and vehicle classes) that govern traffic flow. With a sampling frequency of one minute, DT-GM uses almost the instantaneous probability distributions of the traffic variables (the *natural distributions*, so to speak), rather than using estimated distributions, as is the case with conventional offline simulations. By doing so, DT-GM creates the calibrated simulation frames (nearly exact replicas of real traffic) along the *live* simulation path.

Since DT-GM is based on microscopic simulation, it allows the simulation of the entire motorway system's traffic dynamics, including simulation of traffic control strategies. Moreover, since it takes into account actual real-time traffic data, DT-GM can be integrated into the decision-making process in TM on motorways. Thus, DT-GM can serve as additional real-time feedback to TM, providing not only actual traffic but also traffic conditions predicted by simulations that can be useful in a safety-critical decision-making process. Consequently, a better understanding of traffic behavior and the impact of different control strategies on the spatiotemporal evolution of traffic can be verified by DT-GM and predicted clearly enough in advance, through the detailed microscopic run-time simulation analysis before the control strategies are deployed in real system.

This implies assessing control strategies in an online yet safe manner. Thereby, the concept of parallel Digital Twin Instances (DTI) is introduced. By DTI, simulations that are initialized from the running DT-GM are assumed (see Figure 3.2). Once initialized, they can be run much faster than real time (since simulation allows for that) and, therefore, can be used to anticipate the behavior of the system under a specific control strategy. Such a concept (DTI-GM) may encourage the integration of advanced (*adaptive*) control technologies like complex machine-learning models [54, 94] in real-time TM, thus, pushing the new (*black box* machine learning) algorithms closer to the real world traffic control application on motorways. In this way, DT-GM provides a foundation for strategic TM [97], an approach to ensure overall system performance by confidently selecting a good strategy among several competing strategies.

This fundamentally changes the paradigm of traffic simulations in general. Hence, it makes preparation for the development of the *continuously synchronized* simulations providing a run-time framework for real-time safety analysis, rather than the offline approach that has been applied so far in predominated conventional simulation methods. Accordingly, in this chapter, a conceptual application of DT-GM in real-time TM on motorways regarding the VSL deployment is presented. Additionally, the DT-GM model will be demonstrated *in action* and verified by a virtual simulation field test in chapter 7 of this thesis.

To sum up, this chapter provides detailed methodological steps for the integration of different

technologies in developing of DT-GM, particularly the open data platform ODPMS and simulator SUMO, and presents the possible conceptual applications of DT-GM in motorway traffic control. Additionally, the proposed DT-GM model differs from other introduced DT traffic models in the sense that it is a run-time calibrated digital microscopic simulation model continuously updated with actual high-resolution traffic data collected directly from motorway traffic counters every minute. This makes it possible not only to visualize actual traffic in real time, but also to simulate traffic during run-time and use simulation models for detailed traffic forecasts and early detection of traffic anomalies. In addition, simulation enables the testing of control strategies (e.g., variable speed limits on motorways). Thus, DT enables the run-time safe testing of control strategies considering the current traffic situation in the most cost-effective way. This allows for using traffic simulations for dynamic control optimization and instantaneous traffic analytics during system run-time, thus, providing a framework for an early warning detection of a system (or process) failure. This was not feasible in the past when simulations were generally considered synonymous with offline validation and optimization using historical or artificial data. As such, DT-GM raises the current simulations' capability to the next level and lays the foundation for further research on DT in the transport domain. Thus, developing a simulation technology on motorways that can be fused together with real-time measurements from motorways' sensors or with the Internet of Things (IoT) in general.

Therefore, the contribution of this chapter includes [98]:

- novel paradigm in microscopic simulation traffic modeling, a virtual microscopic simulation based DT instance of the physical Geneva motorway continuously calibrated as real traffic changes over time;
- the importance of open, real-time mobility data in advancing and promoting DT in transportation, emphasizing fine-grained actual traffic data streams from traffic counters provided by ODPMS;
- the detailed methodological process of the development of DT-GM, emphasizing the calibration features of SUMO, which allow (dynamic) continuous calibration of the running simulation scenario via DFC;
- real-time simulation-based support in safety-critical decision-making in transportation systems based on predictive analytics through the concept of parallel DT instances.

3.3.1 Physical twin - Geneva motorway

The section of the A1-motorway located in the Geneva region is used as a base for the DT-GM microscopic simulation modeling. The approximate observed motorway length in both directions is about 13,200 *m*. As shown in Figure 3.1, the motorway network topology consists

of a major grade-separated interchange with junctions to the east (center of Geneva), south (border with France), and north (toward Geneva airport). The main sections of the motorway consist mainly of two or three lanes. Furthermore, the motorway contains four on-ramps and two off-ramps (see Figure 3.1b). It is worth noting that the model contains slight simplifications. For example, some minor on-ramps and off-ramps, for which ODPMS does not provide data, were not included in the modeling in order to reduce the unknowns in the traffic flow model (3.1). As a result, they are ignored in this model due to the small amount of traffic they generate.

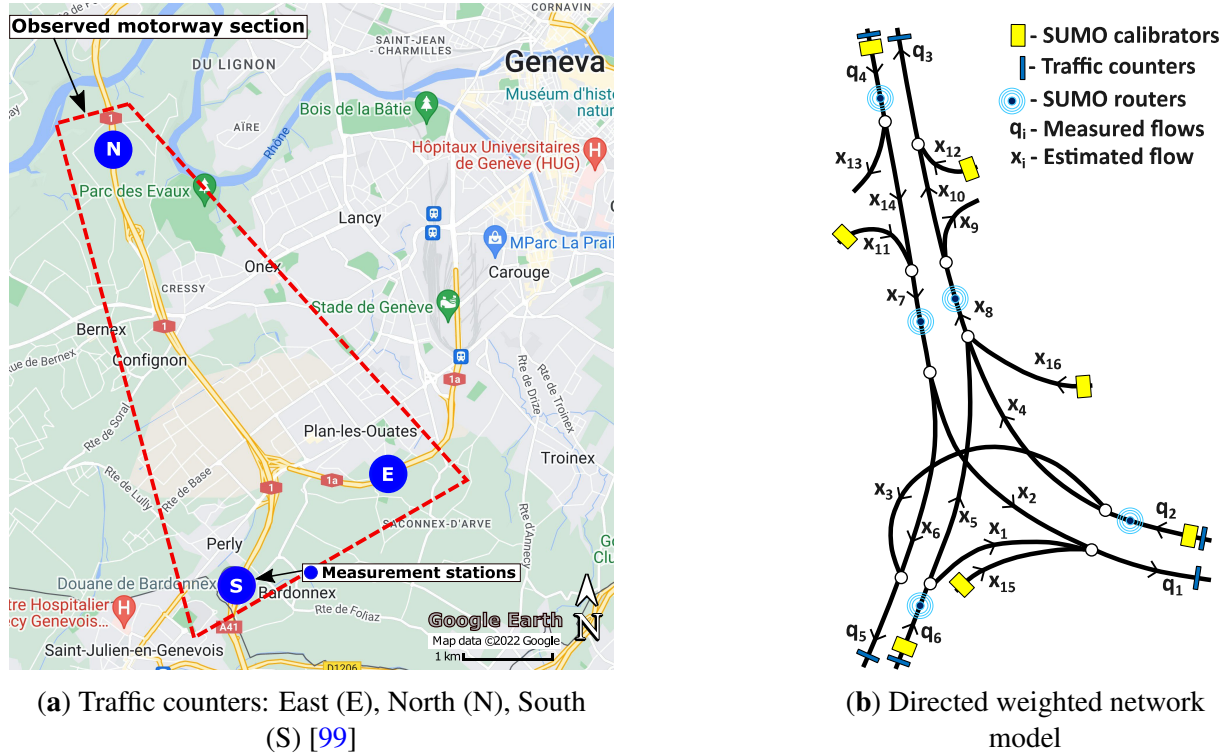


Figure 3.1: Physical motorway (left) and abstract corresponding model (right)

Although the experiment in this chapter does not aim to solve existing problems on an observed motorway section, the most striking feature is the occurrence of congestion once in the morning (and in the afternoon) in the southern region between Switzerland and the French border due to daily commuter traffic over the selected motorway network. This is the reason for choosing this particular location. As such, it will later serve as a benchmark for research on traffic optimization using DT-GM simulations and DWL-ST-VSL in chapter 7.

3.3.2 Digital twin of Geneva motorway

In Figure 3.2, the concept of run-time synchronized microscopic simulation-based DT motorway framework is presented. It consists of the physical world, in this thesis the selected part of the Geneva motorway (Figure 3.1a), the information technology that ensures the communication between the local computer and the ODPMS platform in a continuous process of sending

requests and receiving the new traffic data collected by the traffic counters on the motorway. On the local computer, the running microscopic simulation is continuously calibrated as the dynamics of traffic flows on the physical motorway change in real time. As an example of the foundation of DT-GM, one can include the existing TM system in the proposed scheme, which can use the DT model for the safety-critical decision-making process and can serve as an accurate virtual traffic model for the training and evaluation phase of learning-based traffic controllers. The proposed scheme provides a conceptual foundation for the bidirectional interaction of the virtual model with the physical motorway in real-time and is elaborated in more detail in chapter 7.

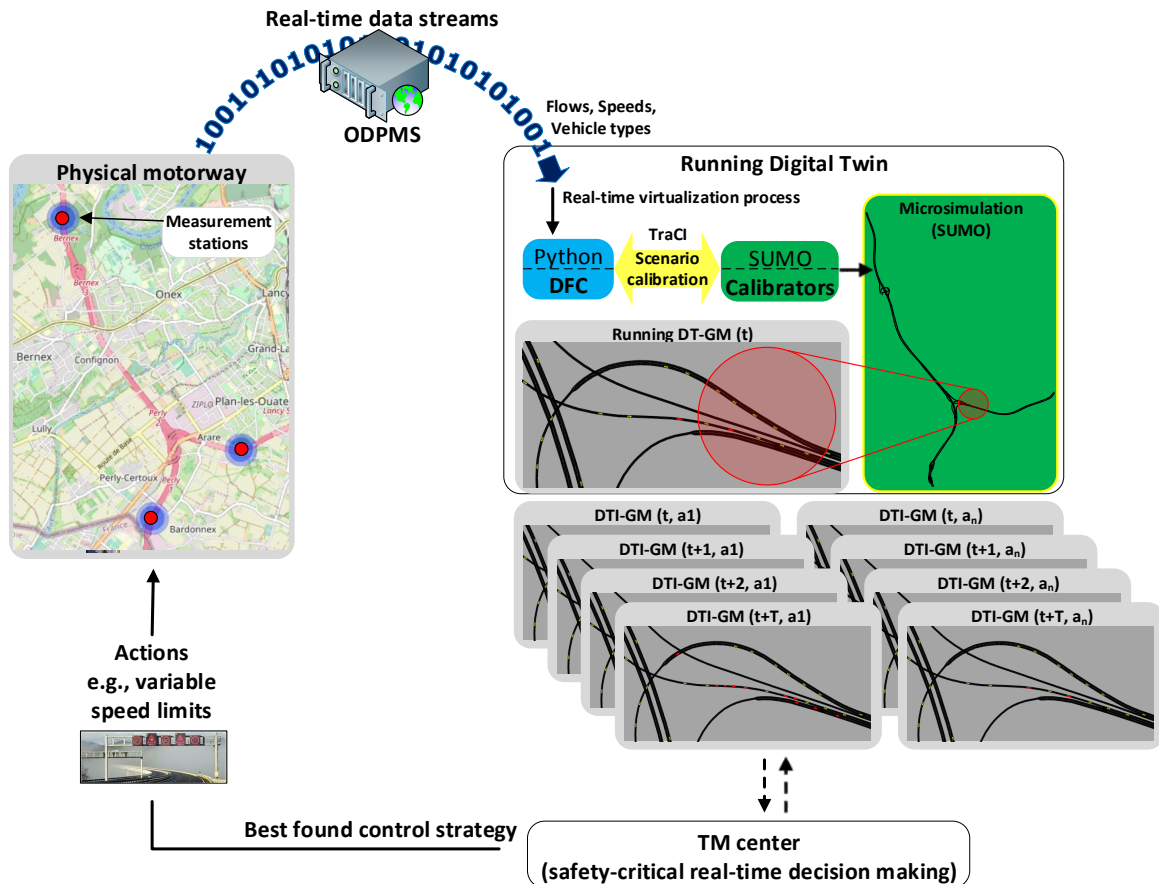


Figure 3.2: Scheme of run-time synchronized DT-GM and concept of bidirectional interaction between virtual and physical motorway using DTI-GM in real-time TM

3.3.3 Concept of parallel DTI-GM in TM

In any control optimization problem, the goal is to move the system along a desired path, that is, to move it along a desired trajectory of states. In most states, one must choose among several actions. The actions in each state essentially define the trajectory of states the system is likely to follow [18]. By analogy, the optimization of VSL on motorways involves choosing the right

actions (the desired speed limits) for particular traffic states [100]. For example, the objective of VSL is to slow down and harmonize the incoming traffic in the congestion zone in order to relieve the congested sections of the motorway and restore the traffic flow to a steady state if possible.

Thus, DT-GM generally provides a basis for TM as additional support (real-time feedback) for observing and predicting the system response in the safety-critical decision-making process through the bidirectional interaction (action/reaction) between DT-GM and real motorway traffic. In this way, DT-GM and DTI-GM (parallel instances of DT-GM) provide a basis for investigating a motorway's system performance issues during its run-time and developing potential improvements, with the goal of generating valuable insights that can then be applied to the real motorway in real-time. This process is depicted in Figures 3.2, and 3.3, where each DTI-GM (running faster than real-time) enters the running scenario with a different action a_n that is deployed from the current time t . By doing so, the boundaries can be extended (time horizon $t + T$) to examine what might happen in future steps, i.e., to evaluate/predict the impact of a given traffic control action on traffic dynamics without the risk of negative effects on physical motorways.

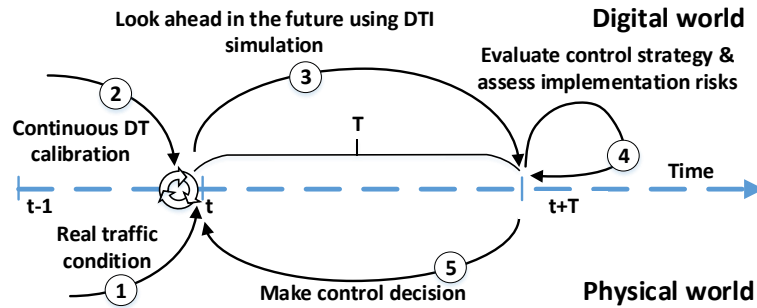


Figure 3.3: Predictive analytics with DTI

In this way, the best traffic control action among many can be selected based on certain objective criteria (e.g., what speed limits are appropriate for the next few minutes on the main sections of the motorway to increase throughput) and safely deployed in the real motorway environment. In general, one can anticipate problems with DT-GM and parallel DTI-GM. If one can detect problems early enough, there are more options (control strategies) to mitigate them and optimize motorway performance. This has not been possible with conventional offline simulations and is not discussed in this way in the existing literature. Thus, the proposed DT concept fundamentally changes the application of microscopic simulation tools in traffic analysis by enabling real-time simulation-based control optimization and analysis of traffic during the motorway's run-time.

3.4 Simulation synchronization with run-time motorway traffic

In this section, the focus is on the capabilities of the SUMO microscopic simulator for dynamic online calibration of traffic flow via TraCI. Therefore, it describes the methodological steps on how SUMO's calibrators are set up and can be used to dynamically generate the desired traffic volume and continuously calibrate the simulated traffic scenario as corresponding physical traffic changes in real-time at specific locations and, thus, become essential building blocks for the proposed DT-GM. Calibrators' locations in the simulation model represent sources that emit the desired number of vehicles on the main section of a motorway and at on-ramps. Thus, they enable the straightforward use of SUMO to create a microscopic DT of motorway systems calibrated in real-time.

Therefore, in the following part, the properties of SUMO's *calibrators* objects are introduced and their use in conjunction with dynamic routing via TraCI by proposing the DFC mechanism is highlighted. Thus, DFC is used to calibrate the running DT-GM model in order to approximate the real motorway dynamics at the locations of available measurements and over the other observed motorway parts. It is also important to note that the technical details of the proposed DT framework implementation are given in the paper [101], in which detailed technical descriptions and steps on how to build a DT of motorway traffic are provided.

3.4.1 Dynamic simulation calibration using actual traffic data streams

What distinguishes the described analysis of the proposed DT approach from the existing literature is the use of available real-time data streams obtained directly from the traffic counters implemented on the motorways. For motorways and national networks, traffic demand is measured in Switzerland by FEDRO.* Traffic counters are set up along major road sections to measure all traffic movements by direction and time. They can count vehicles (per lane), classify vehicles, and determine the speed of vehicles passing a particular location. The actual traffic data are updated every minute and refer to the traffic movements in the last full minute. The data records are, thus, aggregated over 1-minute intervals and recorded every minute on FEDRO's server, and can be accessed in real time through the ODPMS platform exclusively via an available API.† Essentially, the number of vehicles and average speed are reported in two categories: (i) light vehicles such as cars, motorcycles, buses, and small delivery trucks, and (ii) heavy goods vehicles such as trucks and trucks with trailers. These capabilities are leveraged by obtaining real-time traffic data from traffic counters on the motorway in the Geneva region

*<https://www.astra.admin.ch/astra/en/home.html>

†<https://opentransportdata.swiss/en/rt-road-traffic-counters/>

to build the DT-GM. Circles in the region of interest (see Figure 3.1a) represent the locations of installed traffic counters. A single (blue) dot represents multiple sensors (individual sensors per direction and lane). Therefore, one can create a bounded linear model (3.1) for both directions of travel that can estimate unknown traffic flow at the edges between the boundaries of E , N , and S .

This allows to feed the proposed DT-GM model with raw traffic data in real-time taken from the physical system. Moreover, with this fine granularity of data, the original traffic pattern (distribution of traffic flows) is preserved, since the distribution of the real system in spatio-temporal terms is directly contained in the fine-grained traffic data. In contrast to the current state-of-the-art, where a longer aggregation period is used, and, thus, the simulation approach requires an estimation of the distributions of the governing variables in order to simulate and analyze the system at a finer scale. Namely, suppose the aggregation period is one hour, and traffic is generated based on the estimated distribution. In that case, some intervening events (in a finer time resolution) may not be adequately captured by the model, thus, increasing the error in the simulation.

Moreover, in conventional simulation modeling, the internal calibrated features become dominant for a particular scenario defined in the simulation. For example, the calibrated scenario for a specific weekday may not accurately simulate traffic on other days of the week, non-working days, or holidays, and vice versa. In other words, the samples generated by the simulator, and, thus, the expectations of the random events, may vary widely. Thus, conventional simulation models need to account for the different traffic characteristics of the weekday and holiday scenarios separately to overcome the problem. It is also true that if the model is calibrated based on aggregated weekly traffic data, the accuracy of the simulated traffic events at daily resolution (even worse at hourly or minute resolution) may contain significant errors.

As DT-GM is based on continuous, instantaneous sampling and calibration, it, therefore, reduces the sources of error in the model (the finer the granularity, the smaller the error). Therefore, the DT-GM model can adapt online in the spatiotemporal domain and closely resemble the underlying structures in the evolution of real traffic for each minute (regardless of the day of the week) with high accuracy.

Run-Time Flow Estimation

In practice, it would be very costly to survey the traffic flows on all roads, as a large number of traffic counters would have to be installed and maintained. Instead, if the traffic counters are placed in a way that they capture most of the incoming and outgoing traffic, it is possible to reconstruct the traffic flows in between. Thus, from partial data, one can define a system of linear equations using the conservation law of traffic flow and calculate the unknown traffic

flows in a road network. In the context of a motorway, the traffic situation can be simplified, as it is safe to assume that the roads are *one-way* and that a vehicle entering the observed network also leaves the network (no terminal states within the motorway network). In reality, however, a vehicle may linger for a while at a gas station or a public facility such as a rest area, but this can be neglected in the context of the proposed DT-GM modeling. The flow in the network is balanced, i.e., the total flow entering the network is equal to the total flow leaving the network. This is true for all branches in the observed network (Figure 3.1).

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \\ x_{15} \\ x_{16} \end{bmatrix} = \vec{X}_p + \vec{X}_n = \begin{bmatrix} q_1 \\ 0 \\ q_5 \\ q_2 - q_5 \\ q_6 - q_1 \\ 0 \\ 0 \\ q_2 - q_1 - q_5 + q_6 \\ q_2 - q_1 - q_3 - q_5 + q_6 \\ q_3 \\ 0 \\ 0 \\ q_4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{pmatrix} 1 & -1 & 0 & -1 & -1 & 0 \\ -1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{bmatrix} x_6 \\ x_{11} \\ x_{12} \\ x_{14} \\ x_{15} \\ x_{16} \end{bmatrix} \quad (3.1)$$

The complete solution (3.1) of the flow balance in the flow model contains the particular solution \vec{X}_p (where q_1, \dots, q_6 are measured flows by traffic counters) and the special solution \vec{X}_n . For the special solution, the free variables $\vec{X}_{free} = [x_6, x_{11}, x_{12}, x_{14}, x_{15}, x_{16}]^T$ can be chosen arbitrarily. However, from a practical point of view, all traffic flows must be non-negative. This implies adding constraints (3.2) on the complete solution, thus, limiting \vec{X}_{free} .

$$\vec{X} \geq 0 \quad (3.2)$$

Since six free variables are defined and need to be estimated in such a way that one must obtain positive solutions, a linear program to search for feasible solutions using the Simplex algorithm is defined. Thus, in order to find the solution to the given problem, two steps of the Simplex algorithm are performed. In the first step, the inequalities from (3.2) are used as constraints,

while randomized cost coefficients of the objective function were used in each Simplex run. In this way, the extreme points of the feasible region for free variables \vec{X}_{free} can be defined. Based on the calculated bounds for each free variable, the positive interval is defined starting with the minimum positive value and ending with the maximum value. Also, the interval is divided into 10 parts so that each free variable is defined with ten intensity levels. Once the feasible ranges are known, the solution space is further restricted by defining the desired intensity vector of the free variables $\vec{X}_{free-des}$. By doing so, the range of each free variable is further restricted with respect to the desired intensity level. Such a newly constrained problem is again solved by several Simplex runs. The best found solution for the free variables \vec{X}_{free} is determined by calculating the minimum relative error using the equation (3.3) between the vector of the desired intensity and the given set of feasible solution vectors.

$$\frac{\|\vec{X}_{feasible} - \vec{X}_{free-des}\|}{\|\vec{X}_{free-des}\|} \quad (3.3)$$

After all flow variables are computed, DFC calculates the routes' distribution, which is used to distribute the vehicles (flows) generated by calibrators throughout the network in order to satisfy (3.2).

SUMO's Calibrators Objects

In SUMO, calibrators (*trigger-type objects*) enable the modeling of location-dependent changes in traffic flow dynamics and driving behavior. Once defined in the initial simulation scenario, they allow for dynamic adjustment of traffic flows and speeds along with changes in vehicle parameters by assigning different predefined vehicle types.

Calibrator Definition

Each calibrator is uniquely associated with a particular edge (part of a road in the created simulation model) or even by a particular lane on the edge on which it is placed in the simulation model. In order to use the calibrator, it is necessary to define the interval (start, end), which specifies the time during which the calibration takes place. Thus, the interval length defines the aggregation period for comparison of the observed and desired flows. The calibration goal is to ensure that the correct number of vehicles is deployed at the end of the respective time interval, and also in that particular place. At the same time, the space-time structure of the existing traffic should be preserved as much as possible [31]. Thus, a calibrator removes vehicles that exceed the specified traffic volume and inserts new vehicles (of the specified type) when traffic demand in a simulation does not reach the specified number of vehicles. Also, vehicles can be assigned with the desired speed if adjustments are required. This means that the calibrator in its basic

configuration works like a static object and modifies the traffic flow according to the predefined attributes mentioned above: traffic flow [veh/h], speed [m/s] and vehicle types within certain time intervals. Thus, different flows and speeds can be used for different time intervals during the simulation.

Run-Time Calibration of Simulation via TraCI

Moreover, SUMO provides the ability to access the calibrator via the TraCI interface while the simulation is running. This enables the activation of a specific calibrator and the adjustment of flow rate, speed, and even vehicle type in the current time interval during the execution of an actual simulation. In the performed experiment, intervals for calibrator adjustments are set to match the frequency at which actual traffic data is received from the motorway sensors (one minute).

The calibrators are placed in the created microscopic simulation model at the positions where their corresponding traffic counters are implemented in the real motorway (see Figure 3.1). These calibrators are, thus, used to continuously adjust the flows to match the current traffic demand as it changes in the real motorway with a minute resolution. For more technical details on the setup and technical implementation of the calibrator in DT-GM, one can refer to the paper [101], as mentioned above.

Dynamical Rerouting of Vehicles

In this section, the routing mechanism in SUMO is explained in a rather abstract way, and in the next part, details of the implementation are elaborated. The realism of traffic flow behind or between calibrators depends on the correspondence between simulated and real routes. This correspondence importance increases with the network's size and complexity between calibrated edges [31]. Rerouting along with calibrators allows for the desired traffic flow to be distributed in a desired proportion across the network to satisfy other routes' traffic demand (routes defined by particular edges). Thus, once the calibrator starts inserting the desired traffic flow, it should be distributed across the network by assigning routes to individual vehicles. Routes are assigned to a vehicle by the *routing device*, which is an abstraction of the location (a particular edge in the simulated motorway network) monitored by DFC during simulation run-time. Accordingly, DFC assigns the desired route to the vehicle based on the calculated route distribution when the vehicle drives onto that edge.

Principle of Dynamic Flow Calibrator Mechanism for DT-GM

Given the locations of the physical traffic counters (Figure 3.1a), the simulation model is designed to use the available real-time traffic measurements as much as possible. Therefore, the

boundary points of DT-GM are defined by the locations of the traffic counters. At these locations, the model is equipped with calibrators from which the desired traffic flow (measured by traffic counters) is distributed using DFC to match the traffic flow on other observed motorway segments (Figure 3.1b). Thus, DFC uses information about all possible routes from a given point. Subsequently, the information about the traffic flow on the edges (calculated by (3.1)) is used to compute the probability distribution of the traffic flow among all possible routes. Routes and route distributions that are dynamically assigned to vehicles using TraCI are predefined in conjunction with the calibrators. This forms the base for introducing the DFC mechanism to reroute the traffic flows over the observed motorway segment. Accordingly, the number of vehicles that need to continue their direction of travel or switch to other routes is calculated. This process is illustrated in Figure 3.4. Given the initial traffic flow q_0 with speed v_0 arriving at the model's starting point, the calibrator adjusts the traffic flow and speed according to the real-time measurements provided by the traffic counters on the real motorway. The modified traffic flow (q_1, v_1) continues the trip according to the predefined initial route $Route_0$. When the vehicles reach the point where the routing device $Router_1$ is installed, they are assigned (according to the calculated probability) the new route corresponding to the desired traffic flows on the next parts of the network (computed by (3.1)). For example, the probability P_{x_1} that a vehicle is assigned to traffic flow x_1 (on $Route_1$) is calculated as follows:

$$P_{x_1} = 1 - P_{x_2} = 1 - x_2/q_1, \quad (3.4)$$

where P_{x_1} and P_{x_2} are calculated every minute and transmitted via TraCI to a particular edge on which the routing is performed. The probabilities of assigning a vehicle to a particular route among two possible routes are defined as an ordered pair (P_{x_1}, P_{x_2}) and, thus, can take discrete values whose sum equals one (see [101]). Thus, the process of rerouting assumes that the desired probability distributions have been predetermined. This rerouting solution follows the same principle for the entire motorway model in areas with multiple possible directions of travel.

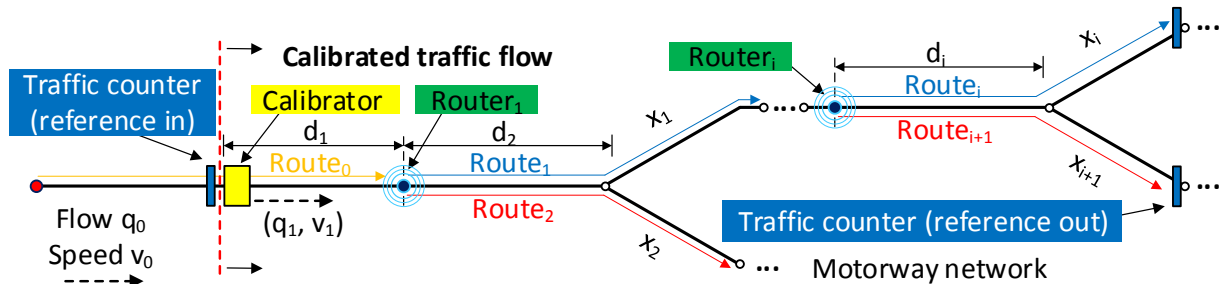


Figure 3.4: Dynamic Flow Calibration principle in SUMO

An additional *marker* is added to each vehicle when it is inserted into the simulation. This

contains the information about the routes assigned to it according to the calculation performed by DFC for that time window. This prevents vehicles that have already been inserted from being assigned the newly calculated routes. Specifically, the vehicles that arrive late in the router area (which may be the case in traffic jams when the vehicles are stuck for a certain period of time) are not assigned with the newly calculated route distribution, but they follow the original route plan (encoded in *marker*) that was set when they were inserted in the simulation. In this way, the traffic flow maintains a balance between the flow (in) and (out) in the DT-GM model, otherwise, a mismatch could occur and the flow model (3.1) could not be satisfied.

The summary of the execution of the DT-GM microscopic simulation model is represented by the Algorithm 3.1. Once the variables are initialized, the algorithm loads the simulation scenario \vec{S} with all necessary SUMO files. The simulation is started and controlled by TraCI. At each simulation step, SUMO calibrates and updates the simulation scenario. For every multiple of 60th seconds of real-time, a new request for actual traffic data is sent to the FEDRO server via ODPMS. The data (traffic flows \vec{q} , speeds \vec{v} , and vehicle type \vec{v}_{type}) are received and forwarded to the DFC mechanism. In this way, the calibrators \vec{C} are entrusted with the task of generating the desired traffic flows. Moreover, DFC computes all traffic flows \vec{X} in the network and, accordingly, the route distribution probabilities \vec{P} to distribute the traffic flows generated by the calibrators through the network such that \vec{X} is satisfied for a given time window of one minute. The process is repeated until the specified end, saving the status of DT-GM. Also, the process can be reloaded by restoring the saved simulation state and running it from that point. Additionally, when the real-time TM concept is enabled, new parallel simulation instances are initialized and started by loading the states of the running DT-GM. Such parallel DTI-GM simulations (running faster than real time) are associated with different control strategies from \vec{T} to test each one and predict their impact on motorway traffic during system run-time.

Algorithm 3.1 DT-GM at each simulation step

```

// Set parameters and load sim. model
Init  $\vec{S}, \vec{T}, \vec{C}, \vec{P}, \vec{q}, \vec{v}, \vec{v}_{type}, \vec{X}, \vec{X}_{free}, \vec{X}_{free-des}$ 
for each simulation step
    if simulation time % 60 [s] == 0 then
        Get new actual traffic data via ODPMS:  $\vec{q}, \vec{v}, \vec{v}_{type}$ 
        // DFC computations
        For given  $\vec{X}_{free-des}$  and  $\vec{q}$  calculate  $\vec{X}_{free}$  and  $\vec{X}$  (see (3.2))
        Update calibrators  $\vec{C}$  and routes distributions  $\vec{P}$  using  $\vec{X}$ 
        // Conceptual run-time analysis in real-time TM
        if  $TM_{active} == \text{True}$  then
            Save current DT-GM simulation state
            Initialize new sim. using the current state of running DT-GM
            Run DTI-GM
            Test different control strategies from  $\vec{T}$ 
            Return best strategy
            Deploy the best found strategy on the physical motorway
        end if
    end if
    Calibrate and update simulation scenario
end for
Save DT-GM simulation state and close simulation

```

3.4.2 Experimental setup

To validate the proposed DT approach, the created DT-GM is compared with a baseline—the actual traffic measured by the installed traffic counters on the motorway in the Geneva region. The simulation framework used in the DT-GM development and simulation experiments consists of the microscopic simulator SUMO, Python programming environment, and the ODPMS server accessed remotely every minute (Figure 3.2).

Simulation Model

To create a simulation model with the exact replica of the geometry of a real road network, geometric elements with appropriate attributes of geographic data from the free geographic database of the world OpenStreetMap (OSM) [102] must be converted into the simulation network model. For this purpose, the NETEDIT module from SUMO is used to create a digital motorway network. The network model consists of interconnected edges representing roads. Within an edge, a lane is specified according to the real motorway topology. Connecting the edges by nodes forms a complete SUMO motorway network. Since the OSM website limits the size of the region to be extracted, additional steps are taken to obtain the final OSM file from which the SUMO motorway network is generated. More technical details on extracting the network from OSM can be found in the previously mentioned paper [101].

Once the motorway network is defined, the traffic flow can be specified, e.g., as repeated emissions of vehicles (*flows*) together with dynamic traffic flow generated by calibrator objects during simulation run-time. Since the calibrator's functions are essential building blocks in the creation of DT-GM, this main component is therefore explained in detail in the following experimental settings. The additional definition of *flows* serves as initial traffic flow in the model and their purpose in DT-GM is again explained in [101].

DT-GM Parameters

It is important to note that the purpose of the demonstrated real-time simulation experiment is not to fine-tune internal simulation parameters, e.g. in terms of driver behavior or vehicle parameters, but rather to investigate the extent to which the proposed DT-GM with the basic simulation configuration (mainly considering default parameters) resembles real traffic by using the underlying motorway system information through fine-grained real-time traffic data and run-time calibration of traffic flow with DFC (Figures 3.2, and 3.4). Therefore, the observation between real and simulated traffic should be considered mainly as a comparative measure of how closely the model resembles reality, rather than an absolute measure of performance.

SUMO Simulation Parameters

Traffic flow data consists of cars and trucks, as these classes were distinguished by traffic counters. At the moment of creating DT-GM, the latest SUMO version was 1.13.0. The longitudinal (speed choice) vehicle behavior uses the Enhanced Intelligent Driver Model (EIDM) car-following model [103], while lateral (lane changing, overtaking, merging) behavior uses the lane change model LC2013 [104]. Additionally, several parameters within the mentioned models were modified. The parameter *tau* defines the time interval between the passage of successive vehicles at the observed road point (*time headway*), measured in seconds. It is defined by the lognormal distribution (with the shape parameter $\sigma = 0.05$ and the location parameter $\mu = 0$) [105, 106]. The headway distribution model is shown in Figure 3.5. In this way, stochastic effects are introduced into the car-following model to better suit the behavior of real drivers. Similarly, adjustments have been made for the lane change parameters (see [101]). The simulation step used in the analysis is set to a quarter of a second, which means that four simulation steps lead to one simulated second. Also, nominal maximum allowable speed limit values on Swiss motorways (120 [km/h]) for main sections were assumed, as access to speed limit information via ODPMS is under development. Finally, actual traffic collected on Thursday, March 24, 2022, and the results of the correspondent simulated traffic replica by run-time DT-GM are summarized in the next section.

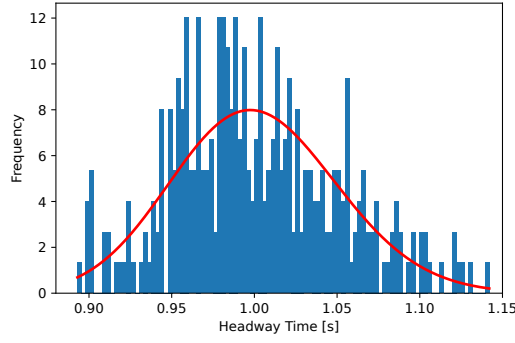


Figure 3.5: Lognormal headway distribution model

Flow Model Parameters

As explained in section 3.4.1, the free variables in the flow model, $x_6, x_{11}, x_{12}, x_{14}, x_{15}$ and x_{16} are defined by intensity levels rather than absolute values. Therefore, the defined intensity level for these variables is approximated by observing the GPS traces on the OSM and Google Maps traffic website as a function of time of day, and experimenting with run-time test simulations. Therefore, they should not be considered as an exact representation of the actual traffic volume on the corresponding motorway sections (edges), but only as an approximation. Although they are an approximation of the actual traffic volume, mathematically, they represent a stable, feasible solution that clearly satisfies the system (3.2). Therefore, eight different intervals with respect to the time of day and corresponding desired intensity vectors $\vec{X}_{free-des}$ for free variables $x_i, i = 6, 11, 12, 14, 15, 16$ are defined as follows:

- from 6 to 8 am with values 10, 1, 2, 7, 3, 1, respectively;
- from 8 to 10 am with values 10, 1, 2, 4, 3, 1, respectively;
- from 10 to 12 am with values 10, 2, 2, 5, 2, 1, respectively;
- from 12 to 4 pm with values 6, 1, 1, 10, 1, 1, respectively;
- from 4 to 6 pm with values 6, 1, 3, 8, 1, 3, respectively;
- from 6 to 8 pm with values 7, 2, 2, 8, 1, 2, respectively;
- from 8 to 10 pm with values 7, 2, 2, 7, 1, 2, respectively;
- for the rest of the day with values 7, 1, 1, 7, 1, 2, respectively.

Also, the Simplex algorithm is performed in two steps, as discussed above, when solving the inequality to find a feasible solution for the system (3.2). In each step, it is executed and solved 300 times, each time with a random initialization of the coefficients of the objective function to find the feasible bounds for x_{free} (see section 3.4.1).

3.4.3 Results and analysis

This section presents the overall joint analysis of results related to the methodology used in the development of the DT-GM by testing the accuracy and timeliness of workday traffic simulated at run-time via DT-GM. Flow analysis is used to analyze the spatiotemporal behavior of DT-GM adaptation with respect to measured real eastbound, northbound and southbound traffic (flows-in and flows-out). Thus, graphs given in Figure 3.7 show the evolution of the flows for the selected workday. Similarly, for traffic flow indicators, the corresponding *GEH* statistics are shown. The *GEH* statistic is a formula used in traffic engineering, traffic forecasting, and traffic modeling to compare measured (real) traffic volumes with model-generated traffic (simulation). In addition, Figure 3.9 shows the spatiotemporal snapshot of the evolution of traffic flow speeds of the actual run-time traffic on the physical motorway and flow speeds in the run-time simulation performed by DT-GM. Thus, the focus is mainly on spatial and temporal correspondence between the DT-GM model and real traffic.

Traffic Flow Analysis

The overall traffic demand in the observed motorway area for the analyzed workday has a symmetric characteristic. Traffic generated by commuters from France has a major contribution to the morning peak hours eastbound direction to Geneva (see Figures 3.6 and 3.7d).



Figure 3.6: Illustration of study motorway area

Additionally, the eastbound is partially loaded by commuters from the north, which use the motorway as a bypass to avoid driving through the city center in order to reach the southern part of Geneva, and by traffic entering the motorway at the on-ramp x_{15} (Figure 3.1b).

Similarly, for traffic on northbound (Figure 3.7e), traffic generated by commuters partially from southbound (flow-in) from France (Figure 3.7c), and eastbound (flow-in) from Geneva (Figure 3.7a) have a major contribution to morning peak hours northbound across the main sections. However, traffic entering the motorway at the on-ramp x_{12} also contributes to the overall picture.

The peak hours in the afternoon have an opposite character, i.e., the commuters return, and the peak hours are symmetrically reversed. The high traffic flow is mainly present in the southbound motorway sections towards the Swiss-French border (Figure 3.7f), partly caused by commuters from eastbound of Geneva and the main motorway traffic from the north.

Comparison between DT-GM and Real Traffic

As can be observed from the results (the graphs are created using the moving average over 20 measurements), there is no significant difference between the traffic generated by the run-time DT-GM and the actual motorway traffic. In the one-day-long simulation on March 24, 2022, both the morning and afternoon peak traffic and the stationary traffic between the peak hours simulated by run-time DT-GM align with the actual traffic conditions at the granularity of one minute.

The small temporal shifts between measured and simulated flows (outflow part) are collectively attributed to the fact of the simplification in the model, where some local motorway on-ramps for which ODPMS does not yet provide traffic data were omitted. Therefore, additional traffic is generated on the included on-ramps (x_{11} , x_{12} , x_{15} and x_{16}) to achieve the required overall traffic flow balance on the observed motorway segment. In addition, the spatial displacement of the calibrators (*traffic emitters*) on main sections in the east, north, and south compared to the symmetric measurement points of flow-out may result in a small delay in the response time of the flow model (3.1). Consequently, the time to reach the desired flow on the exit sides of the motorway section may be slightly shifted to the right (Figures 3.7d, 3.7e, 3.7f), because vehicles have to travel a longer distance to reach the desired points. Nevertheless, while a rightward shift is expected, it's important to note that this occurs on a very small time scale, as the run-time adjustment of DT-GM is based on minute intervals. Thus, even though the obtained data can be *timely* slightly shifted, the achieved result is still pretty accurate.

In addition, the time at which the output values first reach the steady state of desired traffic volumes for a given simulation may be slightly delayed, especially during peak hours. Consequently, additional delayed traffic in the network may show up as a slow increase (or decrease) between the simulated traffic and the reference traffic measured by the traffic counters (Figure 3.7). However, the cumulative error is negligible since one is still at a fine-grained temporal resolution.

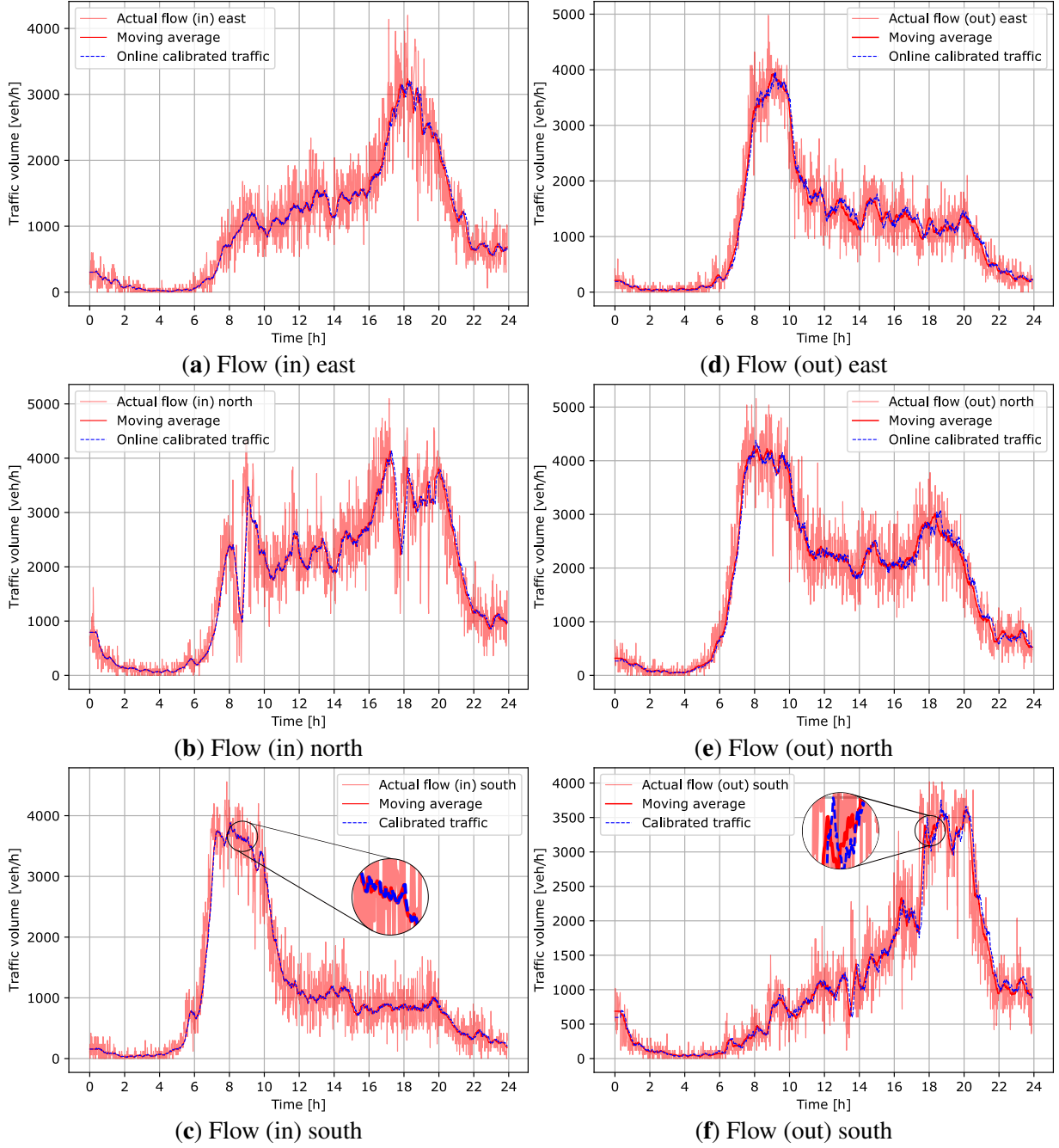


Figure 3.7: Comparison of actual and simulated daily flow with minute resolution

Furthermore, outliers can cause a significant bias in commonly used numerical measures such as the average. Therefore, the Box-Whistler diagram (Figure 3.8) is presented for actual run-time traffic and dynamically calibrated traffic. It gives a rough summary of the data distribution using five numbers: the smallest traffic volume, the lower quartile $Q1$, the median, the upper quartile $Q3$, and the largest traffic volume. Even though there are no recorded errors in received actual run-time data, a data set contains measurements that differ markedly from the others in the set, both east and south (in/out) actual traffic flow (upper fence for outliers: $Q3 + 1.5(IQR)$, where $IQR = Q3 - Q1$). Without going into the nature of the outliers (observational errors or the fact

that outliers themselves may contain important information about the traffic dynamics), those similarities mostly reflect dynamically calibrated traffic flow. The most striking outliers are found in the eastern and southern regions. The main reason could be unexpected traffic patterns due to traffic lights in Geneva (east) urban areas and uncertain traffic dynamics in the south at the border crossing, accompanied by high commuting between these regions. It is also noticeable that in the east, for flow (in) (Figure 3.8a), the calibrators were not able to reproduce outliers at all. The reason could be that the calibrators cannot completely generate higher sudden traffic volume in a short interval due to their specific waiting properties explained in the discussion part. Thus, this additional traffic is generated in subsequent intervals to compensate for the previously incurred loss. In this way, all traffic is maintained, but with some delay. Additionally, if traffic is congested in the lane, there may not be enough room to insert a vehicle safely by the calibrator. Therefore, the calibrators may only insert a portion of the vehicles rather than the required traffic volume. This ensures that invalid congestion is not passed through a calibrator. Such behavior can be controlled in SUMO and should be further investigated, as such congestion can occur on motorways and, thus, poses an issue for the current calibration design and implementation.

However, the calibrators reflect the maximum traffic flow values (in) and (out) in the east and south. In contrast, significant differences in maximum values are observed for traffic in the north, explainable by the same arguments mentioned above. Since the medians are to the right of the center of the box in all cases, the distributions are skewed to the left, i.e., there are some small measurements of traffic volume caused by low traffic at night. Even so, there are high traffic volumes due to peak hours, particularly in the northern region. Thus, all flows are characterized by high variability. Nonetheless, these phenomena are generally quite well reproduced by DT-GM at the granularity of one minute.

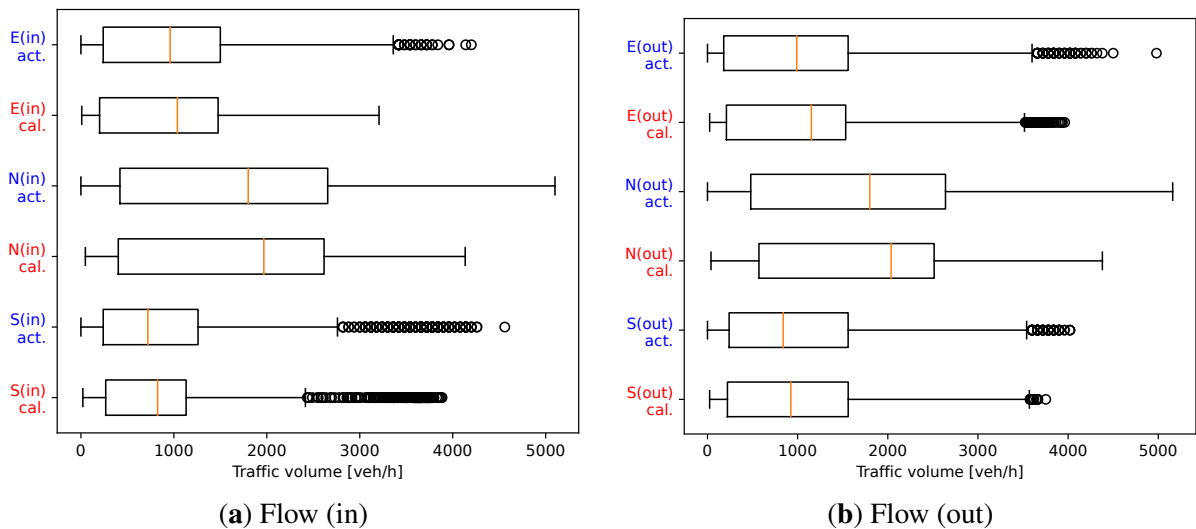


Figure 3.8: Box-Whistler diagram for actual and calibrated traffic at minute resolution

Finally, a numerical evaluation of the deviation between the graphs of actual run-time traffic and traffic simulated (and calibrated) on-the-fly by DT-GM is presented in the next section.

Quantitative Workday Simulation Analysis

The traffic flows generated at the calibrator locations (flow in) match almost exactly the actual traffic (see Figures 3.7d, 3.7e, and 3.7f). Therefore, the GEH statistics calculated by equation (3.5) (Tables 3.2, 3.3, and 3.4) are presented only for the case of outflows to show the accuracy of the model with respect to the proposed DFC mechanism used for distributing the traffic generated by the calibrators over the simulated motorway network.

$$GEH = \sqrt{\frac{2(m-c)^2}{m+c}} \quad (3.5)$$

Thus, a comparison is made between the measured traffic flow leaving the network (flow out) and the measurements recorded by the traffic counters. It can be seen that the *GEH* values are within the acceptable range of 0 to 5 (a *GEH* value of less than 5.0 is considered a good match between simulated dynamic *m* and observed actual run-time traffic volumes *c*). The highest values are observed during peak hours when congestion events delay traffic in the middle parts of the motorway network by increasing the travel time of vehicles. Even though the running simulation is continuously fed (every minute) with new traffic data from the traffic counters, the traffic simulation is well run-time calibrated. Thus, DT-GM allows the scenario of running simulation to evolve as the conditions on the physical motorway in the areas of the traffic counters.

Table 3.2: GEH STATISTIC FOR FLOW (OUT) EAST

Morning time (<i>h</i>)	0	1	2	3	4	5	6	7	8	9	10	11
GEH $\sqrt{veh/h}$	0.0	0.0	0.5	0.8	0.6	1.6	2.8	3.0	0.0	1.3	1.5	0.7
Afternoon time (<i>h</i>)	12	13	14	15	16	17	18	19	20	21	22	23
GEH $\sqrt{veh/h}$	0.9	3.5	0.4	0.7	1.1	0.9	0.9	0.2	2.3	2.1	1.8	0.7

Table 3.3: GEH STATISTIC FOR FLOW (OUT) NORTH

Morning time (<i>h</i>)	0	1	2	3	4	5	6	7	8	9	10	11
GEH $\sqrt{veh/h}$	1.3	0.1	0.6	0.0	1.8	1.6	2.4	1.9	0.4	0.8	0.6	0.2
Afternoon time (<i>h</i>)	12	13	14	15	16	17	18	19	20	21	22	23
GEH $\sqrt{veh/h}$	0.3	0.2	0.1	0.1	0.0	1.3	0.2	3.3	4.1	1.5	1.6	1.5

Table 3.4: GEH STATISTIC FOR FLOW (OUT) SOUTH

Morning time (h)	0	1	2	3	4	5	6	7	8	9	10	11
GEH $\sqrt{veh/h}$	0.0	0.2	0.6	0.3	0.9	1.7	0.0	0.9	0.0	0.7	2.4	0.3
Afternoon time (h)	12	13	14	15	16	17	18	19	20	21	22	23
GEH $\sqrt{veh/h}$	0.7	1.0	1.5	0.2	0.1	2.2	0.2	0.4	3.5	3.8	1.4	0.1

Space-Time Analysis

In Figure 3.9, a space-time diagram of vehicle speeds in the motorway network is shown. Speeds are color delineated, ranging from low speed (dark-red) to free flow speed (green). The spatiotemporal color map is obtained from Google Maps[‡] (Live traffic) on Wednesday, March 30, 2022, at 5.15 pm, while the colors in the run-time DT-GM model are defined as follows: dark-red for slow traffic (speeds less than 50 % of free flow speed), red for 75 % of free flow speed, orange for 90 % of free flow speed and green for free flow speed regarding the nominal speed limit of the edges. Despite the similarity, the comparison with the baseline (speed presented by Google Maps (real traffic)) should be considered primarily as a playground for comparison between the physical traffic dynamics and simulated traffic by DT-GM rather than an absolute measure of performance since exact values for matching the colors with speeds are not available. Particularly, the aggregation period and the length of a segment on which vehicle speeds are aggregated in Google Maps are not available, and one can assume only nominal speed limits on the main section. Such visualization reveals DT-GM potential for estimating traffic dynamics on motorway segments for which direct measurements are not available. For example, DT-GM can be used to identify and position traffic sensors to get a better overall estimation of network traffic or to account for and compensate for sensor failures. In addition, it can estimate time-varying origin-destination matrices by observing the fraction of traffic flow on the path with dynamic flow screening of flows. Such visualization can, thus, serve as a mechanism to alert authorities to revise critical sensor locations in the network. It can also serve as a benchmark for validating other approaches for estimating real-time traffic flow, e.g., estimation methods based on cellular network data. In addition to traffic flow under normal conditions, DT-GM may also reflect traffic anomalies, such as the effects of accidents on traffic flow. When such an anomaly occurs near the traffic counter, the travel speed of vehicles is drastically slowed, which is reflected in the running simulation by a color change from green to red. The reflection time depends on the distance between the location of the accident event and the traffic counter. In addition, it is likely that changes in speed limits and, therefore, in traffic flow dynamics can be reflected by DT-GM in the same way. However, this requires additional systematic research and is beyond the scope of this thesis.

[‡]<https://www.google.com/maps/@46.1680111,6.1025731,6847m/data=!3m1!1e3!5m1!1e1>

Nevertheless, one can observe a high variability of speeds in the motorway network near the main sections, where two or more sections merge into one and vice versa, and in the areas of on-ramps and some off-ramps, as well as at the Swiss-French border, caused by a lower capacity of the border crossing infrastructure. These are accompanied by strong vehicle interactions when traffic volumes are high (Figure 3.9). Such phenomena trigger congestion activation. Once the congestion is triggered and the traffic volume is high, it usually spreads upstream to the main sections. In the main sections, it has also been observed that delayed lane change maneuvers disrupt traffic flow as vehicles urgently slow down and sometimes even stop and wait for a gap in the right lane to perform a safety maneuver to change lanes (areas of merge sections or off-ramps) and continue on the desired route. This requires further analysis as a partial cause of this may be the inherent features of the car-following and lane change models and their parameters used in the simulation experiments, but this is beyond the scope of this thesis.

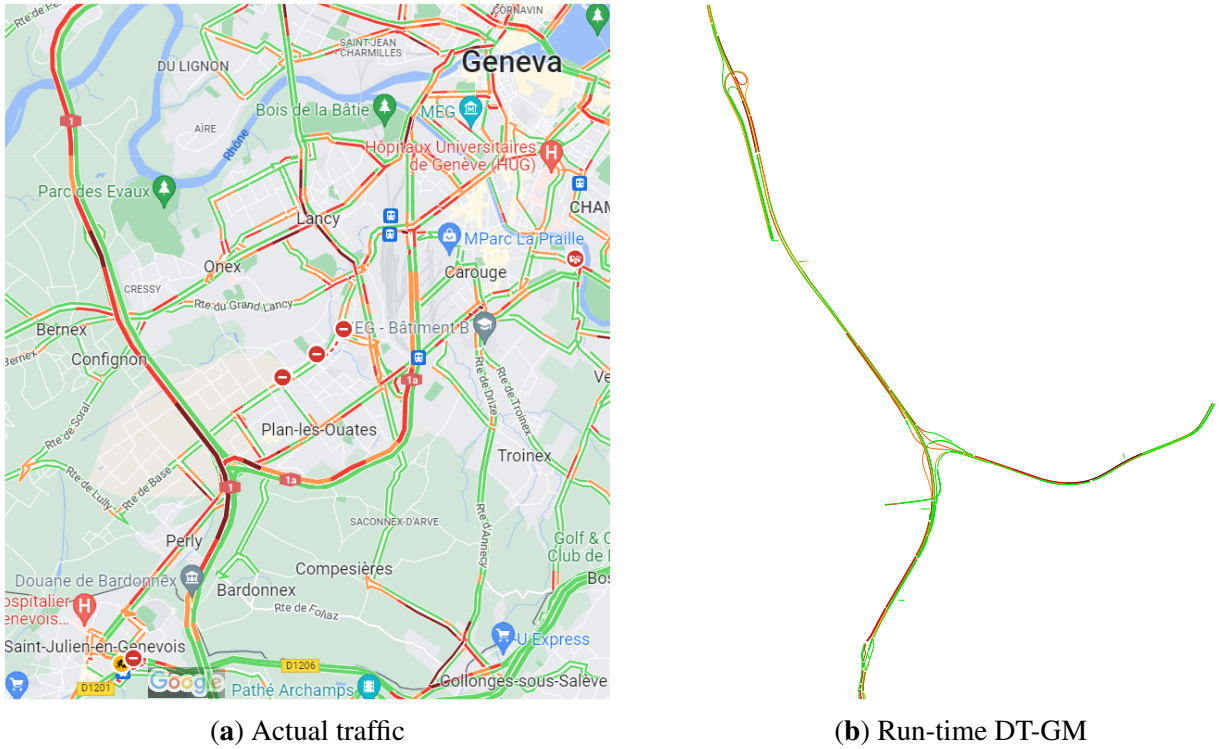


Figure 3.9: Comparison of spatiotemporal speed distribution between actual traffic and run-time DT-GM

3.4.4 Discussion

With a foundation of actual traffic data provided by ODPMS, this research presented and tested, for the first time, a *live* motorway DT replica in the microscopic simulator SUMO. Moreover, preliminary results have shown that the proposed DT-GM itself is able to accurately reflect real traffic dynamics with a very fine temporal resolution during system run-time.

Therefore, in this section, based on the overall joint analysis and the obtained results, the possible further directions for enhancing the proposed DT-GM model are highlighted. It is also worth noting that the comparison with the baseline (real traffic on the studied motorway section) should be considered primarily as a benchmark for comparison between real traffic and the newly proposed DT-GM, allowing for critical review in terms of uncovering gaps and room for possible improvements in design and further development of DT-GM, rather than as an absolute measure of performance.

In general, it can be concluded that the run-time DT-GM model almost exactly replicates the actual traffic in areas of traffic counters on the analyzed section of the Geneva motorway, which supports the methodological approach used to create DT-GM. Furthermore, this confirms that SUMO, which was mainly used for offline motorway simulations, is also suitable for run-time (online real-time) simulations. Thus, DT-GM itself (and its underlying technologies) can be used to map real traffic to a virtual microscopic simulation-based DT model during the run-time of a motorway system.

Furthermore, results indicate further directions for DT-GM enhancement. Information about vehicles' instantaneous location and dynamics is imminent in the context of pervading CV and CAV technology. The further granularity of traffic data (ideally, the event-based data) that ODPMS could provide in the foreseeable future would form the basis for generating instantaneous traffic demand inputs in the running simulation. This would allow for the original distribution of traffic flow on the motorway to be maintained during run-time, allowing DT-GM to evolve as an exact replica of the physical correspondence. In addition, this information will enable an online calculation of parameters, e.g., the distribution of the headway parameter, which is a fundamental microscopic traffic parameter within the car-following model that reflects driver behavior and traffic flow characteristics. Ultimately, headway is a dynamic parameter, i.e., depending on the traffic flow state and roadway characteristics [107]. Thus, further analysis is required to define an adaptive headway over time and space rather than a single constant value [106]. Additionally, the lane change model parameters that define lane change, merging, and overtaking need to be further analyzed (e.g., the look-ahead parameter that allows modeling the delay in driver reaction time [30], which is still experimental to some extent in SUMO), as the current simulation setup mainly uses default parameters of the LC2013 lane change model [104]. Additionally, an in-depth analysis of the parameters of the EIDM car-following model [103] is recommended for future work.

In addition, speed limits need to be taken into account because they strongly influence vehicle dynamics and, thus, traffic flow parameters (throughput, density, etc.) on a given motorway section. Therefore, speed limits are of great importance when modeling traffic flow dynamics on motorways. In the current model, the motorway's real speed limit system could not be

included in the created DT-GM and, therefore, could not be utilized. Instead, the predefined (static) speed limits for Swiss motorways were used for the modeling. The lack of speed limit information across the network could affect the traffic dynamics and, thus, represent a cause for potential errors in the DT-GM if the speed limit changes significantly over time while nominal speed limits are considered. This will be further investigated and also included in the model once ODPMS provides speed limit data.

Moreover, conditions such as weather can affect traffic behavior on motorways. Thus, DT-GM can take into account not only information from sensors about traffic flow on the motorway (traffic counters), but also sensors from the environment, i.e., sensors that describe the conditions under which traffic flow operates. All of these changes can, therefore, be captured by a run-time DT in order for a model to accurately reflect traffic changes and provide better insight into the monitored motorway system. Thus, an interesting future work direction would be to investigate the possible integration of hidden, latent variables of environmental sensors, such as weather conditions or daytime and nighttime visibility conditions on the motorway [108]. In addition, the road maintenance component with information about lane closures or road surface conditions should also be integrated into the microscopic DT model.

So far, it has often happened that *calibrators* wait until the end of the interval and then generate a higher number of vehicles to meet the desired traffic volume on edge, which might form vehicle platoons. This can be observed with short intervals (one minute long calibration period) when the traffic volume that needs to be calibrated by the calibrator varies a lot. For example, when the calibration interval is one minute, the dominance of platoons in moderate traffic is not apparent. However, if the calibrators are demanded to generate a higher traffic volume, they may create a platoon that affects the ongoing traffic in an unrealistic way. Thus, it would be a good way to enable the calibrator to scan the few edges upstream to get an approximate insight of the amount of traffic that will pass in the current calibration interval so that it can uniformly (or using other distributions) generate the desired amount of vehicles during the calibration interval. This is expected to change in the future with developments in SUMO.

Even when traffic counters on motorways provide accurate information about the traffic flow for a given micro location, the current limitation is the sparse coverage of the motorway with traffic counters. Especially where there are a larger number of on- and off-ramps that are not equipped with traffic counters. A satisfactory DT model is subject to two conflicting requirements. It must be sufficiently detailed to represent actual traffic with relative accuracy, and simultaneously, it must be simple enough to make a real-time simulation analysis practical. In this manner, the traffic flow models (3.1) and (3.2) for estimating traffic volumes on unknown routes can be further simplified by integrating new traffic counters on the aforementioned ramps. This might form the basis for a more stable solution (fewer unknowns) of the traffic flow model and, thus,

increase the accuracy of DT-GM even more. This is critical to accurately reflect changes in traffic dynamics across the motorway while reducing computations when the DT-GM model is extended to a larger motorway network.

So far, the final solution of (3.1) is affected by the manually defined desired level of intensities ($\vec{X}_{free-des}$) for free variables in the model (see section 3.4.2). Since different distributions of flows can be observed over the weekdays, further development of defining the adaptive mechanism for $\vec{X}_{free-des}$ is needed to match traffic intensity on inter-mediated network elements. In particular, the bottleneck in DFC is the search of the space for feasible bounds on \vec{X}_{free} , which is performed by the Simplex algorithm and relative error formula. So far, simplifications of the model have been kept to a minimum in order to preserve most of the traffic dynamics (further simplifications reduce the number of unknowns in the model but at the expense of the accuracy of DT-GM). Thus, it has been found that the limiting number of free variables in DT-GM is 6, for which DFC can still find the system's feasible solution (3.2) in a reasonable time ($\approx 1-2$ [s]).

In addition, indications from the design phase and modeling of DT-GM imply that it would be possible to make this mechanism dynamic, i.e., to adjust (reduce or increase) the number of Simplex runs depending on the measured error between \vec{X}_{free} and $\vec{X}_{free-des}$ and defined error threshold. Also, results indicated that it is better to use the formula for the relative error (see section 3.4.1) between a vector of the desired intensity and the vectors from the given set of feasible solutions to search for the best \vec{X}_{free} . Comparing the mentioned vectors using the cosine distance gives worse results during peak periods. The above shortcomings may pose a challenge for large-scale DT simulation of a more complex road network. One possibility to improve the scalability of DT-GM is to use the *divide-and-conquer* approach. It can parallelize the simulation processes of each motorway region and use the (in/out) flow of a particular region as the (out/in) flows for the simulation process of the neighboring regions and vice versa. However, this raises the issue of synchronizing the execution of parallel simulations. Finally, a possible extension of the used mathematical tool (Simplex algorithm) with, e.g., genetic algorithm or advanced machine learning algorithms together with the fusion of traffic flow data with other available data sources, such as mobility behavior through smartphone positioning data [89] could be a promising direction for future work in enhancing the robustness of DFC and, thus, DT-GM itself. IoT [91] and Big Data [90] in urban mobility raise the question of a synergy of DT microscopic simulation and AI technology in the future.

3.5 Concluding remarks

This chapter presents a comprehensive methodological process for developing a motorway DT. Comparing this approach with existing research on recent advances in the development of DT

in transportation leads to several important findings. First, it fills a gap in the literature by presenting DT-GM: a novel microscopic simulation-based digital twin of the Geneva motorway implemented in the microscopic simulator SUMO. Secondly, the results show that the proposed DFC mechanism exploits the full potential of SUMO to dynamically adapt the running traffic flow scenario by leveraging the SUMO's calibrators objects used to calibrate the traffic flow at run-time via the TraCI interface. Thus, DT-GM itself continuously adjusts the run-time traffic scenario as the spatiotemporal traffic changes on the real motorway. Finally, the proposed and created DT-GM model uses the newly available fine-grained real-time traffic data received every minute from the traffic counters on the Geneva motorway via ODPMS. This allows for comparing the performance of the created DT with the actual motorway traffic at run-time. Experimental results confirm the reliability of DT-GM, as it reflects the actual traffic with high accuracy. This means that DT-GM responds to the physical motorway with relatively low latency, which is made possible by the current development of ODPMS.

As such, DT-GM provides the basis for a wealth of new concepts in TM that were not possible before. DTI-GM parallelization provides a basis for the safe evaluation of control strategies during the run-time of the physical system, i.e., before the control scheme is applied to the actual physical system. Therefore, in chapter 7, the application of DT in optimizing VSL in a safety-critical decision context using an advanced learning-based control algorithm is investigated. This will provide data-based evidence of the performance of DWL-ST-VSL in the context of real-world traffic and enable systematic evaluation of complex control processes involving traffic uncertainties, which is particularly important for the development of adaptive traffic control systems, in particular, those based on learning.

Chapter 4

Dynamic spatiotemporal adaptation of variable speed limit zones

The most effective way to explain the model and motivation for the dynamic adaptation of VSL zones is to systematically review what has been learned so far and which simulation-based experiments will be performed to confirm the hypothesis described in the Introduction. To do this successfully, it is necessary to present the work in a logical and orderly manner, as suggested by the following sequence of steps: (i) simulation-based optimization of VSL zone locations has to be performed, (ii) to assess whether there is a trade-off between bottleneck optimization and system performance, an additional multi-criteria analysis that includes different configurations of VSL zones and speed limits is performed, (iii) according to synthesized results of other researchers and results obtained in simulation experiments new conjecture about the interpretability of the need for dynamic VSL zones adaptation through the concept of uncertainty in traffic flow capacity is introduced, and, finally, (iv) the model for dynamic adaptation of VSL zones that will be used in the needed simulation-based experiments in chapters 6 and 7 is proposed.

4.1 Optimizing VSL zones allocation

Before implementing VSL in a real system, it is critical to determine the length and location of VSL zones for optimal VSL performance. The length of the VSL zone and its distance from the bottleneck directly affect traffic dynamics and thus bottleneck control [28]. Therefore, this section analyzes the impact of different lengths and positions of VSL zones on traffic performance in the context of bottleneck optimization using a feedback VSL control system based on the simple proportional speed controller (SPSC-VSL). The different VSL zone configurations and their impact on traffic flow control and vehicle emissions were evaluated in SUMO

on a high traffic demand scenario. The results support the observations of previous researchers on the significant dependence of VSL zone configuration on VSL efficiency. In addition, this analysis provides new simulation-based traffic and vehicle emissions data as evidence of the performance of the SPSC-VSL design in terms of the best placement of successive VSL zones for motorway congestion control.

4.1.1 Case with feedback VSL controller

Simple proportional speed limit controller

One of the desirable characteristics of VSL controllers is their stable behavior under fluctuating traffic conditions. At the same time, it is desirable to have a less complex controller so that it can be implemented and tested in a simulation framework and easily adjusted for real-world deployments. Therefore, the SPSC-VSL was proposed in [109]. SPSC-VSL is a simple feedback-based VSL controller that can respond appropriately to changes in downstream density on controlled motorway sections. To release a disturbance in the downstream motorway section $N+1$, N -controlled sections that are upstream of the section $N+1$ would be active and respond to their downstream density changes using equation (4.1). For the case of the used simulation model section $N+1$ represents section L3 where the bottleneck occurs, see Figure 4.1. Therefore, SPSC-VSL attempts to adjust the speed limit of the upstream flow to reduce the inflow to the downstream motorway region to minimize the differences between the densities measured from the previous control time step $t - 1$ and the currently measured density on the controlled motorway segment. In equation (4.1), the difference of the sum of measured densities of all affected downstream sections from the previous control time step $\rho_j(t - 1)$ and the sum of current densities $\rho_j(t)$ with respect to section i is added to the speed limit value from the previous control time step $V_i(t - 1)$. The impact of density changes on the new speed limit $\bar{V}_i(t)$ can be set by the positive proportional gain K_v :

$$\bar{V}_i(t) = V_i(t - 1) + K_v \left[\sum_{j=i+1}^{N+1} \rho_j(t - 1) - \sum_{j=i+1}^{N+1} \rho_j(t) \right]. \quad (4.1)$$

To achieve a smooth speed transition and avoid large fluctuations between two successive speed limits, the new speed limit in section i is bounded by equation (4.2). The parameter C_v stands for the maximum allowable speed limit change between two successive control time steps and is set to 20 [km/h].

$$V_i(t) = \begin{cases} V_i(t-1) - C_v, & \text{if } \bar{V}_i(t) \leq V_i(t-1) - C_v \\ V_i(t-1) + C_v, & \text{if } \bar{V}_i(t) \geq V_i(t-1) + C_v, \\ \bar{V}_i(t), & \text{otherwise.} \end{cases} \quad (4.2)$$

The flaw of the SPSC-VSL controller is the frequent change in speed limits due to the nature of equation (4.1). To further stabilize SPSC-VSL, an additional variable C_i [veh/km/lane] is introduced to control the activation of the controller. If the density in the adjacent section ($i+1$) at control time t is greater than the predefined activation threshold C_i , recalculation of the speed limit in section i is required. Both the activation variable and the controller gain values were selected from several simulation tests.

Nonetheless, SPSC-VSL is not subject to the curse of modeling (e.g., it does not require the flow-speed fundamental model which might be difficult to estimate for traffic flow on a motorway due to its time-varying nature). SPSC-VSL is also highly scalable; there is no limit to the number of motorway segments that can be controlled. Accordingly, multiple adjacent VSL zones upstream of the congestion area can achieve a smooth speed transition between the downstream congestion and the free flow arriving at the bottleneck [45]. However, to further optimize SPSC-VSL, an analysis of the appropriate length of the VSL zones and their positions is required. Therefore, an analysis methodology for efficient spatial placement of VSL zones in terms of their length and starting point from the bottleneck is presented in the next section.

Analysis methodology

As mentioned above, in this section, the methodology used in required experiments is described in detail regarding the impact of different VSL and acceleration zone lengths on VSL efficiency. The VSL zone is referred to as an area where the speed limit is controlled by a VSL controller. In the simulation experiment, two VSL zones (L1 and L2) controlled by the SPSC-VSL controller (Figure 4.1) are used. To ensure optimal bottleneck operation, vehicles must enter the bottleneck with a speed close to V_{cr} . Due to the frequently oscillating location of the creation of the bottleneck in the mainstream traffic flow within the merging area of the on-ramp r_3 , it is hard to determine the exact location of the bottleneck activation. What was observed during the simulations is that when a little activation of the bottleneck is created, it is likely that at the very beginning, the formed disturbance (congestion) will propagate upstream at least 50 – 100 [m]. As a result, the ending point of the acceleration area in the simulation model is set at the position of 5.3 [km] (100 [m] before starting position of the on-ramp r_3) so that vehicles are still able to accelerate after leaving the VSL zone and can enter the bottleneck tail with V_{cr} as the results from [28] suggested.

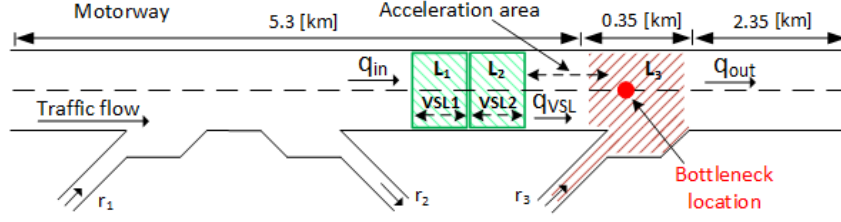


Figure 4.1: Set up of VSL zones for the feedback VSL controller

Thus, in the following experiment, three VSL zone configurations 100 – 100, 300 – 300, and 500 – 500 [m] were chosen. The first value refers to L1, while the second represents the length of the L2 zone, as shown in Figure 4.1. Additionally, three acceleration areas are selected 100, 200, 300 [m] for the experiment. The two VSL zones are adjacent and spatially adjusted so that the VSL2 zone is always close to its downstream acceleration area. Therefore, 9 different VSL configurations are used in the performed simulation experiment. It will be referred to in the following sections as follows. For example, the VSL configuration with two VSL application zones with a length of 300 [m] and an acceleration area of 100 [m] will be abbreviated as 300 – 300 – 100. The same principle applies to other configurations [110].

Motorway model and traffic data

The motorway model used to test selected scenarios is based on the simulation model used in [71]. The total length of the motorway model is 8 [km]. The mainstream consists of two lanes. It contains two on-ramps (r_1 and r_3) and one off-ramp r_2 . The total length of each ramp is 0.5 [km]. The model contains three main sections used for SPSC-VSL controlling the VSL zones (L1 and L2), and the congestion section L3. The VSL zones are equipped with a VMS on which the new speed limits are displayed depending on the current traffic conditions. The congestion is caused by the additional traffic demand on the mainstream and at r_3 , which leads to the activation of the bottleneck in the merge area. The traffic flow used in the simulation experiment was defined according to Figure 4.2. The additional traffic peak in the mainstream flow (between 42-47 simulation minutes) and the highest demand at the on-ramp r_3 lead to pronounced congestion in the merging area. It is important to note that the calibration procedure of the simulated traffic flow is not included, since a synthetic model with corresponding traffic flow was used in this analysis. The objective of this experiment is to develop a methodology to evaluate the impact of different configurations of VSL zones on traffic flow and vehicle emissions, as well as their impact on the performance of SPSC-VSL. To test the effects of VSL zone lengths on a given motorway section, appropriate traffic data must be collected and the simulation model ought to be calibrated. Then, the methodology defined in this section can be applied to a specific motorway in the same way as described here.

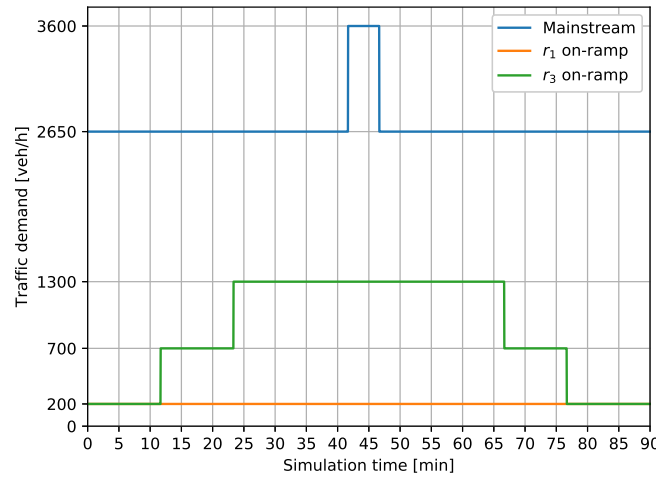


Figure 4.2: Simulated traffic demand

Simulation results

The conducted experiments were evaluated using the corresponding Measures of Effectiveness (MoEs) explained further on. The average vehicle speed and the average traffic density were measured in the area of the bottleneck (section L3). The environmental parameters measured during the simulation are CO , CO_2 , NO_x , and PM_x . Together with the macroscopic traffic parameters (density and average speed), the mentioned MoEs represent an enrichment of the information about the traffic situation since the sudden braking and acceleration of vehicles in congested traffic causes additional fuel consumption and, consequently, a higher environmental impact. The variant without VSL is used as a baseline to evaluate the obtained results.

Influence of VSL and acceleration area lengths on traffic parameters

Regarding Figures 4.3, 4.4, and 4.5, presented in this analysis, it should be noted that the measurements have been connected by using lines for better illustration. However, the part of the line that lies between the points does not correspond to any measurement.

Figure 4.3 shows the TTS measured on the entire motorway network for the simulated cases. There are 3 cases with a lower TTS than in the case without VSL (notation No VSL). The best result is obtained for the VSL configuration 300 – 300 – 100, where each VSL zone has a length of 300 [m] and the acceleration area is 100 [m]. The obtained improvement is 1.38 % compared to the case without VSL. A similar result is obtained in the case of the 500 – 500 – 100 VSL configuration. Thus, a shorter acceleration zone reduces the delay and leads to a better congestion resolution.

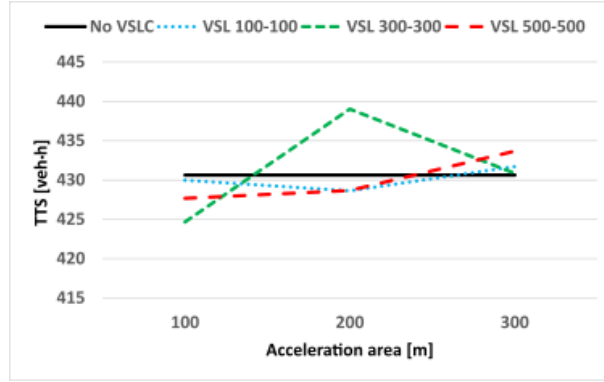


Figure 4.3: Variation of TTS with the lengths of VSL zones and acceleration zones

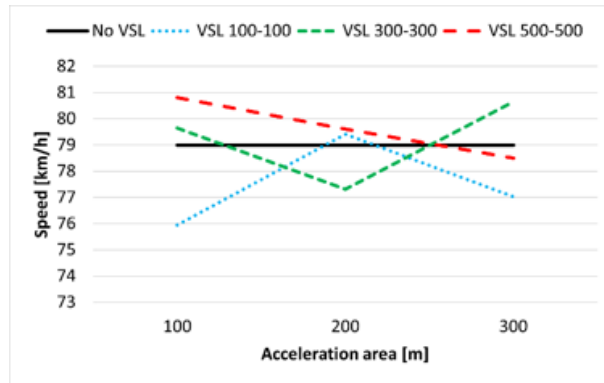


Figure 4.4: Variation of average speed in L3 with the lengths of VSL zones and acceleration zones

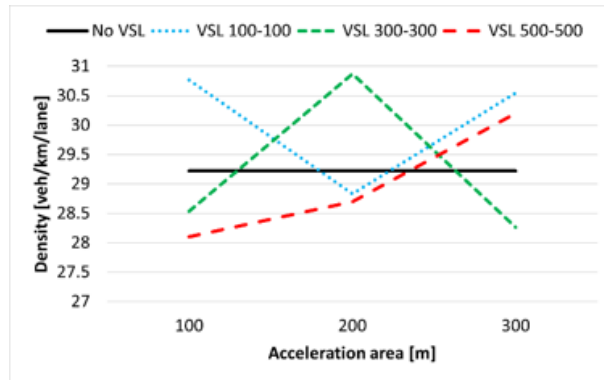


Figure 4.5: Variation of traffic density in L3 with the lengths of VSL zones and acceleration zones

Figure 4.4 shows the change in average speed in the bottleneck area as a function of VSL zone length and acceleration area. The VSL zone configuration 500 – 500 [m] with an acceleration area of 100 [m] gives the best result (notation 500 – 500 – 100). Also, the VSL zone configuration 300 – 300 [m] with an acceleration length of 300 [m] (notation 300 – 300 – 300) achieves a good performance in terms of the measured average speed in the bottleneck area. For all acceleration areas in the case of VSL zones with a length of 100 – 100 [m], the measured speed is

equal to or slightly lower than the case without VSL.

Figure 4.5 shows the average density relative to the different VSL zone configurations. The highest density of 30.87 [veh/km/lane] is obtained for the 300 – 300 – 200 scenario, which corresponds to a deterioration of –5.65 % compared to the density of 29.22 [veh/km/lane] in the case without VSL. The best case is achieved for the 500 – 500 – 100 configuration with a density value of 28.10 [veh/km/lane]. Compared to the case without VSL, the average density in L3 is decreased by 3.83 %. Also, the VSL configuration 300 – 300, where the length of the acceleration zones is 100 and 300 [m], leads to favorable results in reducing the traffic density. The most significant reduction in vehicle emissions is achieved with the VSL configuration 500 – 500 – 100. The positive impact of the controllers on CO_2 emissions is up to 1.25 %, on CO emissions up to 2.97 %, on NO_x emissions up to 1.14 % and on PM_x emissions up to 1.16 % per simulated traffic scenario. For a complete list of results of all cases analyzed, see Tables 4.1, 4.2, and 4.3.

Table 4.1: OBTAINED MOES IN CASE OF VSL ZONES LENGTH 100-100, AND DIFFERENT ACCELERATION ZONE LENGTHS

VSL zones lengths [m]	No VSL	100 - 100					
Acceleration zone length [m]		100		200		300	
		Obtained	Improv. [%]	Obtained	Improv. [%]	Obtained	Improv. [%]
TTS [veh·h]	430.63	429.99	0.15	428.65	0.46	431.74	-0.26
Avg. speed in L3 [km/h]	79	75.95	-3.86	79.41	0.52	77.03	-2.49
Avg. density in L3 [veh/km/lane]	29.22	30.77	-5.3	28.33	1.33	30.55	-4.55
CO_2 [kg]	9,473.84	9,465.31	0.09	9,427.56	0.49	9,530.76	-0.60
CO [kg]	120.31	119.99	0.27	119.63	0.57	120.61	-0.25
NO_x [kg]	20	20	0	19.8	1	20.13	-0.65
PM_x [kg]	452.33	452.07	0.06	449.2	0.69	454.4	-0.46

Table 4.2: OBTAINED MOES IN CASE OF VSL ZONES LENGTH 300-300, AND DIFFERENT ACCELERATION ZONE LENGTHS

VSL zones lengths [m]	No VSL	300 - 300					
Acceleration zone length [m]		100		200		300	
		Obtained	Improv. [%]	Obtained	Improv. [%]	Obtained	Improv. [%]
TTS [veh·h]	430.63	424.69	1.38	439.04	-1.95	430.85	-0.05
Avg. speed in L3 [km/h]	79	79.64	0.81	77.31	-2.14	80.65	2.09
Avg. density in L3 [veh/km/lane]	29.22	28.54	2.33	30.87	-5.65	28.26	3.29
CO_2 [kg]	9,473.84	9,398.6	0.79	9,480.61	-0.07	9,429.61	0.47
CO [kg]	120.31	117.75	2.13	119.72	0.49	119.23	0.89
NO_x [kg]	20	19.8	1	20.18	-0.88	19.84	0.79
PM_x [kg]	452.33	447.1	1.16	454.79	-0.54	449.73	0.57

Table 4.3: OBTAINED MOES IN CASE OF VSL ZONES LENGTH 500-500, AND DIFFERENT ACCELERATION ZONE LENGTHS

VSL zones lengths [m]	No VSL	500 - 500					
Acceleration zone length [m]		100		200		300	
		Obtained	Improv. [%]	Obtained	Improv. [%]	Obtained	Improv. [%]
<i>TTS</i> [veh·h]	430.63	427.7	0.68	428.7	0.45	433.7	-0.71
Avg. speed in L3 [km/h]	79	80.8	2.28	79.6	0.76	78.5	-0.63
Avg. density in L3 [veh/km/lane]	29.22	28.1	3.83	28.7	1.78	30.2	-3.35
<i>CO</i> ₂ [kg]	9,473.84	9,355.55	1.25	9,396.42	0.82	9,437.06	0.39
<i>CO</i> [kg]	120.31	116.73	2.97	117.9	2	118.66	1.37
<i>NO</i> _x [kg]	20	19.77	1.14	19.93	0.37	20.01	-0.05
<i>PM</i> _x [kg]	452.33	447.08	1.16	450.17	0.48	452.14	0.04

Spatiotemporal congestion analysis

Space-time diagrams allow visualization of traffic behavior over time and space (throughout the observed motorway network). For example, the spatiotemporal distribution of congestion is shown in Figure 4.6. Speeds in Figures 4.6a, 4.6c, 4.6e, and 4.6g are color-coded (from dark red for 0 [km/h] to dark blue for vehicle speeds around the nominal maximum speed of 120 [km/h]). The x-axis represents the distance from the beginning of the motorway sections $x = 0$ to $x = 8$ [km], which corresponds to the length of the simulated motorway segment. The on-ramp r_3 is located at $x = 5.4$ [km]. On the y-axis is the time and the simulation lasts 1.5 [h]. Figure 4.6a shows the vehicle speed for the base case without VSL (No VSL). The region with the yellow and red patterns corresponds to a congested area where vehicles travel at lower speeds. Due to changes in the intensity of inflow at the on-ramp, the location of the floating merge point of vehicles at the on-ramp and vehicles in the mainstream may change over time [28]. Therefore, the bottleneck in this experiment forms in the merge area from $x = 5.4$ to 5.65 [km]. The congestion starts at $t = 0.45$ [h] with the activation of the bottleneck in the merge area and propagates upstream to the position $x = 4.5$ [km]. After the traffic demand decreases, the congestion decreases and finally dissipates at about $t = 1.25$ [h]. The best result is obtained with the 300 – 300 – 100 configuration (diagram (e) in Figure 4.6), where the congested area is the smallest compared to the other configurations. A closer look at the corresponding best VSL zone setup and speed limits is shown in Figures 4.7 and 4.8. Since the speeds and densities are correlated macroscopic variables (fundamental speed-density diagram), similar patterns can be seen in Figures 4.6b, 4.6d, 4.6f and 4.6h, where orange-red represents a higher density value (congestion). Looking at these plots vertically near the position of on-ramp r_3 (merge range from $x = 5.4$ to 5.65 [km]), it can be seen that the bottleneck becomes active at about $t = 0.45$ [h]. In this case, the traffic density is about 35 [veh/km/lane] (light blue stripes in the mentioned diagrams).

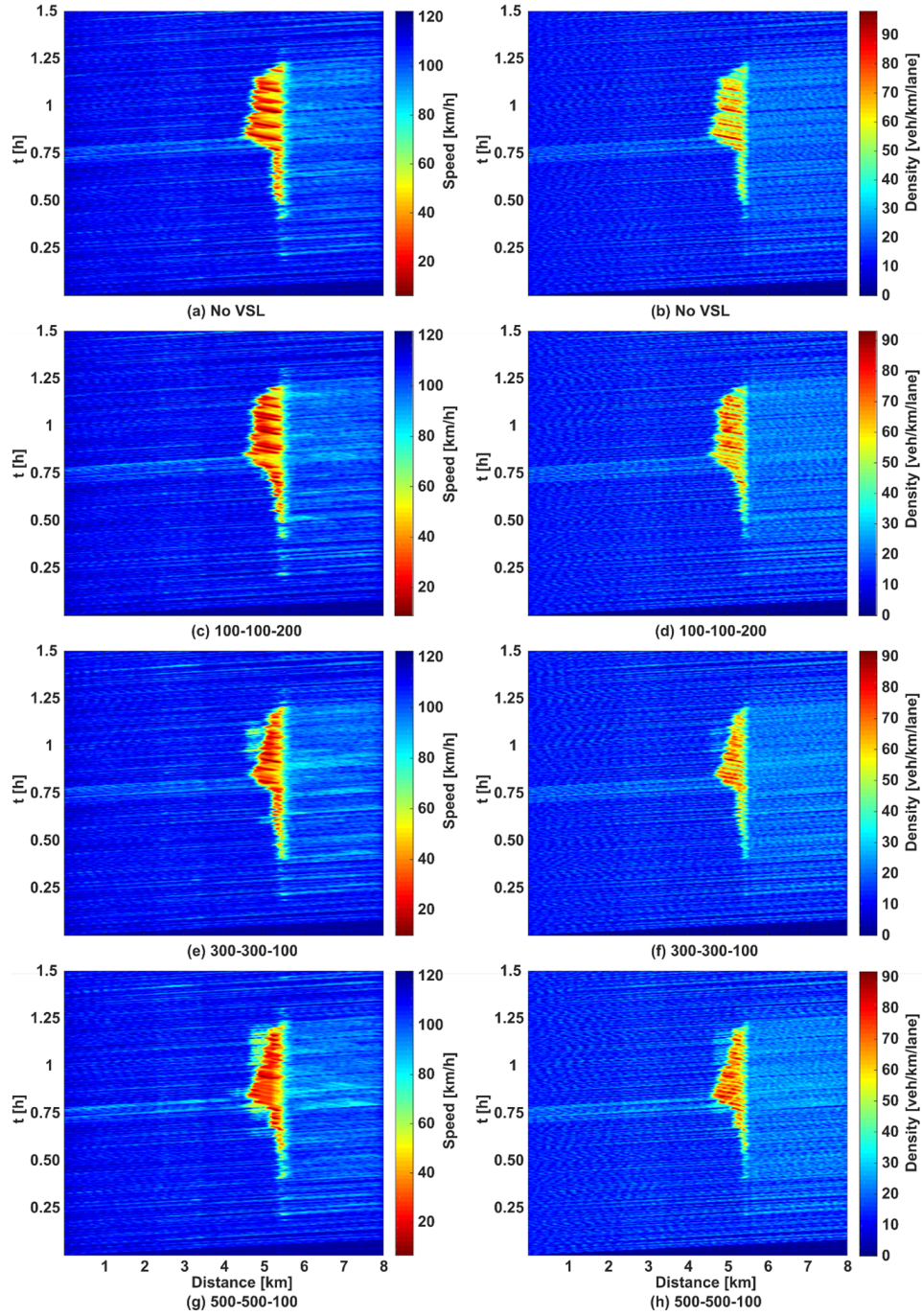


Figure 4.6: Spatiotemporal diagrams for best obtained VSL configurations

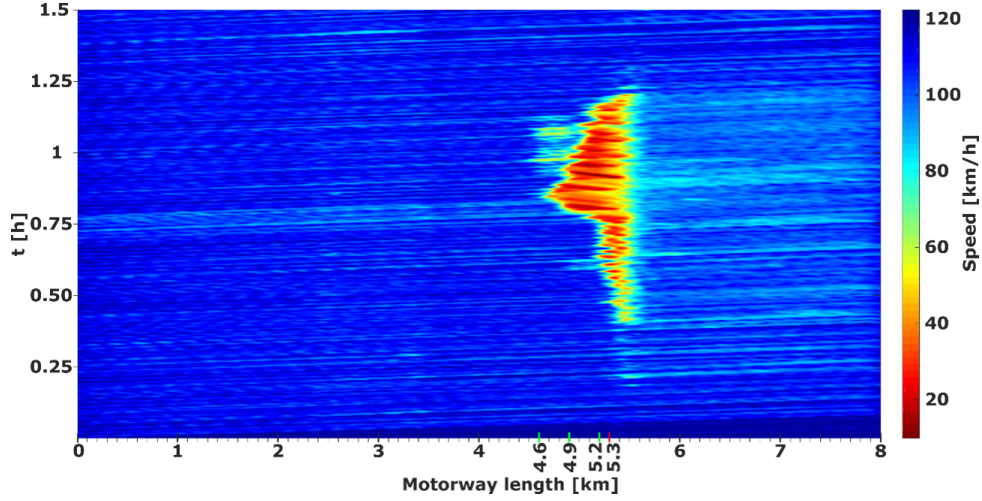


Figure 4.7: Average speed for the best VSL configuration 300 – 300 – 100

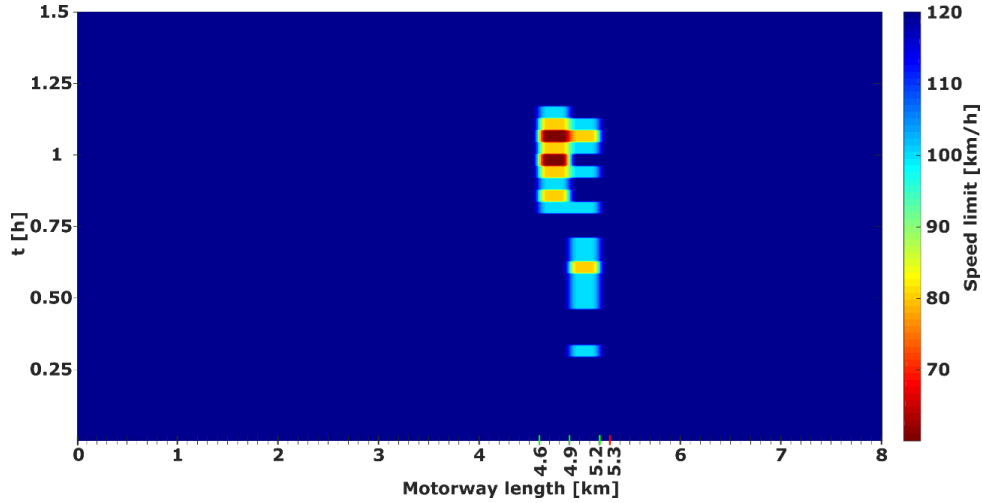


Figure 4.8: Speed limits for the best VSL configuration 300 – 300 – 100

Discussion

As can be seen in Figure 4.6, a traffic jam is inevitable due to the simulated high traffic demand, whereupon the traffic flow enters the unstable congested state. No matter what VSL configuration is chosen, SPSC-VSL cannot prevent the occurrence of congestion. In the case of the best VSL configuration 300 – 300 – 100, the lengths of the VSL zones ($VSL1 = VSL2 = 300$ [m]) and the corresponding speed limits (see Figures 4.7 and 4.8) are sufficient to reduce the traffic outflow from the VSL application zones, which eventually helps to relieve the downstream bottleneck. At the same time, the acceleration area of 100 [m] is sufficient to ensure that vehicles leaving the VSL zone can accelerate and enter the bottleneck at critical speed (where capacity is reached [28]). Consequently, relieving the bottleneck is more efficient than other VSL zone configurations, resulting in the best congestion resolution. It should be noted that this experiment is different from the one described in [28], where the lowest possible speed limit is 10 [km/h], while the nominal speed is 100 [km/h]. Therefore, it is shown that the length

of the acceleration area below 175 [m] is not sufficient to ensure enough distance for the vehicle to reach the V_{cr} at the bottleneck when very low-speed limits are deployed [28]. Since the lowest speed limit in this simulation experiment is 60 [km/h], the length of the acceleration area of 100 [m] might seem sufficient for vehicles to accelerate and reach the bottleneck with V_{cr} . In contrast, with a longer acceleration area, some of the vehicles might accelerate and reach a speed above the V_{cr} and, thus, they need to perform deceleration again to adjust the speed to the lower speed of the bottleneck. Thus, the congestion discharging effect of the bottleneck is suboptimal. As a result, the congestion lasts longer and propagates much farther upstream. Moreover, in the configuration with 500 – 500 VSL zones, the VSL zones are too long and the vehicles that the VSL itself has slowed down travel longer. Consequently, the *TTS* is higher even though the density and measured speeds in the bottleneck area are improved.

This fact implies that there is a VSL zone configuration where both downstream bottleneck travel time and upstream traffic travel time affected by VSL could be efficiently minimized. In other words, this implies that there might exist a trade-off between local optimization of the bottleneck and the entire network as a whole. The recent results in [29] show that the use of longer VSL application areas also implies the usage of longer acceleration areas in the case of higher-speed limits deployment. Therefore, further research on a control structure that adjusts the configurations of the VSL zones (more spatial variation) in the case of SPSC-VSL is desirable. A positive impact of VSL on some aspects of traffic flow (*TTS*, average speed, and density) is correlated with environmental characteristics (vehicle emissions). Based on the results of this experiment, a multi-objective analysis is performed in the next section, which includes two objectives: local bottleneck optimization and the system as a whole with respect to different VSL zone configurations with a finer spatial VSL zones granularity.

4.1.2 Multi-criteria analysis for optimal VSL zones allocation

The first step of the suggested methodology is to perform a simulation analysis of the proposed VSL zone model (see Figure 4.9) to identify the factors that have the strongest correlation with the proposed optimization criteria. Based on these significant factors, a set of solutions is generated for which an attempt is made to determine the correlation between the length and distance of the VSL zone from the bottleneck as well as speed limit values. The analysis is divided into two parts. First, the specific response of the system to different configurations of VSL zones for an active speed limit value of 80 [km/h] during the congestion period is analyzed. The case without VSL was taken as a baseline for the comparison of results (Figure 4.10).

The second simulation experiment examines different configurations of VSL zones along with various speed limits. Both approaches are evaluated using two objectives: (i) the performance of the overall system, measured as total *TTS* aggregated during the simulation, over the entire

motorway segment, and (ii) the bottleneck congestion expressed as the average traffic density over the entire simulation. Both parameters, i.e., their numeric values, are obtained as the average of five simulations run with different simulation seeds. The first criterion has a global character, i.e., it considers the cumulative amount of TTS of the vehicle in the network, whereas the second criterion (simulation-mean traffic density in the bottleneck) has a local character. Finally, for the search of non-dominated solutions in two-dimensional space, the Matlab software tool is used. Particularly, the function *paretosearch* from the program support for multi-criteria optimization to determine the optimal Pareto front (optimal set of solutions) is used. The obtained solutions were further analyzed to obtain an overall picture of the influence of VSL zone configuration on traffic flow performance in the context of simulated traffic congestion.

Simulation experiment setup

In the simulation experiment, different configurations of VSL zones (zone length (L [m]) and distance from a bottleneck (D [m])) were compared to analyze their effects on congestion optimization in the bottleneck (a merge area in the vicinity of on-ramp located at $X3 = 5.3$ [km]). The simulation environment used consists of the microscopic simulator SUMO, in which a synthetic test model of the motorway was created. The VSL zone and speed limit were uniquely defined for each simulation. Each simulation, i.e., each traffic experiment, lasts 1.5 [h]. In this experiment, the motorway model and traffic demand with medium load defined in detail later in chapter 6 is used.

VSL zone and speed limit configurations

The available configurations of VSL zones range from $X1 = 3$ to $X3 = 5$ [km] with a minimum length of 100 [m] and a maximum length of 1,000 [m]. Thus, the main characteristics of the VSL zone configuration (VSL application area and vehicle acceleration area) are clearly determined by L and D (see Figure 4.9). Where s and e stand for the start of the VSL zone and its end, respectively. The VSL activation scheme is based on a simple activation rule: if the lane-weighted average traffic density in the bottleneck is greater than 30 [veh/km/lane], activate VSL, i.e., apply speed limit, otherwise, turn it off, i.e., the nominal speed limit of 120 [km/h] is used.

Analysis of Pareto optimal solution

Figure 4.10 shows a set of solutions obtained by the simulation experiment. On the x-axis is TTS , and on the y-axis is the mean value of traffic density in the bottleneck. Each measured point represents data for a particular configuration of the VSL zone. The selected points a , b and c represent the Pareto front, i.e. a set of optimal non-dominant solutions according to the

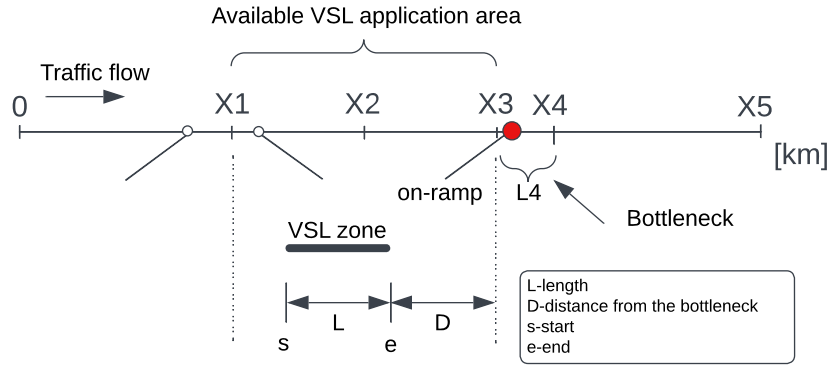


Figure 4.9: Scheme for VSL application zone design

criterion of minimization of both criteria. The configurations of the VSL zones for points *a*, *b*, and *c* are shown in Table 4.4, where the measured parameters *TTS* and density are also shown and compared with the case without VSL. It can be seen that points *a* and *b* have approximately equal *TTS* values, while the average density in the bottleneck is lower compared to the density without VSL. It follows that for the specific experiment, there are two solutions (points *a* and *b*) for establishing VSL zones that have an adequate function to stabilize the traffic in the bottleneck without affecting the total travel time through the network. Although point *c* reduces the traffic density in the bottleneck the most (a decrease of -3.56% compared to the case without VSL), the dependence of the increase in the length of the VSL zone on the increase in *TTS* is evident, which is 0.25% larger than without VSL.

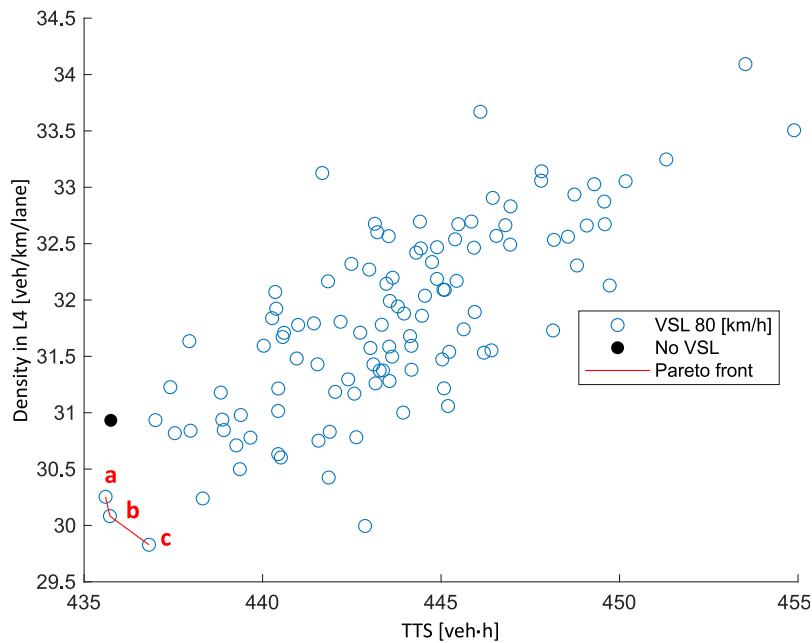


Figure 4.10: Pareto front of optimal non-dominated solutions for different configurations of VSL zones for an active speed limit of 80 [km/h] during the congestion period

Table 4.4: Characteristics of non-dominated optimal solutions (*Pareto front*) for different configurations of VSL zones with a speed limit of 80 [km/h]

	L	D	VSL	TTS	Reduction	Density in L4	Reduction
	[m]	[m]	[km/h]	[veh·h]	[%]	[veh/km/lane]	[%]
No VSL	-	-	-	435.7	-	30.9	-
a	200	800	80	435.6	-0.02	30.3	-1.94
b	300	700	80	435.7	0	30.1	-2.59
c	400	600	80	436.8	0.25	29.8	-3.56

Analysis of Pareto optimal solution with different speed limits

To assess the impact of different speed limit values on the system, an additional analysis that includes variation in speed limit range from 60 to 100 [km/h] has been performed. Figure 4.11 shows a set of solutions obtained by the simulation experiment in which the speed limits were varied in addition to the different configurations of the VSL zones. Different speed limits for individual measurements are marked with different colors. The right side of the graph shows a scale where the blue color represents a speed from 60 [km/h], to the yellow color representing a speed limit of 100 [km/h] (in between there are speed limits with a discrete step size of 10 [km/h]). Thus, as in the previous case, every measured point represents data for each individual unique configuration of the VSL zone with an additional unique color indicating the speed limit. The isolated points a' and b' represent the Pareto front, i.e., a set of optimal, non-dominant solutions. The configurations of the VSL zones for points a' and b' are shown in Table 4.5, where the measured parameters TTS and density are also shown and compared with the case without VSL. It is obvious that points a' and b' with the speed limit of 100 [km/h] have better (lower) values of TTS and density in the bottleneck, compared to the case without VSL. It follows that for the specific experiment, there are two solutions (points a' and b') for establishing VSL zones that have an adequate function to stabilize the traffic in the bottleneck without affecting the total travel time in the network. Moreover, compared to the previous solution, there is an additional reduction in TTS and traffic density when the speed limit is increased from 80 to 100 [km/h]. It can be concluded that the variables for the configuration of the VSL zone and the speed limits are interdependent and their setup strongly depends on the characteristics of the traffic problem to be optimized. Also, the trade-off between local bottleneck optimization and global system performance is evident from the obtained results. For example, one can significantly reduce the speed limit without further improving the performance of the bottleneck, but the traffic on the motorway as a whole will suffer.

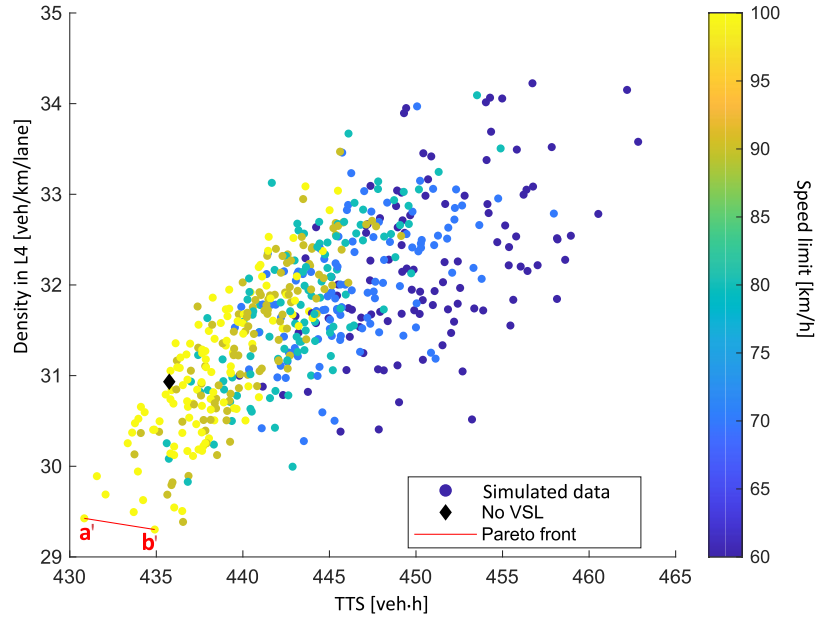


Figure 4.11: Pareto front of optimal non-dominated solutions for different configurations of VSL zones and speed limits

Table 4.5: Characteristics of non-dominated optimal solutions (*Pareto front*) for different configurations of VSL zones and speed limits

	L	D	VSL	TTS	Reduction	Density in L4	Reduction
	[m]	[m]	[km/h]	[veh·h]	[%]	[veh/km/lane]	[%]
No VSL	-	-	-	435.7	-	30.9	-
a'	200	300	100	430.8	-1.12	29.4	-4.76
b'	700	700	100	434.9	-0.18	29.3	-5.18

4.2 Nondeterministic capacity behavior and VSL zones allocation

In this section, the impact of spatiotemporal capacity variations on motorways on optimal VSL zones allocation is addressed. As presented in chapter 2, traffic breakdown and, therefore, the operational capacity, is a random parameter, i.e., even measured at the same location, it can vary from time to time.

Accordingly, for a better understanding of the hypothesis about dependencies in the allocation of VSL zones and nondeterministic capacity, an illustrative example from Figure 4.12 can be used. The phenomena of a bottleneck capacity drop (area A1) caused by on-ramp traffic is depicted.

The function $q_{cap}(x, \bar{t})$ illustrated above the hypothetical motorway segment is, therefore, associated with fluctuating motorway capacity for a given fixed moment \bar{t} . Therefore, to adjust the upstream arriving flow $\bar{q}_{in}(x, \bar{t})$ (this illustrative example assumed constant uninformed mainstream demand indicated by a horizontal dashed line with the arrows in the direction of traffic) to fit into the reduced bottleneck capacity reduced by the amount $\Delta q_{drop}(\bar{t})$, VSL should not be implemented wherever. This is because the upstream capacity itself is limited. Accordingly:

- favorable areas that can be used as storage to accommodate an additional amount of vehicles (equal to $\Delta q_{drop}(\bar{t})$) are areas (motorway regions) A2 and A3 (Figure 4.12);
- in case of temporary flow reduction, the length of VSL zones dictates the duration of flow reduction (Figure 2.5);
- there is no apparent motivation to reduce the flow rate from the VSL area lower than the dropped capacity $\Delta q_{drop}(\bar{t})$.

Therefore, to avoid VSL-induced additional traffic breakdown (and capacity drop) upstream of the bottleneck, VSL should be applied in regions where additional capacity exists with an appropriate VSL zone length proportion and speed limit value. Therefore, regions A2 and A3 (Figure 4.12) should be associated with sections on which VSL zones should be adapted for the particular moment \bar{t} . However, given that a moment later \bar{t}_1 , the curve $q_{cap}(x, \bar{t}_1)$, $\bar{q}_{in}(x, \bar{t}_1)$ and $\Delta q_{drop}(\bar{t}_1)$ might look differently, and, thus, A2 and A3. This implies the need for an additional adjustment of VSL zones in time. Accordingly, one can conclude that the optimal positioning of the VSL zone is not just a question of how far it should be from the bottleneck as presented, for example, in [28, 29], but is also dependent on spatially and temporally constrained upstream motorway capacity by the given prevailing conditions.

Regarding the effect of a temporary reduction of traffic flow (presented in chapter 2) if the higher traffic demand at the on-ramp lasts longer, the temporary effect will disappear after some time. Therefore, the bottleneck will again be exposed to a stronger loading. Thus, the following questions arise. (i) Is it possible to prolong the temporary flow rate reduction from the VSL area by activating different VSL zones? For example, once an active VSL zone has lost the ability to reduce traffic flow can activation of a new capacity-free VSL zone prolong temporary flow reduction? (ii) Is it possible to accommodate additional vehicles in a distributed way across the regions with available capacities (e.g., regions A2 and A3) to achieve the needed lower flow rate to meet the capacity drop $\Delta q_{drop}(\bar{t})$?

One can postulate that this may be achieved by activating VSL zones using appropriate switching procedures, i.e., activating/deactivating, shrinking/extending, and reallocating VSL zones in certain sequences in a timely manner. By doing so, partially temporary flow reduction achieved at different locations (e.g., at A2 and A3 in Figure 4.12) can be summed up as a semi-permanent

flow reduction effect that prolongs the temporary flow reduction effect for a substantial portion of time equal to the duration of for example peak flow rate at the on-ramp. Apparently, such additive property could be useful in VSL traffic control on motorways where the regulation does not allow for implementing speed limits less than e.g., 60 [km/h]. In other words, this can accomplish the low-valued controlled flow (severe flow reduction) achieved by standard MTFC approaches [13, 28, 38] that rely on much lower speed limit values (far below 60 [km/h]).

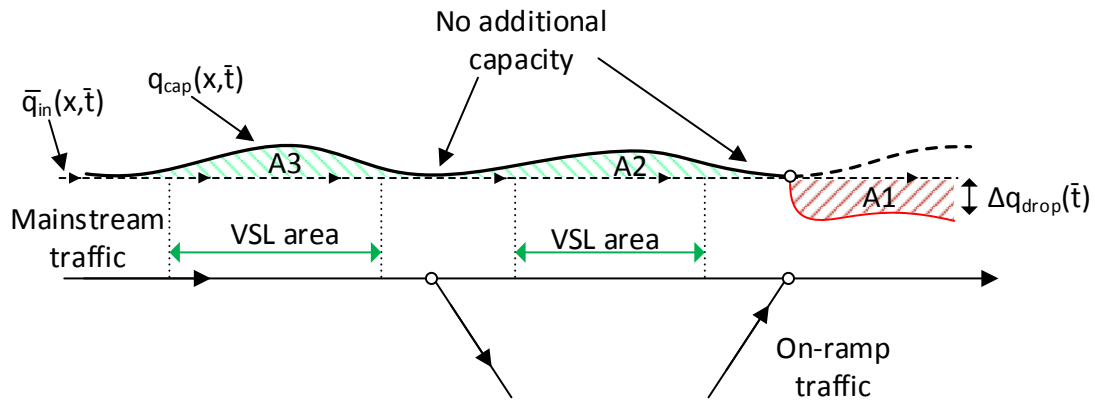


Figure 4.12: Impact of capacity as a function of position and time on VSL zones positioning

The question, then, is: how can one quantify uncertainty in spatiotemporal motorway operational capacity? Given the relationship (2.1) from the fundamental traffic diagram, there is also a relationship between the critical parameter of traffic flow $q_{cap} = \rho_{cr} v_{cr}$. As pointed out in [4], it follows that ρ_{cr} is a more stable parameter, while q_{cap} is not. This means that v_{cr} varies with the change in q_{cap} . Therefore, the sensitivity of v_{cr} to stochastic traffic characteristics or environmental effects may be similar to the sensitivity of q_{cap} . Accordingly, the spatiotemporal capacity can be expressed as a function:

$$q_{cap}(x, t) = \rho_{cr}(x, t)v_{cr}(x, t). \quad (4.3)$$

Although a strong analytical formulation of this conjecture about the correlation between the upstream constrained capacity and the position of VSL zones is not provided, simulations have been used to explain the proposed postulation. Moreover, it can be assumed that important information about the flow capacity and its uncertainty is embedded in $\rho_{cr}(x, t)$ and $v_{cr}(x, t)$ and their spatiotemporal patterns to support the proposed conjecture. Therefore, $\rho(x, t)$ and $v(x, t)$ can be used as state variables for the proposed DWL-ST-VSL since RL agents could benefit from them by extracting knowledge about preconditions (from traffic patterns) that lead to undesirable capacity constraints, i.e., detecting probabilistic traffic breakdown, and proactively adjusting speed limits and VSL zones accordingly.

4.3 Model for dynamic spatiotemporal VSL zones adaptation

By synthesizing all the research results and the simulated results, conclusions can be drawn, based on which the dynamic spatiotemporal VSL zones adaptation model is proposed, which must allow for the timely allocation of sufficient VSL zones combinations. In particular, this is confirmed by the results of the multi-criteria analysis, which shows that multiple acceptable solutions (*pareto front*) to the objectives can be achieved by using completely different VSL zone configurations and speed limits. Thus, the results show that VSL zones should have degrees of freedom in both directions along the motorway. In this way, VSL can adjust the zone configuration closer or farther from the bottleneck (Figure 4.13).

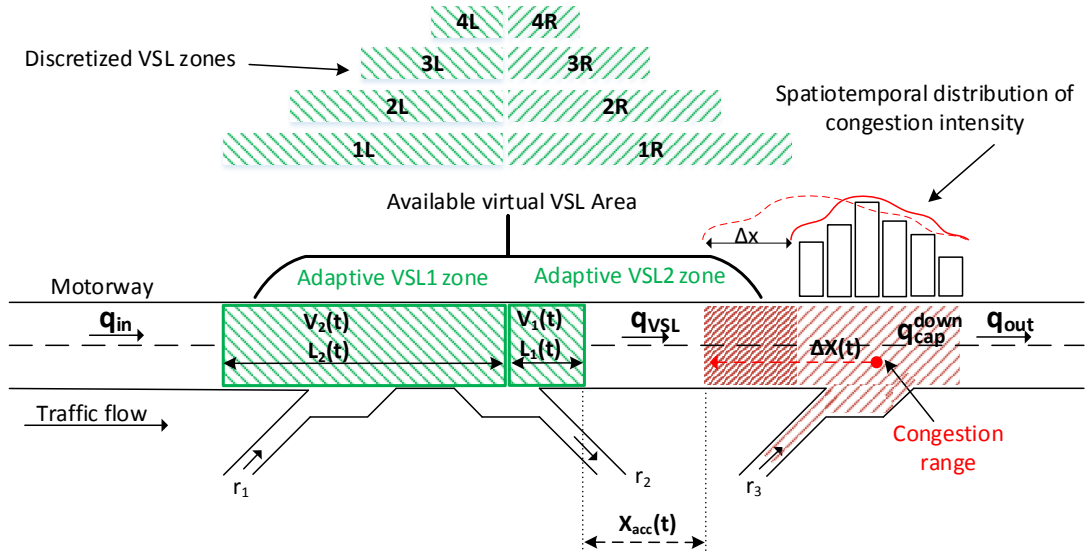


Figure 4.13: Scheme for dynamic VSL zones adaptation

Thus, according to findings in this chapter, the dynamic VSL zones adapting model must be able to: (i) precisely respond to the bottleneck, i.e., adjust the VSL zone to the bottleneck location; (ii) adjust VSL zones to accomplish a congestion tail that moves upstream $\Delta X(t)$ (this can be useful for the mitigation of shockwaves, i.e., moving bottlenecks); and (iii) account for local spatiotemporal disturbances that might occur upstream of the bottleneck.

Accordingly, in Figure 4.13, a virtual VSL area is associated with the spatial domain on which the various VSL zone configurations can be activated. To illustrate this, one can imagine a different combination of active discrete VSL zones within the virtual VSL area. For clarity, in Figure 4.13 the left and right groups of available VSL zones are shown since they will be used by adjacent learning agents to control their own section. For example, an upstream agent may

use the configurations of 1L,..., and 4L, while a downstream agent (closer to the bottleneck) may use the VSL zone of 1R,..., and 4R. More on this will be elaborated in the following chapters.

By doing so, dynamic spatiotemporal adaptive VSL zones do not just account for the lengths of VSL application area ($L_1(t)$, $L_2(t)$) and acceleration area ($X_{acc}(t)$). Moreover, different VSL configurations allow for temporary flow reduction prolongation since different VSL zones can be activated once previously triggered VSL zones get overloaded and if they have lost the flow reduction effect. This goes in line with the proposed conjecture that VSL installation should take into account the uncertain capacity constraints of the motorway upstream sections.

The model scheme of dynamic VSL zones that is presented here represents the underlying logical structure of how VSL application areas should behave. Technically, this is an interface between the VSL controller and the controlled motorway environment that determines how and in what range VSL zones can be adjusted. As such, it will be an integral part of the main program logic of the proposed learning-based controllers. The task now is to develop a strategic control logic capable of simultaneously optimizing the speed limits $V1$ and $V2$ as well as the VSL zone lengths $L1$ and $L2$, and, thus, the acceleration area X_{acc} in the given example. As such the presented adaptive model satisfies all the basic criteria for setting VSL zones defined in the research literature so far.

4.4 Concluding remarks

For optimal traffic flow control on the motorway using VSL, the initial conditions of traffic dynamics in the area of interest must be taken into account when calculating the new speed limits and the position of the VSL zone. However, the spatial *boundary conditions* of constrained capacity must also be taken into account. Thus, the facility where the speed limit must be implemented is split into two segments: (i) the point to be optimized, e.g., a downstream bottleneck, and (ii) the upstream motorway segment where VSL is to be deployed. Thus, the constraints must include both a feasible region of additional upstream capacity where VSL can be applied and the point of interest for optimization.

Accordingly, the need for a spatiotemporal dynamic adaptation of VSL zones that accounts for the nondeterministic nature of motorway capacity using the fact that capacity is not constant but changes with position and time can be postulated. To accomplish this, a model for dynamic spatiotemporal adjustment of VSL zones is proposed. It is a rather simple model, but with it, one can achieve and justify quite a number of things about dynamic VSL zones adaptation, which will be presented in the following chapters.

Multi-criteria analysis for the simulated bottleneck provides evidence that exists the trade-off

between local bottleneck optimization and the system as a whole. Accordingly, the main hypothesis regarding the operation of dynamic spatiotemporal VSL zones is as follows. To achieve a reduction in traffic flow equal to or less than the capacity of the activated bottleneck, it is not always necessary to reduce the speed of upstream traffic to values equal to or less than the speed of the bottleneck. It is supposed that an adequate reduction in traffic flow can be achieved by maintaining the speed of upstream traffic flow at a higher level by dynamically and sequentially adjusting the VSL zones along with the speed limit values. Of course, the adjustment should be done taking into account the spatiotemporal traffic characteristics, which will be discussed further on. Thus, one can try to simultaneously take advantage of both effects - speed homogenization and MTFC - by dynamically adjusting the VSL zones to improve the operation of the bottleneck without affecting the overall system performance.

Chapter 5

Reinforcement learning based VSL

This chapter presents the essential elements for understanding RL techniques and solving VSL control problems with their help. Accordingly, a single-agent RL controller for VSL is introduced. The twin curses of modeling VSL using continuous traffic variables are discussed, for which an application of function approximation techniques in RL to deal with such problems is presented.

5.1 Reinforcement learning

RL can be seen as a form of simulation-based dynamic programming technique for making decisions in situations that have some dynamics - that evolve over time [18]. In an AI community, RL is viewed as a *machine learning* technique. Hence the word learning, while reinforcement can be associated with an agent that learns through trials and errors as it shall be seen later. It is useful in solving large-scale nondeterministic processes [111]. Decisions are made sequentially. The concept itself is called stochastic control, and the main method to deal with such problems is dynamic programming. In that context, RL can be seen as an approximation method for solving classical dynamic programming problems. It combines the principle of the Monte Carlo method with the principle of dynamic programming, which in RL is called the temporal difference method. In RL, a simulation can be used to generate samples of the value function of a complex system, rather than finding an explicit model (state transition probability). Samples are then discounted or averaged to obtain the expected value of the value function. Therefore, transition probabilities are not required in RL.

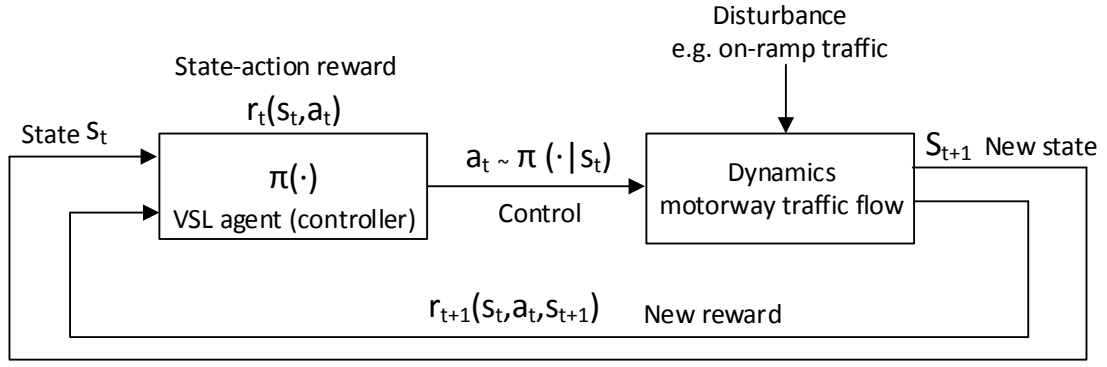


Figure 5.1: Concept of applying RL in VSL control

This avoids the curse of dimensionality, i.e., a potentially large number of states which leads to the well-known curses of dynamic programming: the curse of modeling and the curse of dimensionality [18]. Although it is known as a model-free technique, the RL optimization process heavily relies on simulations. Therefore, an implicit model of the controlled process is persisted in a simulator classifying the RL as a simulation-based optimization technique.

RL agent in general learns its optimal behavior (e.g., VSL control optimization) by trial-and-error interactions within its environment (see Figure 5.1). Thus, in the standard RL paradigm, the agent is connected to its environment via the perception and action framework. At each step of interaction, the agent senses the environment and then selects an action to change the state of the environment. This state transition generates a reinforcement signal (reward or penalty) received by the agent [112]. Therefore, the agent's goal is to learn how to move the system along a desirable trajectory of states to maximize the return. The agent's environment is assumed to be modeled as a Markov Decision Problem (MDP). An MDP is a tuple $\langle S, A, P, R \rangle$ where S is a finite set of n discrete states and A is a finite set of actions available to the agent. The actions are stochastic and Markovian in the sense that an action a_t in a given state $s_t \in S$ results in a state s_{t+1} with fixed probability $P(s_{t+1}|s_t, a_t)$, while $R(s_{t+1}|s_t, a_t)$ represents the reward received after state transition. Policy π is a function representing a mapping from states to actions $\pi : S \mapsto A$, which optimizes performance, for example, expected accumulated reward received. RL methods differ according to the exact measure and optimization criteria (total reward optimization, discounted reward optimization, or average reward optimization) used for select actions.

5.1.1 Q-learning algorithm

In this thesis, the focus is on using the Q-learning algorithm to optimize the VSL control process. Q-learning is an off-policy model-free RL algorithm that learns to associate an action a_t with the expected long-term payoff (reward) for performing that action in a given state s_t [113]. How good an action is in a given state is expressed as a Q-value stored in the Q-learning function $Q(s_t, a_t)$. Q-function is learned using the following iterative update rule:

$$Q_i(s_t, a_t) := Q_i(s_t, a_t) + \alpha_n(r_{t+1} + \gamma \max_{a' \in A} Q_i(s_{t+1}, a') - Q_i(s_t, a_t)). \quad (5.1)$$

The performed action a_t in state s_t stimulates a state transition to the new state s_{t+1} , from which an optimal action is a' . Depending on this transition, the agent receives a reward r_{t+1} . The parameter α_n is the learning rate that controls how fast the Q-values are adjusted. The discount factor γ controls the importance of future rewards. Various exploration/exploitation strategies (e.g., ϵ -greedy) are used to search the solution space, i.e., to ensure that the agent sufficiently explores its environment and learns the appropriate action in a given state.

5.1.2 Function approximation for RL

The update rule for Q-learning implies that equation (5.1) requires a lookup table to store the learned Q-values for all possible (s_t, a_t) pairs. When RL is applied, for example, to traffic control on motorways, states can be defined using traffic parameters such as normalized traffic density and average traffic speed suitable to describe traffic conditions. As a result, the sets S , and A become vast or infinite, and the stochastic algorithm (5.1) loses its efficiency in its basic lookup formulation. The dimensions of the state representation grow as a function of the discretization of the traffic flow parameters. As the number of states or discretization granularity rises, the size of the lookup table increases exponentially, revealing the curse of dimensionality. It is worth noting that the state space becomes infinite by the inclusion of continuous variables. Actions have a similar problem. The number of actions increases the size of the table. And if the actions are continuous, the table becomes infinitely large. But, continuous actions along with high-dimensional state modalities are desirable in many real-world control applications. Accordingly, to visit every state-action pair sufficiently for good Q-value function estimation can take a considerable amount of computer time or could be impossible [113]. The most common solution to these problems is to replace the table with an approximation function. Thus, learning the optimal Q-values requires some form of function approximation (linear or nonlinear). Approximations address the large dimensionality problem of storing state-action values in the computer's memory and provide a way to obtain solutions in a reasonable computational time [18]. Moreover, the approximation function of Q-values allows the computer to

work with continuous state/action variables, which, in the end, plagues many real systems with possible large solution space, e.g., modeling VSL control as RL problem [24, 53]. Moreover, nonlinear function approximation techniques, like neural network approximator, may improve control if an underlying controlled process is nonlinear and nonstationary, as can be the case with motorway traffic control [21, 54, 114].

5.2 Modeling VSL as RL problem

In this section, the single-agent RL-VSL controller is introduced. Particularly, Q-learning to solve the VSL control (QVSL) problem is used. To assess the benefit of using function approximation in RL both are presented, QVSL with full state representation (QVSL-FS) using tabular formulation (Q-matrix) and the QVSL with the function approximation (QVSL-FA) techniques.

5.2.1 Variable speed limit control as an MDP

The VSL problem has to be defined as an MDP with the assumption that an agent makes the control decisions. The agent has to activate different speed limits at the end of every decision interval. For every possible state of the environment, the agent can select a particular action and obtain a reward depending on how well the selected action (speed limit) has performed. The agent cares for accumulated discounted rewards gained from a sequence of executed actions. The transition time between states after activating VSL control equals the control time step. The state may change every time when the agent takes an action that affects the current traffic state. Thus, the VSL decision process can be formulated as an MDP problem and solved by applying RL [115].

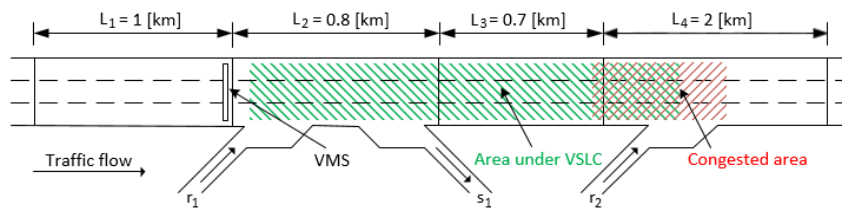


Figure 5.2: Controlled motorway stretch divided into four sections

For QVSL, the controlled stretch of the urban motorway divided into smaller sections (see Figure 5.2) represents the model of the agent's environment. Actions are speed limits that the agent can send on VMS. For each state $s_t \in S$ of the environment, the agent can select an action a_t from a finite set of actions A . For simplicity, the executed action a_t at time t is a single speed limit posted on VMS at the beginning of the controlled sections (L_2 and L_3 in Figure 5.2) simultaneously for all analyzed approaches. In this experiment, a VSL setup with fixed zones

is examined. Traffic density and speed measured in several consecutive sections (L_2 , L_3 , and L_4 in Figure 5.2) are used for state representation.

The reward r_t received by the agent at time step t after state transition to the new state s_{t+1} has the following form for all implemented QVSL approaches:

$$r_t = \begin{cases} -\delta, & \text{if } (a(t-2) = a(t-1)) \wedge (a(t-1) \neq a(t)) \\ -\delta, & \text{if } |a(t-1) - a(t)| > 20 \\ 0, & \text{if } 105 < \min\{v_i(t+1) \mid i = 2, 3, 4\} < 110 \\ \delta, & \text{if } \min\{v_i(t+1) \mid i = 2, 3, 4\} \geq 110 \\ -TTS(t), & \text{otherwise} \end{cases}, \quad (5.2)$$

with slight modification compared to the reward function used in [116] by including the positive reward δ in the case when the agent recognized free-flow conditions. The constant δ should be at least equal to the maximal expected TTS value, measured between the previous and the current control time step. The first condition of r_t prevents oscillations of the speed limit, and the second presents punishment if the difference between two consecutive speed limits is too large. If the average speed v_i of the observed motorway sections i is between $105 \text{ [km/h]} < v_i < 110 \text{ [km/h]}$, it can be assumed that there is currently no congestion. Therefore, the agent does not receive punishment in such states. In the case where measured average speeds are above 110 [km/h] , free-flow traffic is assumed, and the agent receives a reward. These two conditions can be true if the agent executes actions that allow for free flow traveling speed of vehicles, and there is a free flow condition across affected sections. In all other cases, the agent receives a punishment that is proportional to TTS spent between the previous and current control time step measured across the affected area (L_2 , L_3 and beginning of L_4 in Figure 5.2) during every control time interval. As such TTS can be used to define the objective function of the traffic flow optimization [116], [117].

In the beginning, the agent does not know anything about the environment and which action to take. An exploration-exploitation strategy has to be applied to test all possible actions. In this thesis, the ϵ -greedy mechanism is applied in all compared approaches. First, the agent tries random actions, and, with time, starts to use what it has learned. The exploration parameter ϵ bounded with $0 < \epsilon < 1$ is used to manage the share of exploration and exploitation during learning. When $\epsilon \rightarrow 1$ the agent acts only randomly and can even worsen the current LoS on the urban motorway. When $\epsilon \rightarrow 0$ the agent uses the knowledge stored in a Q-matrix (the memory of what the agent has learned through experience). During multiple simulations, ϵ is gradually decreased according to the function $\epsilon = \exp \frac{(1-n)}{600}$ (n denotes the number of current simulation). When ϵ reaches 0.03, it remains constant.

Q-Learning with full state representation

For the implementation of QVSL-FS, discrete states and actions need to be defined. Actions belong to one of the following two sets:

$$A_1 = \{130, 110, 100, 90, 80, 70, 60\}, \quad (5.3)$$

$$A_2 = \{130, 110, 100, 80, 60\}, \quad (5.4)$$

where the set A_1 has been used for the analysis of QVSL-FS and QVSL-FA, and the smaller set A_2 only in QVSL-FS with constant learning rate $\alpha = 0.5$. This has been done to make a Q-matrix with fewer number of elements (only 3,125).

A Q-value is assigned to each state-action pair as a measure of the quality of each combination. The learned Q-value function $Q: S \times A \mapsto \mathbb{R}$ represents a mapping from state-action pairs to expected long-term return obtained by executing a specific action in a given state [113, 118]. For a non-deterministic environment, the basic idea of Q-learning is to update the Q-value iteratively by using newly received training sample $(s_t, a_t, s_{t+1}, r_{t+1})$ according to equation (5.1).

In the case of QVSL-FS, the learning rate $\alpha_n = \frac{1}{n}$. The rate of α_n should be decreased over time to ensure convergence to an optimal policy as explained in [113]. In this analysis, γ is set to 0.8, and remains the same for all tested approaches. Q-values for implemented QVSL-FS are stored in a five-dimensional Q-matrix with 6,125 elements according to $|S \times A_1|$, where S is a finite set of states defined as:

$$S = \{\rho_2(t), \rho_3(t), \rho_4(t), a(t-1)\}, \quad (5.5)$$

where $\rho_i(t)$ represents the density of the traffic flow at the i th section for time step t . Different density values are coded in five grades based on the values (10, 15, 22, 30). The critical density for the applied model is $\rho_c = 29$ [veh/km/lane]. The last term in (5.5) $a(t-1)$ is the speed limit from the previous control time interval.

The Q-learning algorithm (5.1) converges to the optimal Q-values if every state-action pair is visited plenty of times and the learning rate is decreased appropriately over time. After the Q-values for sufficiently many state-action pairs have been estimated during the learning process, the optimal action for a particular state is determined as the one with the largest Q-value. Then, the Q-learning agent can be applied for optimal control using its knowledge.

Q-Learning with function approximation

When RL is applied for traffic control on motorways, states of the system can be defined using traffic parameters such as normalized density and average speed, which are suitable to define precisely current traffic states. Once dealing with continuous variables, the computation of Q-values requires some form of function approximation (Q-function parametrization).

To model the Q-learning with function approximation, feature-based state representation is used as a method for constructing the basis for the function approximation. Features capture important properties of continuous space of the agent's environment. In the case of linear function approximation methods, using the methods coarse and tile coding, the state space can be mapped into a vector of binary features. For example, if the state s is inside a circle (see Figure 5.3), then the corresponding feature has the value 1 and is said to be present, otherwise, the feature is 0 and is said to be absent. This kind of 1-0 valued feature is called a binary feature.

On the other hand, for example, the radial basis function (RBF) uses the Gaussian function to map the state vector \vec{s}_t into features represented with real numbers within the interval $[0, 1]$. Mentioned approaches are used to create basis functions whose linear combinations approximate the Q-values. Thus, in the setting of Q-learning with function approximation, the goal is to approximate the Q-value by learning the parameter vector $\vec{\theta}$ of an approximate value function Q_θ as:

$$Q_\theta(s, a) = \vec{\theta}^T \vec{\phi}_{s,a} \approx Q(s, a), \quad (5.6)$$

where $\vec{\phi}_{s,a}$ is an m -dimensional column vector that captures important properties from the state-action pair (s_t, a_t) , and $\vec{\theta}$ is a parameter whose dimension is identical to $\vec{\phi}_{s,a}$ [119]. At the beginning all components of $\vec{\theta}$ are set to zero. Now, the task is to learn $\vec{\theta}$ by applying the incremental stochastic gradient descent update rule (5.7) allowing for approximation of the Q-value function (5.6):

$$\vec{\theta}_{t+1} := \vec{\theta}_t + \alpha_n (r_{t+1} + \gamma \max_{a' \in A} (\vec{\theta}_t^T \vec{\phi}_{s_{t+1}, a'}) - \vec{\theta}_t^T \vec{\phi}_{s_t, a_t}) \vec{\phi}_{s_t, a_t}, \quad (5.7)$$

where α_n is the learning rate that decreases with the number of learning episodes. In the case of the presented simulation experiment, the episode refers to one simulation that comprises the uncongested traffic condition, breakdown, congested condition, and traffic recovery.

Feature-based state representation

Generalization of states is essential when Q-learning is applied to continuous state space. Feature-based representation captures important state properties. Often to binary numbers (0/1) (coarse and tile coding) or to the interval $[0, 1]$ (RBF). The used state vector \vec{s}_t was slightly modified compared to the one used in [116]. Three additional density states were added for improving the representation of the current traffic states. In all three methods of feature constructions (coarse coding, tile coding, and RBF), the state set now becomes a state vector $\vec{s}_t \in \mathbb{R}^{m_0}$, where $m_0 = 8$ is the number of state components. Also, the chosen state vector comprises average speeds and past two actions as components of the current state:

$$\vec{s}_t = \left(\frac{a(t-1)}{v_f}, \frac{a(t-2)}{v_f}, \frac{\rho_2(t)}{\rho_j}, \frac{\rho_3(t)}{\rho_j}, \frac{\rho_4(t)}{\rho_j}, \frac{v_2(t)}{v_f}, \frac{v_3(t)}{v_f}, \frac{v_4(t)}{v_f} \right). \quad (5.8)$$

The numerators of the first two components of the state vector (5.8) are the previously executed speed limit values. Variable $\rho_i(t)$ is the current density, and $v_i(t)$ is the current average speed in motorway section i . Every element in the state vector (5.8) is normalized into the interval $[0, 1]$ by appropriate denominators, where $v_f = 130$ [km/h] is the free flow speed, and $\rho_j = 80$ [veh/km/lane] is the jam density. Using the earlier mentioned feature-based methods, the important properties from (5.8) are captured into the feature vector $\vec{\phi}_{s,a}$.

This feature vector is defined differently for the three chosen state representations as follows [53].

Coarse coding

The receptive field corresponds to circles in the state space. For example, according to two-dimensional state space, if the state coordinate is inside the circle, then the corresponding feature has the value 1 and is said to be present. Otherwise, the feature is equal to 0. The receptive field can overlap, enabling generalization between different states as can be seen in Figure 5.3. The binary feature vector corresponding to 2D state space in the example below becomes $\vec{\phi}_s = (0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0)$.

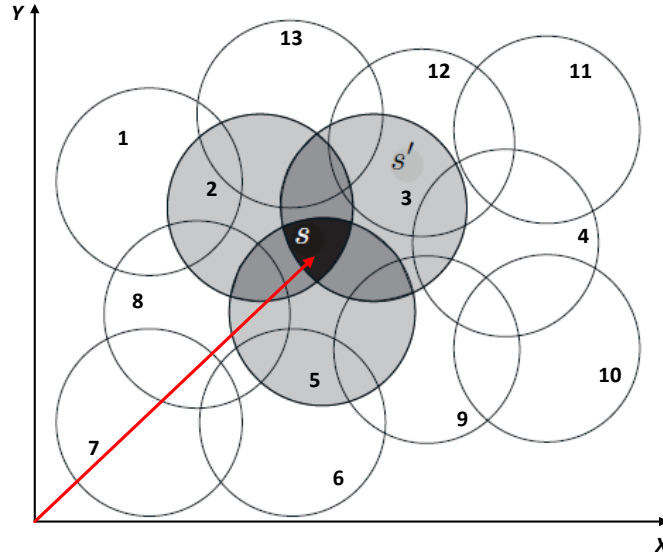


Figure 5.3: An example of a binary feature vector corresponding to 2D state space

Extending of coarse coding in an eight-dimensional state space corresponds to \vec{s}_t (5.8), and the receptive fields are now hypersphere tiles with radius $r = 0.7$. The state space was filled with $k = 900$ points. Each of those represents the center of a hypersphere. The points are placed randomly within the state space, thus, achieving an overlap effect. The dimension of the feature vector (5.9) is $m = 1 + |A_1|l$ with $|A_1|$ representing the cardinality of the action set, while $l = k$, plus one extra element for bias. The same state vector is used in all three approaches with function approximation. Each possible combination of actions and hypersphere tiles has a unique component in $\vec{\theta}$.

Tile coding

The receptive field of the features is grouped into partitions of the input space. Each partition is called tiling, and each element of the partition is called a tile. According to a two-dimensional state space example, the simplest tiling is a uniform grid, where receptive field (tiles) are squares [111]. One tiling in eight-dimension fills the state space with p^{m_0} equally spaced hypercube tiles. The binary feature vector in this case is \mathbb{R}^m , where $m = 1 + |A_1|l$. Here $l = kp^{m_0}$, parameter $p = 3$ is the number of tiles (hypercubes) along one dimension, parameter $k = 64$ is the number of tilings, and m_0 is the dimension of the state vector (5.8). To ensure a higher resolution of the state space partitioning, k tilings are created. Each tiling is shifted by the displacement vector $\vec{d} = (\frac{1}{10}, \frac{3}{10}, \frac{5}{10}, \frac{7}{10}, \frac{9}{10}, \frac{11}{10}, \frac{13}{10}, \frac{15}{10})$, meaning that it is shifted from the previous tiling by $\frac{\omega}{k}$ times \vec{d} [111]. Tile width ω was defined as $\omega = \frac{1.4}{(p-1+\frac{1}{k})}$. With those small shifts, the state space is filled with k overlapping tiles. A single point within \mathbb{R}^{m_0} , corresponds to the coordinates of the state vector (5.8), will fall in precisely one tile in every of the k tilings.

These k tiles correspond to k features in (5.9) that become active when a particular state occurs, and particular action has been executed. As mentioned, the feature vector is in \mathbb{R}^m and has the following form:

$$\vec{\phi}_{s,a} = (1, \phi_1(s, a_1), \dots, \phi_l(s, a_1), \dots, \phi_1(s, a_7), \dots, \phi_l(s, a_7)), \quad (5.9)$$

where an index of action a indicates the speed limit which has been executed in respective state s . The index l stands for possible active tiles at time t . The first component stands for the bias term to properly scale the function values. The number of total elements in (5.9) seems a bit bigger, but computing the dot product of $\vec{\theta}^T \vec{\phi}$ in (5.7) gains computational advantages because most binary features in (5.9) are always zero.

RBF

This approach is a natural generalization of coarse coding to continuous-valued features. The feature can gain value from the interval $[0, 1]$. Typical RBF uses the following Gaussian response:

$$\phi_{i,s,a} = \exp\left(-\frac{\|\vec{s}_t - \vec{c}_i\|^2}{2\sigma_i^2}\right), \quad (5.10)$$

where $\{\phi_i(\|\vec{s}_t - \vec{c}_i\|) \mid i = 1, 2, \dots, k\}$ is a set of k arbitrary functions, known as radial basis functions. The i th function is a component of the feature vector $\vec{\phi}_{s,a}$. The response of (5.10) depends on the distance between the state vector \vec{s}_t and the center of the basis function \vec{c}_i , and relative to the width σ_i of the radial basis function with respect to the center \vec{c}_i [111] (a one-dimensional example of RBF is shown in Figure 5.4).

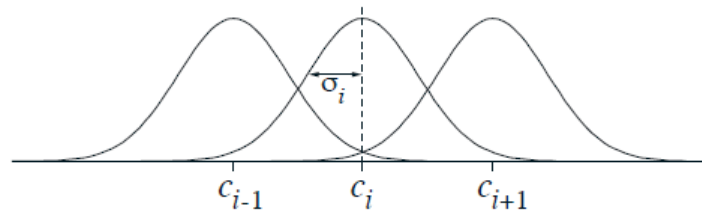


Figure 5.4: An example of one-dimensional radial basis functions [111]

The feature vector is in \mathbb{R}^m , where $m = 1 + |A_1|l$. In the implementation of QVSL using feature construction with RBF, $l = k = 64$, and σ_i have been set to 0.5 according to [120]. This method reduces performance when there are more than two state dimensions like in QVSL because all components (associated with the currently executed action) in (5.9) are active at time t . But so far, the number of RBF functions (5.10) within the feature vector (5.9) has been reduced compared to the previous two cases.

In all mentioned QVSL-FA approaches (coarse and tile coding, and RBF), the components in (5.9) associated with actions that are not executed at time t will be zero, except the bias term.

Simulation setup

To simulate the described QVSL implementations, in this experiment, a simulation framework consisting of the microscopic simulator VISSIM and MATLAB [117, 121] was used. Every simulation lasted 2.5 [h]. All QVSL approaches are learned during 5,000 simulations (episodes) with different seeds to generate a wider spectrum of traffic events, i.e., a stochastic environment for the agent.

Model of the urban motorway

The characteristics of the used motorway model were taken from [122] and modified in order to be suitable for the applied simulation framework [117]. It is a three-lane motorway section as shown in Figure 5.2, divided into four sections. VSL is active within the green area in sections L_2 and L_3 , and at the beginning of L_4 . The first section L_1 is without VSL. A new speed limit value is sent to the VMS during every control time step, T_c (300 [s] in this experiment). All traffic data are collected with the frequency of $T = 30$ [s]. Section L_2 contains one on- and off-ramp (r_1 and s_1). The second on-ramp r_2 is placed in section L_4 , and here congestion is created by changing the input flow at this on-ramp. Congestion gradually propagates upstream into section L_3 and creates a disturbance in it.

The simulated traffic flow consists of 96% cars, 2% trucks, and 2% of buses. The mainstream flow has a constant demand of 4,500 [veh/h] during the whole simulation, of which 95% of traffic remains on the mainstream while the other 5% exit the motorway through the off-ramp s_1 . The on-ramp r_1 has a constant traffic demand of 1,350 [veh/h] while traffic demand at the on-ramp r_2 changes during simulation. It starts with a constant value of 300 [veh/h], whereupon it increases linearly reaching the maximum of 1,250 [veh/h]. This value remains for half an hour after which the traffic demand linearly decreases to its starting value and stays constant to the end of the simulation. One has to notice that the applied traffic simulator generates vehicles for the envisaged traffic flows stochastically using a particular seed for each repeatable simulation.

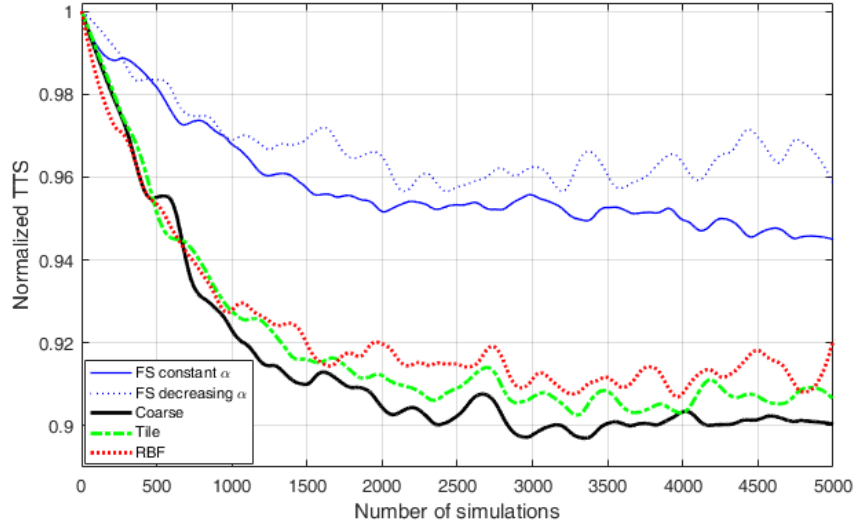
5.2.2 Simulation analysis

Normalized TTS values obtained during the learning process are given in Figure 5.5. A polynomial interpolation was used for a comprehensive illustration. All approaches reduce TTS , but linear function approximation approaches have a steeper decrease rate. The two blue curves represent the results of the QVSL-FS algorithms. The first has a constant learning rate α , and the VSL controller learns faster due to the lower number of elements in the Q-matrix as a con-

Table 5.1: VSLs PERFORMANCE RESULTS

	No VSL	QVSL-FS $\alpha = 0.5$		QVSL-FS dec. α		QVSL-FA Coarse		QVSL-FA Tile		QVSL-FA RBF	
		Obtained	Red. [%]	Obtained	Red. [%]	Obtained	Red. [%]	Obtained	Red. [%]	Obtained	Red. [%]
Max. TT [s]	387	322	16.7	285	26.2	304	21.2	319	17.5	313	19.1
Avg. TT [s]	184	184	0	185	-0.6	177	3.6	179	2.3	180	1.9
TTS [veh · h]	749	729	2.6	736	1.7	708	5.5	714	4.6	723	3.5
Max. Queue [veh]	36	31	13.9	11	69.4	16	55.6	21	41.7	36	0
Avg. Queue [veh]	4.2	4.0	4.1	0.9	77.0	1.2	71.3	1.2	70.7	2.5	39.2

sequence of the smaller action set A_2 . The second also has a decreasing learning rate α_n . A slighter decrease in TTS could be related to an inappropriate (too fast) decrease in the parameter α_n , where the agent cannot correct the Q-values associated with random (bad) actions at the beginning of the learning process. In general, QVSL-FA methods with state generalization outperform QVSL-FS. All QVSL-FA methods exhibit a similar rate of TTS decrease, though the rate is somewhat steeper with coarse coding.

Figure 5.5: Convergence of the normalized TTS during the training process

The impact of VSL on the mainstream traffic parameters is most evident in sections L_3 and L_4 as shown in Figure 5.6. Two cases are shown. The first case is that of no-control (denoted black by a dashed line) and the second is for QVSL with tile coding (indicated in green). Tile coding was selected for presentation since the typical gradual decrease of the speed limit, and its increase without large unallowed changes is most apparent. This is a desirable behavior of VSL that has to be ensured. In the density graph in Figure 5.6 for section L_4 , the positive effect of VSL is evident with actions that actively reduce the density compared with the case of no-control. The timely applied sequence of speed limits upstream of the bottleneck keeps the traffic flow speed in the bottleneck L_4 at a higher value compared to the case of no-control. The gradual reduction of vehicle speeds coming into section L_4 allows the congestion to dissolve more quickly than in the case of no-control. After the congestion has dissolved, a steep increase in speed in sections L_3 and L_4 can be observed in comparison to the case of no-control.

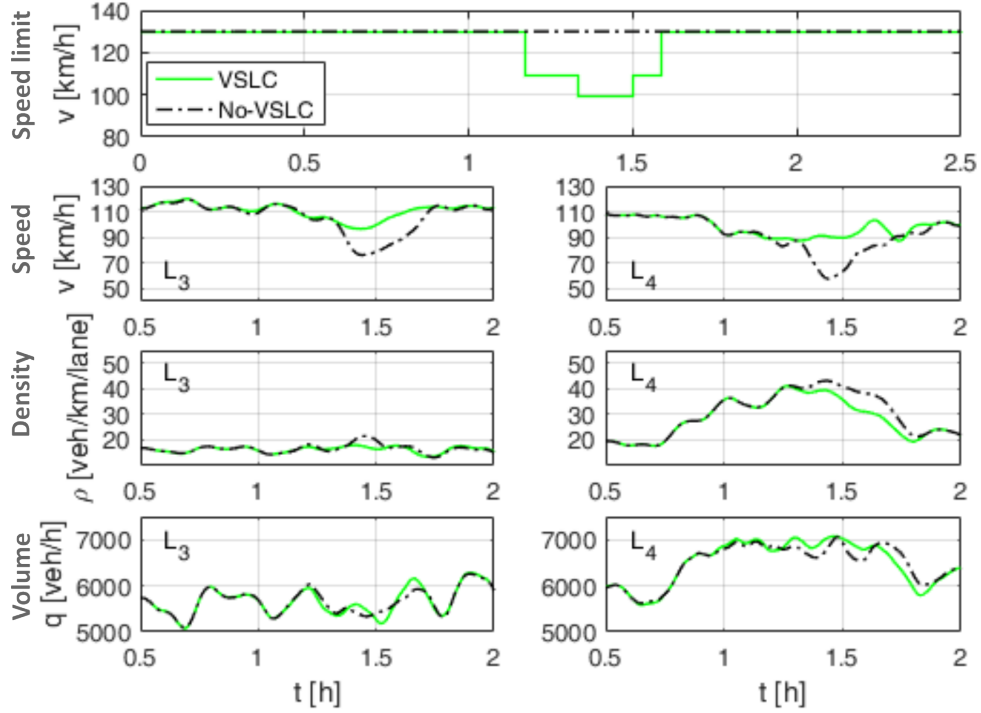


Figure 5.6: Traffic parameters for tile coding in sections L_3 and L_4

Additional MoEs Travel Time (TT) on mainstream, TTS , and queue length at the on-ramp r_2 were measured during simulations. The performance has improved regarding these additional MoEs as shown in Table 5.1. QVSL improves the LoS on the mainstream (lower TT) and at the congested on-ramp (shorter queues). Only QVSL-FS with decreasing α_n has a minor deterioration of TT . Coarse coding shows the best performance regarding obtained convergence rate and value of TTS .

5.3 Concluding remarks

In this chapter, four different approaches of modeling QVSL to learn the optimal VSL policy with the goal of minimizing TTS on the simulated motorway segment are applied. Mainly the standard lookup formulation of modeling QVSL as well as the function approximation methods that allow working with the continuous state space domain are presented. Particularly linear function approximation with different state representation methods based on feature construction techniques, Coarse and Tile coding, and RBF, are demonstrated. The simulation results show that function approximation can learn the needed control policy faster compared to standard full-state representation. However, QVSL-FS, in theory, will achieve equal or even better results but it requires much more computations (training episodes as evident from results in Figure 5.5).

Feature construction for linear methods is a practically applicable method to generalize a con-

tinuous higher dimensional state space. In this case, generalization is required to capture essential properties of multiple traffic state variables. However, as the dimensions of the state space increase, the feature vector may grow exponentially due to finer partitioning of the state space resolution. Accordingly, the designer of such a controller must identify representative traffic parameters that provide comprehensive information about traffic state conditions in order to reduce the dimensions of the state vector or the size of the lookup table (Q-matrix). This would enable state space dimension reduction and, consequently, allow for finer partitioning (higher resolution) of the state space, thus, improving the generalization of unvisited state space regions.

Chapter 6

Cooperative multi-agent learning VSL control system

This chapter extends beyond the notion of a single-agent RL introduced in the previous chapter. Therefore, this chapter is entirely about MARL techniques and their application in VSL control. First, the W-learning algorithm that is primarily used for solving control tasks characterized by competitive control policies is introduced. This is an essential foundation to introduce the DWL algorithm that is used to solve VSL control with dynamic VSL zones. In general, several experiments will be presented: (i) whether MARL has some benefit and to which extent compared to single agent RL, (ii) whether the agents' cooperation in DWL can improve the VSL control, and most importantly (iii) whether the DWL-ST-VSL with a dynamic spatiotemporal adaptation of VSL zones can improve VSL control.

6.1 Multi-agent based RL

If multiple intelligent RL agents form one system, it can be called a MARL system, in which agents can interact and collaborate using direct or indirect communication [123]. As shown in the previous chapter, RL is a technique that can automatically detect patterns in the data received from the system (process), and then use the discovered patterns to predict the future of the data or to perform other types of decision-making under uncertainty. Thus, its extended version *MARL*, to address the uncertainty in the motorway flow capacity and to optimize traffic via VSL, is elaborated in this chapter.

6.1.1 W-learning

WL algorithm proposed in [55] was designed to manage competition between multiple tasks. In particular, an individual policy is implemented as a separate Q-learning process designed by its own state space. As mentioned, the goal is to learn Q-values for state-action pairs for each policy, where a single policy can be viewed as an agent. At every control time step, each policy nominates an action based on Q-values. Applying WL for every state s of each of their policies, the agent learns what happens concerning the reward received if the nominated action is not performed (rated using a W-value for a given state $W(s)$). Thus, an agent only needs local knowledge—what state s_t it was in, whether the nominated action was obeyed or not, the state transition s_{t+1} , and the received reward r_{t+1} .

Hence, all policies recommend new actions. Nevertheless, only one action is executed (suggested by the *winner policy*) based on the highest W-value (if not, this policy will suffer the highest deviation). Each policy updates its own Q_i function using the winning action a_k and its own received reward r_i . W_i values are updated only for policies that were not obeyed ($i \neq k$) using the following update rule:

$$W_i(s_t) := (1 - \alpha_W)W_i(s_t) + \alpha_W(1 - \alpha_Q)^\omega(Q_i(s_t, a_i) - (r_{i,t+1} + \gamma \max_{a' \in A} Q_i(s_{t+1}, a'))), \quad (6.1)$$

where learning rate α_W and delaying rate ω ($\omega > 0$) control the convergence of W_i .

Thus, WL can be seen as a fair resolution of competition. Competition results in fragmentation of the state space between the different agents, thus, allowing any collection of agents. Eventually, they will divide up the state space among them based on the deviations they cause in each other. The winner of a state (determined by the highest $W(s)$) is the agent that is most likely to suffer the highest deviation if it does not win. Eventually, agents are aware of their competition indirectly by the interference they cause.

6.1.2 Distributed W-learning

The DWL algorithm proposed in [27] enables an agent $A_i \in A = \{A_1, \dots, A_n\}$ to learn to select actions that match its local policies while learning how its actions affect its neighbors $A_j \in A$, and to give different weights to the preferences of its neighbors when selecting an action. To prompt an agent A_i to consider the action preference of its neighbors (i.e., to cooperate), each agent implements, in addition to its own local policy $LP_i = \{LP_{i1}, \dots, LP_{il}\}$, a *remote* policy $RP_i = \{RP_{ij1}, \dots, RP_{ijr}\}$ for each of the local policies LP_{jl} used on each of its neighbors. To help neighbor A_j implement its local policy, remote policy RP_i receives a reward r_{ijr} every time a neighbor's local policy LP_{jl} receives a reward r_{jl} ($r_{ijr} = r_{jl}$).

RP_i enables heterogeneous agents to cooperate, implement different policies, and have different actions and state spaces. Thus, the DWL scheme lets an agent adapt to the other agents since their dynamics are generally changeable. Each agent implements its policy as a combination of a Q-learning and a WL process. Q-values are associated with each of its state-action pairs, while W-values are associated with states. In the learning process, an agent A_i also learns Q-values for remote-state/local-action pairs and W-values for local/remote states, through which it acquires the influence of its local actions on the states of its neighbors A_j . Thus, DWL does not need a global knowledge or central component. It relies on local learning and interactions with its neighbors, local rewards from the environment, and local actions.

To learn how its actions affect its neighbors, at each control time step, the agent receives information about the current states of its neighbors and the rewards they have received. All local and remote policies nominate an action with an associated W-value. Nominations for LP_i actions are treated with full W-values. In contrast, RP_i nominations are scaled by a cooperation coefficient C ($0 \leq C \leq 1$) to enable an agent to weigh the action preferences of its neighbors. $C=0$ indicates a non-cooperative local agent, i.e., it does not consider the performance of its neighbors when picking an action. For $C=1$, the local agent is entirely cooperative, implying that it cares about its neighbors' performance as much as its own.

The action performed at the given control time step (one that wins the competition between policies) is selected based on the highest W-value (W_{win}) after scaling the remote W-values by C using the following equation:

$$W_{win} = \max(W_{il}, C \times W_{ijr}), \quad (6.2)$$

where W_{il} and W_{ijr} are W-values nominated by LP_i and RP_i policies of agent A_i , respectively.

6.2 Modeling VSL as a W-learning problem

Using a multi-agent approach based on the WL for VSL control states s_t , actions a_t , and reward functions r_{t+1} are defined as follows.

6.2.1 Agents environment formulation

State description

Since the measured speed can give an agent information about actions it takes, the speed as one of the state variables is used. State variables describing traffic flow speed are coded into variable V_i corresponding to measured average speed $\bar{v}_{i,t}$ [km/h] at time t in motorway sections

L_1 and L_2 ($i = 1, 2$) as presented in Figure 6.1.

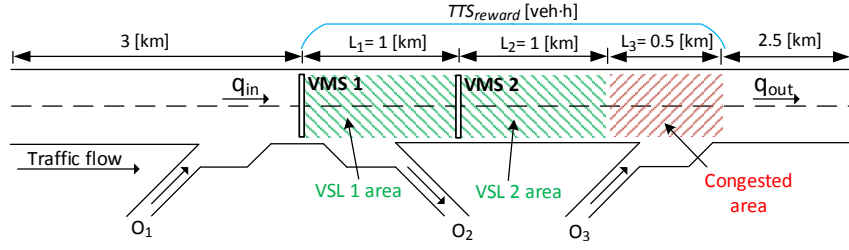


Figure 6.1: Motorway model

$$V_i = \begin{cases} v_1, & \text{if } 0 \leq \bar{v}_{i,t} < 54 \\ v_2, & \text{if } 54 \leq \bar{v}_{i,t} < 72 \\ v_3, & \text{if } 72 \leq \bar{v}_{i,t} < 90 \\ v_4, & \text{if } 90 \leq \bar{v}_{i,t} < 108 \\ v_5, & \text{if } 108 \leq \bar{v}_{i,t} \end{cases} \quad (6.3)$$

State variable ρ_3 contains information about current traffic density $\bar{\rho}_{3,t}$ [veh/km/lane] measured in congested section L_3 (red cross-hatched area in Figure 6.1) as a strong indicator for traffic congestion. According to $V_1 \times V_2 \times \rho_3 \mapsto S$ the number of possible states is $|S| = 225$.

$$\rho_3 = \begin{cases} \rho_1, & \text{if } 0 \leq \bar{\rho}_{3,t} < 15 \\ \rho_2, & \text{if } 15 \leq \bar{\rho}_{3,t} < 20 \\ \rho_3, & \text{if } 20 \leq \bar{\rho}_{3,t} < 25 \\ \rho_4, & \text{if } 25 \leq \bar{\rho}_{3,t} < 30 \\ \rho_5, & \text{if } 30 \leq \bar{\rho}_{3,t} < 35 \\ \rho_6, & \text{if } 35 \leq \bar{\rho}_{3,t} < 40 \\ \rho_7, & \text{if } 40 \leq \bar{\rho}_{3,t} < 45 \\ \rho_8, & \text{if } 45 \leq \bar{\rho}_{3,t} < 50 \\ \rho_9, & \text{if } 50 \leq \bar{\rho}_{3,t} \end{cases} \quad (6.4)$$

Action space

Each element in the action set (6.7) contains two variables. The upper part represents the speed limit in section L_1 , while the variable below is the speed limit in section L_2 . In that way, the winner agent in WL-VSL (one with the highest $W_i(s_t)$ value in state s_t) will define the speed limits for sections L_1 and L_2 . An additional restriction is added to action set A_{WL} . The absolute

speed difference ($\|a_{t,L1} - a_{t,L2}\| \leq 20$) between posted speed limits in consecutive sections must not be greater than 20 [km/h]. This is necessary to ensure smooth and safer speed transition (harmonization) for flow entering the bottleneck region. This may seem like an additional increase in state-action space, but in this way, there is no need to model restrictions directly into the reward function, keeping the reward function simpler and focused on optimization parameters. At the same time, WL-VSL is able to operate safely following the authority's recommendation about maximum allowed speed changes.

$$A_{WL} = \left\{ \begin{Bmatrix} 60 \\ 60 \end{Bmatrix}, \begin{Bmatrix} 60 \\ 80 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 60 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 80 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 100 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 80 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 100 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 120 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 100 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 120 \end{Bmatrix} \right\} \quad (6.5)$$

Finally, the action executed by the winner agent (A_k) at time step t is chosen from available subset $A_{WL}^* \subset A_{WL}$. E.g. if $a_{k,t-1} = A_{WL}(9)$, then available action subset at time t is $A_{WL}^* = \{A_{WL}(6), A_{WL}(7), A_{WL}(8), A_{WL}(9), A_{WL}(10)\}$. Similar is done in [21], where action selections are state-dependent.

Reward function

In [19] and [21], TTS has been successfully used as a measure of performance for RL-VSL. Based on these insights, one part of the used reward function is also based on TTS. An extra factor sensing the average speed induced by the posted speed limit in the section of each agent is added. In this way, one can ensure that the reward slightly varies for each agent at a time. An agent can minimize punishment (negative sign in reward) by achieving a higher average speed in a controlled area (L_1 and L_2), or by decreasing the TTS . Combining the speed part and TTS , an agent is forced to find a balance between both. E.g., in the case of an inactive bottleneck, the penalty will be smaller if the posted speed limit is higher. On the other hand, when a bottleneck is active but speed limits are high, drivers have no information about the downstream congested area of the motorway. Vehicles will enter the jam at high speed. Vehicles that decelerate suddenly will create shock waves, greater congestion and higher TTS . In this way, reward (6.6) effectively provides the agent with a more direct signal about the consequences of its actions. Thus, TTS from sections L_1 and L_2 is scaled since section L_3 is half as long. Incorporating speed into the reward system enables the agent to more directly perceive the impact of its actions, potentially accelerating the learning process. In (6.6), $\bar{v}_{i,t+1}$ is the average speed (see (6.3)), δ [km/h] is the scale parameter, and TTS_i is total time spent [veh · h] measured in section $i = 1, 2, 3$ (see Figure 6.1) between two control time steps t and $t + 1$.

$$r_{i,t+1} = - \left[\left(1 - \frac{\bar{v}_{i,t+1}}{\delta} \right) \left(\frac{1}{2} (TTS_1 + TTS_2) + TTS_3 \right) \right] \quad (6.6)$$

Winner action

In the case of WL-VSL, Q_i -matrix has dimension $m \times n$, where $m = 225$ is equal to the number of states and $n = 10$ is the number of actions in A_{WL} . W_i has dimension 225×1 . Using gained knowledge from Q_i , all agents are able to suggest new actions, but based on highest $W_i(s_t)$ value, only the leader's action a_k is executed. After the state transition $s_t \mapsto s_{t+1}$ each agent receives its unique reward r_i depending on the consequences of executed action a_k (speed limits). After that, all agents are updating their Q_i -function (for winner action a_k), while only agents who were not obeyed update their W_i value [55].

How fast agents learn the Q_i -function depends on the parameters α_Q , and γ . The convergence of W_i is controlled by α_W and parameter ω . Exploration/exploitation is controlled using the ϵ -greedy mechanism explained in the next section.

6.2.2 Simulation setup

The proposed WL-VSL is compared with baselines independent RL-VSL agents (IND-VSL) and single-agent RL-VSL (SGL-VSL), to assess its benefit over previous work (single-agent RL-VSL) [53], and to assess the benefits of the cooperation using WL over IND-VSL.

6.2.3 Simulation model

The applied simulation framework consists of the microscopic simulator SUMO and the programming environment Python [124]. The motorway model used for testing WL-VSL is based on the model used in [53]. It is divided into three main sections L_1 , L_2 and L_3 . Two sections ($L_1 = L_2 = 1$ [km]) are equipped with VMSs and controlled by VSL, while a bottleneck is created in motorway section $L_3 = 0.5$ [km] (see Figure 6.1). Every simulation lasted 1.5 hours. All VSL approaches were learned on 6,000 simulations. For testing the learned control policy, an additional 200 runs using only exploitation ($\epsilon \mapsto 0$) were performed.

A new speed limit value is sent to the VMS during every control time step T_c (in this MARL-related part $T_c = 150$ [s]). Parameters taken for evaluation of VSLs strategies are overall TTS , measured for the entire motorway stretch (including all ramps), and TTS_{reward} as part of objective optimization (reward function). Macroscopic parameters such as average flow speed and density, as well as occupancy on the acceleration lane in the bottleneck area, are taken as valid indicators of the efficiency of upstream applied VSL to control the downstream bottleneck.

Traffic scenarios

To assess WL-VSL learning stability and benefits of cooperation using WL over baseline approaches, two traffic scenarios, dynamic and static, were used for evaluation.

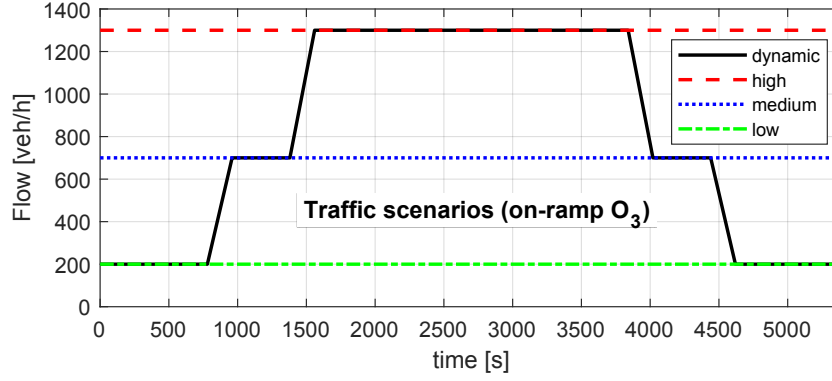


Figure 6.2: Tested traffic scenarios

Dynamic traffic scenario In downstream section L_3 (see Figure 6.1), the presence of a bottleneck is being induced by an increase in traffic demand at on-ramp O_3 . The induced downstream bottleneck is the main test for the proposed WL-VSL. In this scenario, the traffic demand at on-ramp O_3 varies over time (see Figure 6.2). For the highest demand 1,300 [veh/h], the mainstream traffic flow is influenced by merging vehicles, and congestion occurs (active bottleneck). Traffic flow consists of 94% passenger cars, 3% trucks, and 3% of buses. The mainstream flow entering the bottleneck area has a constant demand of 2,700 [veh/h]. The driving dynamics of vehicles in SUMO were modeled using the *Krauss* Car-Following Model.

Static traffic scenario The static scenario with constant traffic demand is used to confirm stability when WL-VSL operates in different traffic regimes. This is needed because before and after peak hours, traffic is stable without high deviations in demand. Therefore, WL-VSL has to be able to perform optimally and safely in such conditions. For that, three different cases were used for testing all approaches regarding the traffic demand at on-ramp O_3 ; low = 200, medium = 700, and high = 1,300 [veh/h], respectively.

Baselines SGL-VSL and IND-VSL

For SGL-VSL and IND-VSL, the available action set is $A = \{60, 80, 100, 120\}$ [km/h]. In the case of SGL-VSL, the same speed limit value is sent on both VMSs in motorway sections L_1 and L_2 . In the case of two independent VSL agents (IND-VSL), each agent is able to post the speed limit in its section without any coordination between them. The executed action $a_t \in A$ should follow the authority rules as explained in the action section. Eventually, an agent is limited to choosing an action from the $A^* \subset A$ according to which action was mandatory in the previous time step a_{t-1} . In IND-VSL, each agent received a reward, as explained in the reward section. Because SGL-VSL is a single-agent RL and the same speed limit is set in both sections and, thus, has a similar effect on the flow in both sections, it takes average speed in L_1 in order to calculate reward (6.6). Q_i -values in case of SGL-VSL and IND-VSL are stored in a 225×4

Q_i -matrix, according to the number of possible states and actions. How fast the agents learn Q_i depends on α_Q and γ , as mentioned above. In this MARL-related part, the values are kept constant $\alpha_Q = 0.5$ and $\gamma = 0.8$, according to [27, 53]. Most of the time, an agent picks good (learned) actions but also does random exploration. This is controlled by the parameter ε which is decreased with the number of simulation runs N using function $\varepsilon = \exp \frac{-\ln(20)N}{6000}$. Bigger ε values let the agent behave randomly in the beginning. As ε gets smaller, the agent starts to utilize its learned knowledge.

WL-VSL parameters

For updates of Q_i and W_i , values in case of WL-VSL, $\alpha_Q = 0.5$, $\alpha_W = 0.5$, while $\omega = 0.3$ has been slightly lower compared to [55] and selected from multiple tests. Parameter δ , used to scale measured average speed $\bar{v}_{i,t+1}$ in the applied reward function, is set to 150 [km/h]. As the maximal allowed speed limit is 120 [km/h], this way, it is ensured that in case of light traffic, part of reward function $(1 - \frac{\bar{v}_{i,t+1}}{\delta})$ will not switch the sign if any vehicle goes slightly faster than allowed. Parameter ε is reduced as in the baseline cases.

6.2.4 Results and analysis

In this part, the simulation results are presented and the efficiency of WL-VSL through the observation of traffic parameters measured within the bottleneck and the entire motorway network is investigated.

Dynamic traffic scenario

Results obtained during the learning process in the case of dynamic traffic scenarios are given in Figure 6.3 and in Table 6.1.

Analysis of control policy

WL-VSL From Figure 6.3, it can be seen that WL-VSL has started applying actions slightly before the density began to increase, indicating proactive control noticed in [20]. Mostly the speed limit is active in section L_1 . While density was high, speed limits have been alternated between values 100 and 120 [km/h]. As density gets lower, WL-VSL recognizes free-flow conditions, and the speed limit is increased to the maximum of 120 [km/h], which remains mandatory until the end of the simulation. The alternating actions between sections could be due to the extensive length of the managed motorway region (sections L_1 and L_2) where both agents operate. If speed limits in both sections were active at the same time, this could gain higher TTS (because of the lower speeds in L_1 and L_2) in such amount that is better to split

activity to one of the sections. Agents in that case compromise, either section L_1 or L_2 will reduce the speed limit when needed.

Baselines SGL-VSL and IND-VSL In case of IND-VSL, only the agent in L_2 reduces the speed from 120 to 80 [km/h] and starts increasing speed limits immediately after reaching the peak density value (see Figure 6.3). This is desirable behavior of the IND-VSL controller since it recognizes when a bottleneck is active or tends to be inactive. The agent in L_1 remains inactive at all times, possibly for the same reason explained in the previous paragraph: the sections are too long.

SGL-VSL is inactive most of the time. The extended length of the managed segment may be the reason as a single agent controls both sections simultaneously. Thus, it cannot delegate decisions to other agents, and in this way, it will produce a higher TTS . Therefore, the agent often determines that inactivity is the optimal strategy to meet its defined objectives.

Density analysis

WL-VSL Traffic density is analyzed since it is included as a state variable in part of the RL algorithm used for all VSL approaches. The second reason is that density is a reliable indicator for the presence of a bottleneck [20]. In the density graph given in Figure 6.3, the positive effect of actions selected by WL-VSL is most visible. Actions selected by WL-VSL efficiently managed to reduce the density during the congested period compared with all other baselines. Compared with no-control, density was reduced up to 18% using WL-VSL. This is desirable because, by reducing the density, it is possible to dissolve the active bottleneck faster or prevent the activation of the bottleneck. This result proves that WL-VSL can be used for controlling existing bottlenecks.

Baselines SGL-VSL and IND-VSL Compared to WL-VSL, both IND-VSL and SGL-VSL achieve a slightly lesser reduction in density. By applying the lower speed limits, IND-VSL tries to create an artificial bottleneck before the real one. This shows less effect in the reduction of the density in the bottleneck region. In addition, performing an action too early and after the creation of a bottleneck in the case of SGL-VSL cannot reduce density as much as WL-VSL did.

Speed analysis

WL-VSL The timely applied sequence of speed limits in motorway sections L_1 and L_2 keeps the traffic flow speed measured in section L_3 at a higher value compared with base cases. The increase in average speed using WL-VSL is 7.3% compared to no-control, indicating more stable traffic flow within section L_3 . This is confirmed with the easier merging of vehicles from

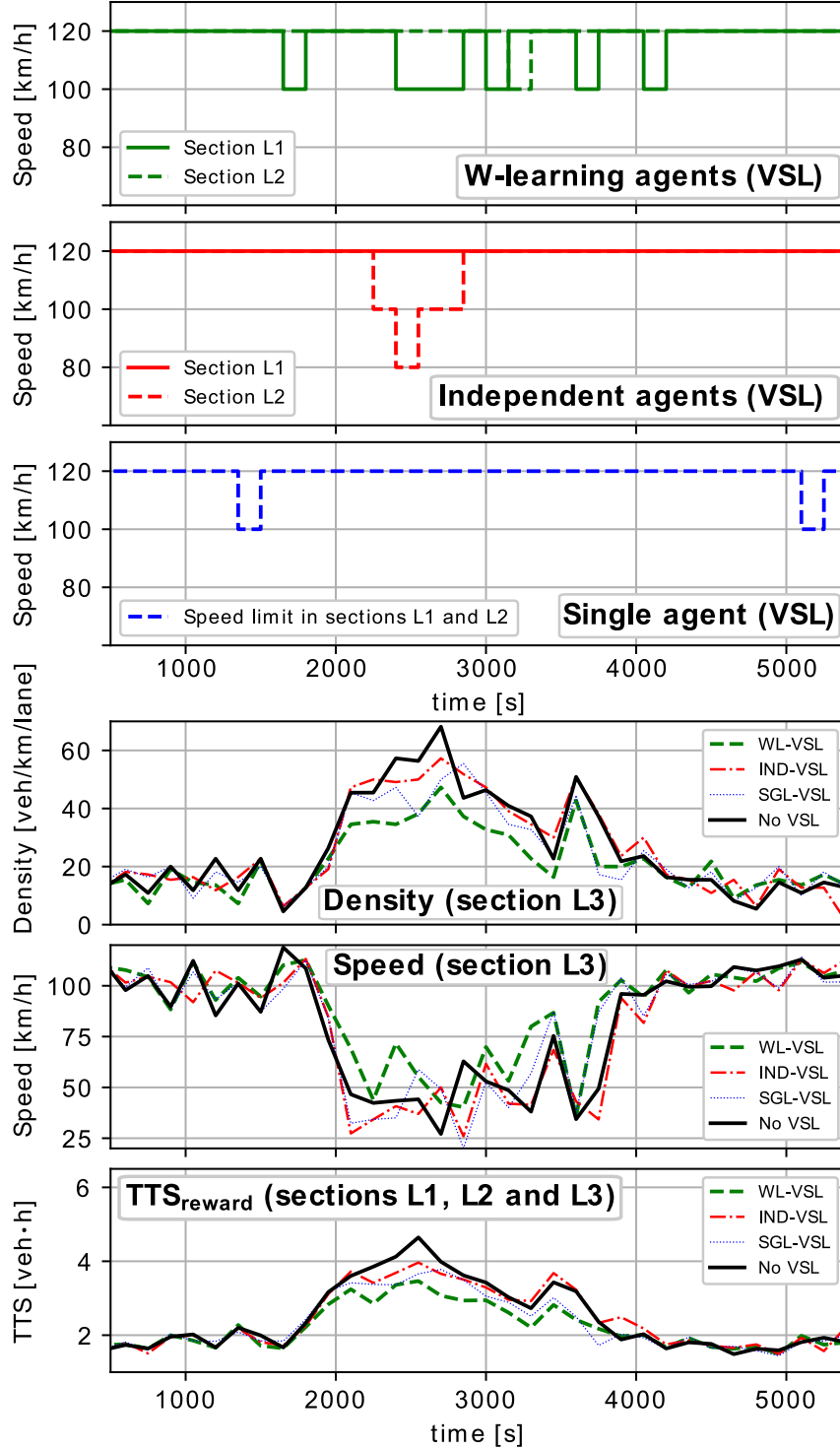


Figure 6.3: Results in the dynamic traffic scenario

on-ramp O_3 into mainstream traffic in section L_3 , where less occupied acceleration lane (for 28%) using WL-VSL is achieved compared with the no-control case.

Baselines SGL-VSL and IND-VSL In the case of SGL-VSL, there is no significant improvement in the average speed. In the case of IND-VSL, the average speed measured within L_3 was lower for 2% compared with no-control. The reason for that is a stronger decrease in the speed

Table 6.1: VSLs PERFORMANCE RESULTS

	No VSL	WL-VSL		IND-VSL		SGL-VSL	
		Obtained	Improv. [%]	Obtained	Improv. [%]	Obtained	Improv. [%]
TTS_{reward} [veh · h]	2.37	2.14	9.70	2.35	0.84	2.24	5.48
Avg. Speed in L_3 [km/h]	83.4	89.5	7.31	81.7	−2.03	83.7	0.36
Avg. Density in L_3 [veh/km/lane]	25.3	20.7	18.18	25.1	0.79	23.4	7.51
Avg. occupancy of accel. lane in L_3 [%]	7	5	28.57	8	−14.28	7	0

limit in section L_2 (80 [km/h]). Occupancy of the acceleration lane, in that case, was higher by 14% compared with no-control, indicating more difficult merging of vehicles in the mainstream flow. Thus, creating the stronger artificial bottleneck in section L_2 in front of the congested section (L_3) deteriorates the traffic in the downstream bottleneck. As shown in [28], in addition to speed limits, positioning and choosing the appropriate length of the section covered by the VSL is the crucial part of the optimal utilization of the VSL system. Therefore, the position and length of section L_2 can be one of the reasons for the suboptimal behavior of applied speed limits in section L_2 in case of IND-VSL.

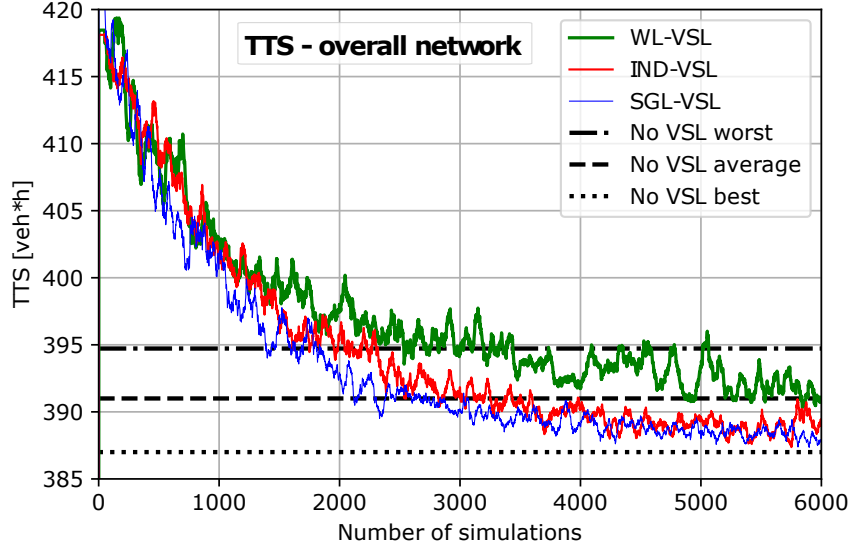
TTS reward analysis

WL-VSL From Figure 6.3, it can be seen that WL-VSL reduces the amount of TTS_{reward} significantly during the period of high density in section L_3 compared to baselines. This is desirable since the reward function (6.6) is modeled in a way to minimize TTS within the motorway sections L_1 , L_2 , and L_3 . An average improvement of 9.7% (Table 6.1) compared to the no-control case is achieved by using WL-VSL.

Baselines SGL-VSL and IND-VSL Improvement of measured TTS_{reward} is also achieved for both (SGL-VSL and IND-VSL) but in a lower amount compared to WL-VSL.

TTS - overall network

In Figure 6.4, a comparison of convergence of overall TTS measured cumulatively during the simulation in the entire motorway network (including all on- and off-ramps) is shown. At the beginning of the learning process, all VSLs performed worse compared to the no-control case since they explore the world in which they operate. As simulations progress, TTS consistently decreases. After the 6,000 simulations (using dynamic traffic scenario), TTS for all VSLs converges around the average TTS value obtained in simulations without VSL control. Eventually, compared with the starting values, the overall TTS has been improved by around 7% for all baselines and WL-VSL cases. It can be seen that WL-VSL has a slightly slower convergence rate, but after a higher number of simulations, it gradually approaches other methods. The reason for a longer convergence time to a steady state may be the larger action set (10 actions)

Figure 6.4: Convergence of TTS during training process

which considers safety restrictions during action selection in WL-VSL. In case of SGL-VSL and IND-VSL, the number of available actions is 4. Due to the bigger state-action space, WL-VSL needs more time (simulations) to explore and learn optimal control policy. Eventually, overall TTS confirms that WL-VSL does not worsen upstream traffic while being applied for local downstream bottleneck optimization.

Static traffic scenario

After training the agents on the dynamic traffic scenario, their action policy is tested on a static traffic scenario. Results are presented in Figure 6.5. In the static traffic scenario, three different traffic loads influencing traffic demand at on-ramp O_3 were used (low, medium, and high, see Figure 6.2). As the results follow the same pattern, the medium scenario (700 [veh/h]) is selected, which corresponds to realistic traffic conditions before and after the congested period.

Figure 6.5 shows that all controllers maintain stable values for the observed traffic parameters (density, speed, and TTS_{reward}). WL-VSL and IND-VSL are activated on a slight increase of the density value in the downstream section L_3 at the beginning of simulations. When the density decreases, both controllers increase the speed limit to the maximum value of 120 [km/h] since density in the bottleneck area L_3 remains low. WL-VSL strategy is active for a longer period and adapts speed limits on both sections L_1 and L_2 . When IND-VSL is applied in the static scenario, the agent on section L_1 is active, while in the dynamic scenario, the agent on section L_2 is active. SGL-VSL stays inactive for the full duration of the simulation. In general, all approaches, including WL-VSL, did not worsen traffic performance, which confirms that WL-VSL could be applied as a stable and efficient VSL control strategy in both static and dynamic traffic scenarios.

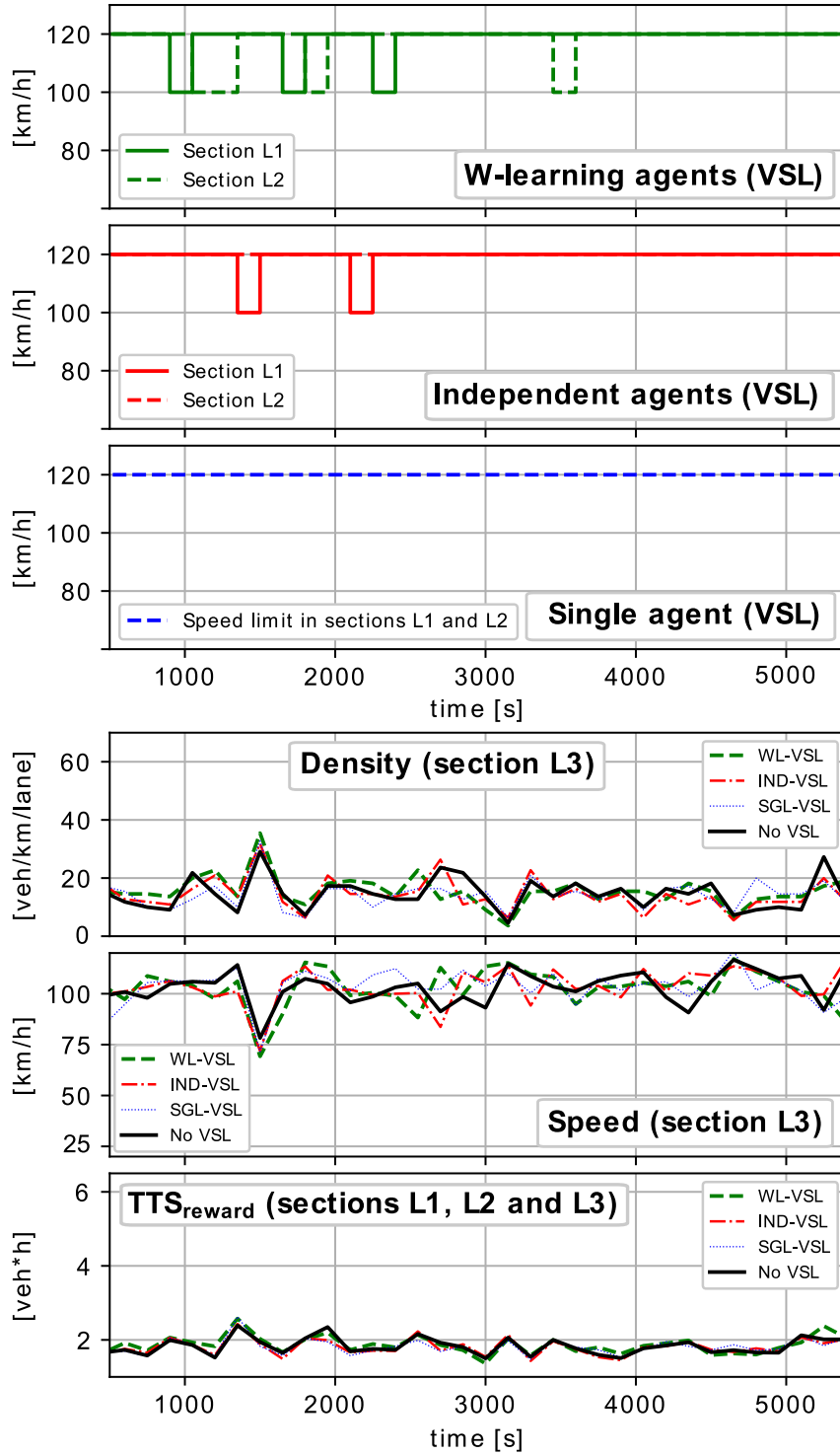


Figure 6.5: Results in the static traffic scenario

6.2.5 Discussion

The discussion in this section is restricted by the design of WL-VSL: a novel multi-agent RL-based VSL control approach. WL-VSL is implemented on a motorway simulation scenario where two agents are learning, using the W-learning algorithm, to control two segments upstream of a congested area jointly. The proposed approach is compared to a single agent and

two independent agent implementations, and it has been shown that WL-VSL outperforms these baselines with respect to measured traffic parameters TTS, density, and average speed in the downstream congested motorway area. Roughly speaking, WL-VSL confirms that the multi-agent VSL approach has benefits over single-agent VSL and, thus, can be used to control the active bottleneck efficiently. When facing a high density at the bottleneck, WL-VSL has been proven to produce a more stable traffic flow downstream by adapting speed limits gradually throughout the VSL application area and in this way limiting the need for more restrictive actions such as a sudden decrease in speed limit values, which could negatively impact the flow upstream of the controlled sections. Therefore, in the context of the next section, it is assessed if additional cooperation among agents using DWL can improve VSL control even further.

6.3 Modeling spatiotemporal VSL as distributed W-learning problem

In this section, the control algorithm for the spatiotemporal dynamic VSL zones adaptation model proposed in section 4.3 is described. For this, the development process of DWL-ST-VSL is explained in a methodological way, starting with the introduction of DWL. Then, the DWL-ST-VSL control problem as a two-agent control problem in resolving bottlenecks is formulated. To verify the extended applicability of DWL-ST-VSL and to assess whether the inclusion of additional agents can improve the spatiotemporal VSL control, experiments with four agents are performed.

6.3.1 Configuration with two agents

DWL has been successfully applied to the problem of controlling multiple traffic lights [27], yet never to VSL. In the proposed DWL2-ST-VSL framework, two neighboring agents ($A_i, i = 1, 2$) control the speed limit and VSL zone configuration (position and length) each on their own motorway section. Each agent perceives its local environment through the agent states and receives a reward (see Figure 6.6).

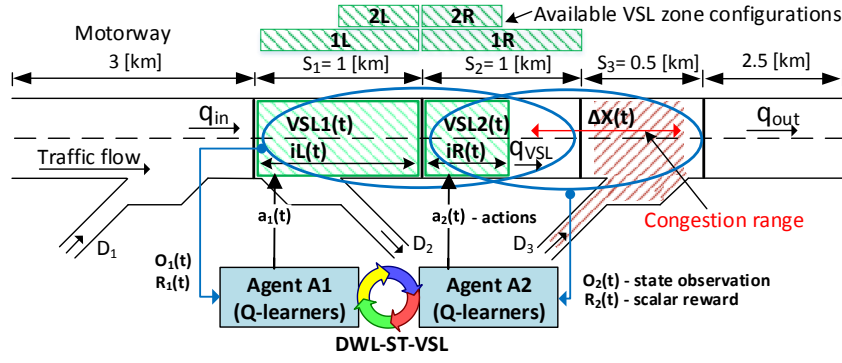


Figure 6.6: DWL2-ST-VSL configuration scheme

To solve the multi-agent optimization problem modeled in DWL2-ST-VSL on multiple conflicting policies, the agent states s_t , actions a_t , and reward functions r_{t+1} are defined as follows.

State-action space and rewards description

In continuation, the state-action space of the agents is elaborated along with the associated reward functions.

State description

Information about the speed of the neighboring agent segment can improve the learning process, especially when agents cause interference by randomly executing speed limits (exploration phase). Moreover, speed indicates traffic flow disruption caused by congestion. Speeds are encoded in the variable V_n , which corresponds to the measured average vehicle speeds $\bar{v}_{n,t}$ at time t in motorway section S_n ($n = 1, 2, 3$), as shown in Figure 6.6. It has been decided on four intervals defined with boundary points (50, 76, 101 [km/h]) where each speed measurement can fall into one interval.

The state variable ρ_n contains information about the current traffic density $\bar{\rho}_{n,t}$ measured in the motorway section S_n . Each density measurement can fall into one of twelve intervals defined with boundary points (15, 20, 23, 26, 29, 32, 35, 38, 45, 55, 65 [veh/km/lane]). Density uniquely determines whether the traffic state is in the free-flow, saturated, or oversaturated state. Since the agents' reward depends on TTS, it is beneficial to have an overview of the density of the motorway section, because density is an independent variable in the computation of TTS. Also, the agent state contains information about the agent's action from the previous control time step.

Therefore, A_1 's local policy LP_{11} at time t senses state $s = (\rho_1, \rho_2, V_2, a_{1,t-1})$, while LP_{12} $s = (\rho_1, \rho_2, V_1, a_{1,t-1})$. A_2 's local policy LP_{21} senses state $s = (\rho_2, \rho_3, V_3, a_{2,t-1})$, while LP_{22} $s = (\rho_2, \rho_3, V_2, a_{2,t-1})$, see Figure 6.6.

Action space

Each element in the action set (6.7) contains two variables. The upper one represents the speed limit in section S_n , while the lower one represents the VSL zone configuration (indexes for the left/right configuration labels iL and iR , see Figure 6.6). Agent A_1 controls the speed limit and the length of the VSL zone in section S_1 , while A_2 controls section S_2 . In this way, the agent winning policy (either local or remote), will define the speed limit and the VSL zone configuration for a given motorway section. The absolute difference between two speed limits within a section (n) must not exceed 20 [km/h]. This is necessary to ensure a smooth and safe speed transition between the upstream free-flow and the downstream slower flow caused by the bottleneck. This constraint implies that the next possible action in $\gamma \max_{a' \in A} Q_i(s_{t+1}, a')$ (see (5.1)) must be bounded based on a_{t-1} . Thus, each time the Q-value is updated, a possible subset of the allowed actions is considered. In this way, it is not necessary to model constraints directly in the reward function. Thus, the reward function focuses on the optimization parameters. At the same time, DWL2-ST-VSL is able to operate safely by adhering to the recommended rules for maximum allowed speed changes. It is important to note that spatial variation in speed limit between two adjacent VSL zones on the motorway is not considered in this setup.

$$A_{i,DWL} = \left\{ \begin{Bmatrix} 60 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 60 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 2 \end{Bmatrix} \right\} \quad (6.7)$$

Reward function

In [21], TTS was successfully used as a performance measure for RL-VSL. Therefore, the TTS measures are used for the reward. The variable $TTS_{n,t+1}$ measures TTS between two control time steps t and $t + 1$ on section n , signaling to the agent how good its action was. Every agent must learn to balance between two conflict policies simultaneously. In the case of an inactive bottleneck, the penalty will be lower for a higher speed limit. Conversely, when congestion occurs, it is necessary to gradually reduce the speed limit upstream in order to control the inflow into the congested area so that the traffic volume is maintained near the operational capacity of the active bottleneck. Therefore, each policy seeks to optimize its objective as follows.

Modeling local and remote policies

To optimize the local agent's goal, local policies are used. To determine the neighbors' preferences, the agents also learn the remote policies. According to the defined reward functions, each policy tries to optimize its goal as follows.

Local policy for free-flow control The first local policy LP_{i1} of an agent A_i , aims to learn the speed limit to ensure a reduction of TTS by promoting, when possible, speed in free-flow traffic (average vehicle speeds above $102 \left[\frac{km}{h}\right]$). To achieve this goal, LP_{i1} reward is:

$$r_{LPi1,t+1} = \begin{cases} 0, & \text{if } \min\{\bar{v}_{n,t+1} \mid n = i, i+1\} \geq 102 \\ -TTS_{n,t+1}, & n = i \text{ otherwise} \end{cases} . \quad (6.8)$$

In a certain percentage, LP_{i1} is activated in saturated flow during the transition from free-flow to congested flow and vice versa. Therefore, it prepares traffic for the second policy (LP_{i2}), which dominates in oversaturated conditions. After congestion, LP_{i1} helps to restore traffic back to free-flow as soon as possible by gradually increasing the speed limit.

Local policy for congested traffic control The second local policy, LP_{i2} , aims to reduce TTS in the downstream section in the presence of an active bottleneck. Thus, an agent must learn appropriate speed limits to restrict the bottleneck inflow until the discharge capacity is restored. Otherwise, it will increase its penalty proportionally:

$$r_{LPi2,t+1} = -\beta TTS_{n,t+1}, \quad n = i+1, \quad (6.9)$$

where β determines how sensitive the agent is to congestion.

Remote policies The cooperation between agents is based on remote policies. Thus, an agent A_i learns two additional remote policies (RP_{ij1} and RP_{ij2}) that complement its neighboring agent's local policies. In order to know how its local actions affect its neighbors' states, the agent updates the remote policies on the information it receives about its neighbors' current states and the rewards that neighbor agents have received (Figure 6.6).

Winner action selection

In the case of DWL2-ST-VSL, an agent A_i 's experiences (Q-values for local-state/action pairs and Q-values for remote-state/local-action pairs) for each policy are respectively stored in Q_{ik} matrices ($k = 1, \dots, 4$). At the same time, for each of the states of its policies, an agent learns W-values of what happens in terms of the reward received if the action nominated by that policy is not performed [55]. This is expressed as a W-value ($W(s_{i,t})$) and stored in W_{ik} matrices in each case. With the knowledge gained from these matrices, all policies (local and remote) propose new actions. The action $a_{k,t}$ that wins the competition between policies at this time step is the one with the highest W-value (W_{max}) (computed using (6.2)) [27]. After the state transition $s_t \mapsto s_{t+1}$, each agent's local policy receives its unique reward ($r_{LPi1,t+1}$, $r_{LPi2,t+1}$) and state ($s_{LPi1,t+1}$, $s_{LPi2,t+1}$) depending on the consequences of the executed action $a_{k,t}$. The remote

policies RP_{ijr} obtain rewards and state information from their neighbor agent by querying the neighbor's local policies states/rewards ($s_{LPj1,t+1}$, $r_{LPj1,t+1}$, $s_{LPj2,t+1}$, and $r_{LPj2,t+1}$). Then, all policies update their Q-values (for the winning action $a_{k,t}$), while only the policies that were not obeyed update their W-value. The above process is repeated for all agents.

The summary of the execution of the DWL2-ST-VSL model is presented by Algorithm 6.1. In addition, readers are also referred to the original paper [27] where the DWL algorithm is proposed.

Algorithm 6.1 DWL2-ST-VSL at each learning step

```

// Set DWL2-ST-VSL parameters and load sim. model
for each simulation step
  if simulation time % 150 [s] == 0 then
    for each  $A_i$  in  $A$ 
      for each  $LP_{il}$  in  $LP_i$ 
        Get  $LP_{il}$ 's traffic states,  $s_{il,t}$ 
        Get rewards from  $A_i$ 's environment
        Update  $Q_{LPil}(s_{il,t-1}, a_{il,t-1})$  for  $LP_{il}$ 
        Update  $W_{LPil}(s_{il,t-1})$ 
        Nominate action  $a_{il,t}$  with max  $Q_{LPil}$  for  $s_{il,t}$ 
        Get  $W_{LPil}(s_{il,t})$ 
      end for
      // Get nomination by remote policies
      for each  $RP_{ijr}$  in  $RP_i$ 
        Get  $RP_{ijr}$ 's traffic states,  $s_{ijr,t}$ 
        Get rewards for  $s_{ijr}$  from  $A_j$ 's environment
        Update  $Q_{RPijr}(s_{ijr,t-1}, a_{ijr,t-1})$  for  $RP_{ijr}$ 
        Update  $W_{RPijr}(s_{ijr,t-1})$ 
        Nominate action  $a_{ijr,t}$  with max  $Q_{RPijr}$  for  $s_{ijr,t}$ 
        Get  $W_{RPijr}(s_{ijr,t})$ 
      end for
      // Select and perform action (set new speed limit and VSL zone configuration)
      Compute and execute winning action  $a_k$  (using formula ( 6.2))
    end for
  end if
end for

```

6.3.2 Configuration with four agents

So far, DWL has been successfully applied to the problem of controlling urban intersections on a larger scale network with a higher number of agents [27]. DWL has also proven successful in the VSL control optimization problem [26] on a smaller motorway segment in the configuration of two VSL agents (DWL2-ST-VSL). Nevertheless, it has never been tested for its extended applicability to motorway traffic control with a higher number of deployed VSL agents. Thus,

in the proposed extended DWL4-ST-VSL framework, four neighboring agents ($A_i, i = 1, 2, 3, 4$) control the speed limit and VSL zone configuration (length and position) on their own motorway section.

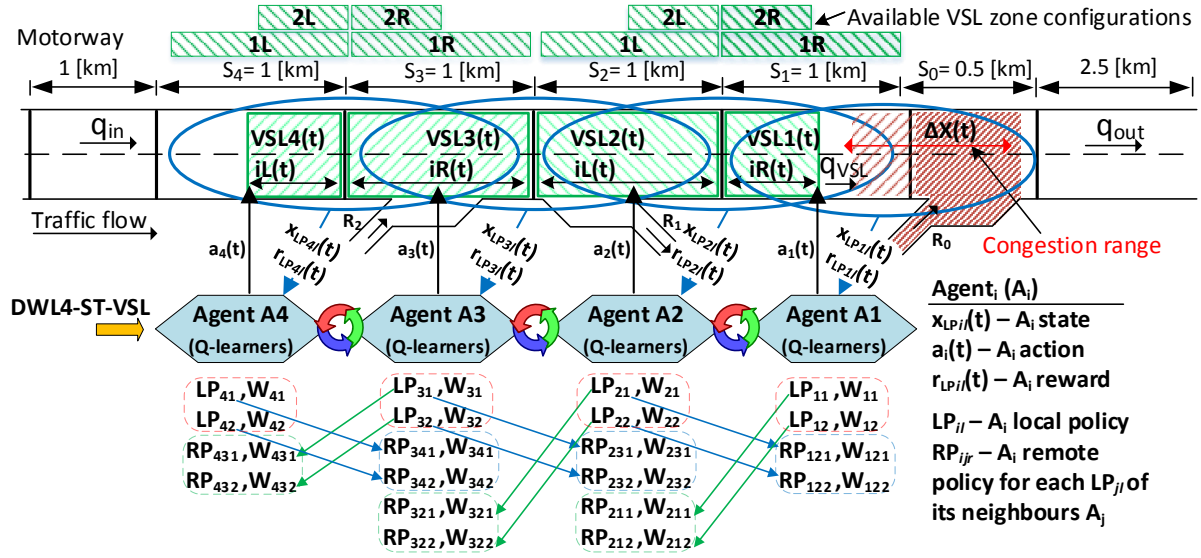


Figure 6.7: DWL4-ST-VSL scheme and configuration of motorway segment

Each agent in DWL4-ST-VSL perceives its local environment through agent states and rewards (see Figure 6.7). Thus, in the proposed multi-agent control optimization problem, the agent states s_t , actions a_t , and reward functions r_{t+1} are modeled as follows.

State-action space and rewards description

In continuation, the state-action space of the DWL4-ST-VSL configuration is elaborated along with the associated reward functions.

State description

As stated in [21], defining a compact Markovian state representation for motorways is difficult because many external factors influence traffic flow: e.g., weather conditions, motorway geometry (curvature, slope), etc., which are hard to model precisely. Augmenting the state by additional information, such as observing more sections (e.g., the density measured on the motorway section further upstream from the congestion location and the on-ramp queue length, primarily to provide a predictive component in terms of motorway demand [24]) or including information from the past in states, may improve the algorithm's performance. Though this increases solution space, it can be overcome by the function approximation technique [21, 53]. However, in DWL modeling, the observation of the agent's neighborhood is available through remote policies. Nevertheless, the observability of the state must be assured. An example of

a partially observable state is the usage of traffic flow rate for states. From traffic flow theory, macroscopic variables describe traffic conditions (speed, density, flow). As a result of the non-linearity of the fundamental diagram (flow-density relationship) [38], the same traffic flow rate can be observed for a density value below critical density with high speed (stable flow) and a density value above critical density with low speed (unstable flow). Thus, the traffic condition is uniquely determined by using traffic density information. Therefore, speed and density measurements to omit the agents' confusion are used in the proposed approach, thus, uniquely determining traffic conditions. As a result, the negative effect of imperfect and incomplete perception of agents' partially observable states in MDP modeling is reduced.

Including the speed measurement of the neighboring segments into the state can enhance the learning process, particularly at the beginning of the learning process, when agents cause interference by randomly performing actions (exploration). Besides, low speed indicates traffic flow disruption provoked by congestion. Speeds are encoded in the variable V_n , which corresponds to the measured average vehicle speed $\bar{v}_{n,t}$ at time t in motorway section S_n ($n = 0, 1, 2, 3, 4$), as shown in Figure 6.7. Each speed measurement can fall into one of four intervals identical to those defined earlier in case of DWL2-ST-VSL configuration.

Current traffic density $\bar{\rho}_{n,t}$ measured in the motorway section S_n , is stored in the variable ρ_n . Each measurement can fall into one of twelve intervals defined identically as earlier, in the case of two agents configuration. Additionally, the state space contains information about the agent's action from the previous control time step, thereby enabling modeling restrictions on the action space by making it state-dependent, which is explained in more detail in the following section.

Therefore, A_1 's local policy LP_{11} at time t senses state $s = (a_{1,t-1}, V_0, \rho_0, \rho_1)$, while LP_{12} $s = (a_{1,t-1}, V_1, \rho_0, \rho_1)$. A_2 's LP_{21} senses state $s = (a_{2,t-1}, V_1, \rho_1, \rho_2)$, while LP_{22} $s = (a_{2,t-1}, V_2, \rho_1, \rho_2)$. Similarly, A_3 's LP_{31} senses state $s = (a_{3,t-1}, V_2, \rho_2, \rho_3)$, while LP_{32} $s = (a_{3,t-1}, V_3, \rho_2, \rho_3)$. Finally, A_4 's LP_{41} senses state $s = (a_{4,t-1}, V_3, \rho_3, \rho_4)$, while LP_{42} $s = (a_{4,t-1}, V_4, \rho_3, \rho_4)$ (see Figure 6.7).

Action space

Each element in the action sets (6.10) and (6.11) consists of two variables. The upper one represents the speed limit $[km/h]$ in section S_n , while the lower one represents an active VSL zone (indexes define the left (iL)/right (iR) configuration; see Figure 6.7). Agent A_1 controls the speed limit and the length of the VSL zone in section S_1 , while A_2 controls section S_2 , and so on. In this way, the agent's winning policy (either LP_i or RP_i) will define the speed limit and the

VSL zone configuration for a given motorway section.

$$\mathcal{A}_{1,2,DWL} = \left\{ \begin{Bmatrix} 60 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 60 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 80 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 2 \end{Bmatrix} \right\} \quad (6.10)$$

$$\mathcal{A}_{3,4,DWL} = \left\{ \begin{Bmatrix} 90 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 90 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 100 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 110 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 110 \\ 2 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 120 \\ 2 \end{Bmatrix} \right\} \quad (6.11)$$

Q-values in (DWL2 and DWL4)-ST-VSL are stored in a $Q_{|S| \times |A_{DWL}|}$ matrix, where S is a finite set containing the indices of the coded states of the Cartesian product of the input traffic variables ($|S| = 4,608$ and $|A_{DWL}| = |A_{1,2,DWL}| = |A_{3,4,DWL}| = 8$). This seems to be a large solution space for learning optimal Q-values using (5.1). Nevertheless, the feasible solution space was reduced by constraining the action selection in the nomination process explained further on. Thus, the resulting Q-matrix can be considered a sparse matrix, and there is no need to search the whole space.

The consecutive speed limit change within a section (n) must satisfy the constraint $|a_{t-1,n} - a_{t,n}| \leq 20$ in the case of agents A_1 and A_2 , which use action set $A_{1,2,DWL}$. In the case of A_3 and A_4 ($A_{3,4,DWL}$), the constraint is $|a_{t-1,n} - a_{t,n}| \leq 10$. This ensures a smooth and safe speed transition between the upstream free-flow and the congested downstream flow, characterized by lower vehicle speeds due to the bottleneck. Thus, the final set of actions allowed for the agent A_i at time t depends on the previously executed actions of the agent. This constraint also implies that the next possible action (a') in the update process of the W- and Q-values (see update rules (5.1) and (6.1)) must be bounded based on a_t . Thus, each time the Q-value is updated, a possible subset of the allowed actions is considered. E.g., if $a_{k,t-1} = A_{1,2,DWL}(7)$, then the available action subset at time t is $A_{1,2,DWL}^* = \{A_{1,2,DWL}(5), A_{1,2,DWL}(6), A_{1,2,DWL}(7), A_{1,2,DWL}(8)\}$. Therefore, the previous action in the state space is used to uniquely distinguish between states' transitions given the constrained subset of actions between control time steps. This constraint is implicitly modeled in the update rule (5.1). It addresses a unique row in Q-matrix ($Q(s, a)$) and the reachable entries in that row, corresponding to a given action index. Feasible entries in the particular row correspond to the original indexes of elements from the original action set. Thus, only such entries in $Q(s, a)$ are reachable in updating Q- and W-values and in the action nomination process while using *argmax* in the Q-learning algorithm [125]. Otherwise, the oscillation in the values of elements in a particular state (row) will be present. Thus, Q-values will not converge to a stationary policy, and action nomination in a particular state will constantly switch no matter how long the learning period is. Eventually, a stable agent diminishes the nonstationarity effect in the learning process of the other agents.

In this way, it is not necessary to model constraints directly in the rewards. However, it is

still ensured that DWL4-ST-VSL operates according to the advised safety rules on maximum allowable speed limit changes.

As in the case of DWL2-ST-VSL, the constraints on the spatial difference of speed limit values between two adjacent VSL zones on the motorway are not explicitly considered in the DWL4-ST-VSL setup. It is assumed that agents communicate information about congestion intensity and locations via remote policies. Thus, the difference in spatial speed limits should be reasonable in terms of optimal traffic flow control. This is also aided by DWL's ability to implement two sets of speed limits with different granularity simultaneously. Action set (6.10) is for agents A_1 and A_2 , which are closer to the bottleneck. The finer action set (6.11) is for upstream agents A_3 and A_4 . The more precise actions aim to slightly adjust the speeds of the arriving vehicles before they enter the VSL application areas controlled by downstream agents. In this way, agents smooth out the incoming traffic towards the congestion point, thus, avoiding the undesirable sudden deceleration of vehicles and effects such as shockwaves.

Reward Function

The TTS as an objective was successfully used in the previous setup, and is, therefore, also used in the setup of rewards for DWL4-ST-VSL. As in the previous case, each agent must learn to strike a balance between two conflicting policies. In the case of an inactive bottleneck, the penalty for a higher speed limit will be lower. On the other hand, if congestion occurs, the speed limit must be gradually lowered in the upstream sections to control the incoming traffic toward the bottleneck and, thus, maintain the operating capacity of the bottleneck at the highest possible level.

Modeling local and remote policies

According to the defined reward functions, each agent's policy tries to optimize its goal as follows.

Local policy for stable-flow control

The local policy LP_{i1} of an agent A_i aims to learn the speed limit to ensure a reduction of TTS by promoting, when possible, higher traveling speeds in stable-flow conditions. To achieve this goal, the LP_{i1} reward is:

$$r_{LP_{i1},t+1} = \begin{cases} 0, & \text{if } \min\{\bar{v}_{n,t+1} \mid n = i, i-1\} \geq 102 \\ -TTS_{n,t+1}, & n = i \text{ otherwise} \end{cases}, \quad (6.12)$$

thereby favoring average vehicle speeds above 102 [km/h].

In a certain percentage, LP_{i1} is activated in saturated flow during the transition from free-flow to congested flow and vice versa. Therefore, it prepares traffic for the second policy (LP_{i2}), which dominates in oversaturated (congested) conditions. After congestion has started to resolve by deploying LP_{i2} , and the congestion intensity reduces to a certain level, LP_{i1} helps restore traffic to free flow (higher traveling speeds) as soon as possible by gradually increasing the speed limit. Thus, LP_{i1} seeks to reduce traffic recovery time. Finally, the states perceived by LP_{i1} satisfy the minimum requirements to determine whether the flow in the agent's neighborhood is a stable flow or deviating from it. Consequently, the agent can recognize when the higher speed limits for free-flow can be implemented or not.

Local policy for congested traffic control

Local policy LP_{i2} aims to reduce TTS in the downstream motorway section in the case of an active bottleneck. Thus, an agent must learn and apply appropriate speed limits to restrict the inflow into the bottleneck until the discharge capacity is restored. If not, congestion will grow, and, consequently, it will increase its penalty in proportion to:

$$r_{LP_{i2},t+1} = -\beta TTS_{n,t+1}, \quad n = i - 1, \quad (6.13)$$

where coefficient β controls the agent's sensitivity to congestion. Instead of using only downstream congestion information, LP_{i2} uses information about the upcoming traffic flow (current speed and density) from the section $S_n, n = i$. This can be considered a prediction of the forthcoming traffic flow (how fast and with what volume it will arrive) into the downstream congested section $S_n, n = i - 1$. In this way, the description of traffic conditions (states) is extended to include more unique traffic characteristics for more efficient congestion control.

Remote policies

Same as DWL2-ST-VSL, in DWL4-ST-VSL, an agent A_i learns additional remote policies ($RP_{ij1}, \dots, RP_{ijr}$) that complement its neighboring agent's local policies. In order to know how A_i 's local actions a_t affect the neighbors' states, the agent updates the remote policies by the information it receives about its neighbors' current states and the rewards that neighbor agents have received (Figure 6.7). The performed experiments suggest that agents' communication is perfect (no loss of information and no breakdown of agents is assumed).

Winner action

In DWL4-ST-VSL, the action selection process is the same as in the case of DWL2-ST-VSL. Thus, an agent A_i 's experiences for each policy are respectively stored in Q_{ik} matrices. In the case of agents A_1 and A_4 ($k = 1, \dots, 4$), while for A_2 and A_3 ($k = 1, \dots, 6$).

6.3.3 Simulation setup

To evaluate whether the dynamic assignment of VSL zones and cooperation between agents with DWL have an advantage over static VSL zones with non-cooperative agents, DWL4-ST-VSL is compared with WL-VSL [71]. To verify the advantages of learning approaches over classical VSL control, DWL4-ST-VSL is also compared with previously described SPSC [109]. It is important to note that the calibration procedure of the simulated motorway section is not included because a synthetic model with different traffic loads was used for this analysis. The objective of this simulation experiment is to evaluate the impact of dynamically adjusting the VSL zone configurations and the different number of agents in DWL-ST-VSL on the optimization of traffic flow within an active bottleneck and the motorway as a whole.

Simulation model

The used simulation framework consists of the microscopic simulator SUMO (version 1.8.0) and the Python programming environment. Here, the software version is important because the simulation output in the new SUMO version may slightly differ from the simulation in the previous version, as the simulator source code is constantly being improved and updated.

The motorway model is based on the model used in [26]. It is divided into 5 main sections, $S_n, n = 0, 1, 2, 3, 4$. To ensure all combinations of VSL zones (see Figure 6.7) and to measure spatiotemporal characteristics of the traffic flow under the impact of dynamic VSL zones, in these experiments, the entire simulation model is divided into smaller links (each 50 meters long). The speed limit is simulated along with the computed configuration of the VSL zones for the chosen control time by directly assigning the allowed speeds to the corresponding links. The new speed limit and the configuration of the VSL zones are, thus, calculated by agents for each control time step $T_c = 150$ [s] that is chosen from multiple tests [26]. The used T_c value is in the range of the foremost values found by the sensitivity analysis of control cycle lengths performed in [56]. The bottleneck is generated on the motorway section S_0 . Each simulation lasts 1.5 hours, and all learning-based VSL approaches were trained in 14,000 simulations.

Traffic scenarios

To evaluate the DWL4-ST-VSL control solution's feasibility and behavior, and to determine whether agent cooperation and dynamic VSL zones allocation with DWL have advantages over VSL control approaches with static VSL zone configuration (WL-VSL and SPSC), the testing was done under medium and high traffic loads. The input traffic data used were synthetic data, and the calibration process of the simulated model is not within the scope of this analysis. Therefore, the driver behavior and vehicle characteristics were modeled using the *Krauss* car-following model with the default settings in SUMO [31].

Medium traffic load

In the downstream section S_0 (Figure 6.7), a bottleneck is induced by an increase in traffic demand at the on-ramp R_0 . The generated bottleneck is the primary test for DWL4-ST-VSL with dynamic VSL zone allocation. In this traffic scenario, the demand at on-ramp R_0 changes over time (see Figure 6.8). For the highest demand at on-ramp R_0 , 1,315 [veh/h], slower vehicles entering the motorway interact with the mainstream traffic in the merge area. Consequently, this causes disturbances, which trigger the activation of the bottleneck, and congestion appears. Traffic flow at ramps R_1 and R_2 remain constant for both traffic scenarios, with the flow of 385 and 230 [veh/h], respectively. The mainstream flow entering the bottleneck area has a constant rate of 1,385 [veh/h/lane]. The traffic flow consists of 94% cars, 3% buses and 3% trucks.

High traffic load

The induced congestion is much more significant in this traffic scenario than in the medium traffic scenario. In particular, an increase is generated by a 7.22% higher traffic mainstream demand entering the bottleneck area relative to the medium traffic scenario. This is the test for DWL4-ST-VSL emphasizing the dynamic adjustment of VSL zones. Since the congestion tail propagates much more upstream through the motorway, it can be expected that different VSL zone configurations will be used compared to the medium traffic scenario.

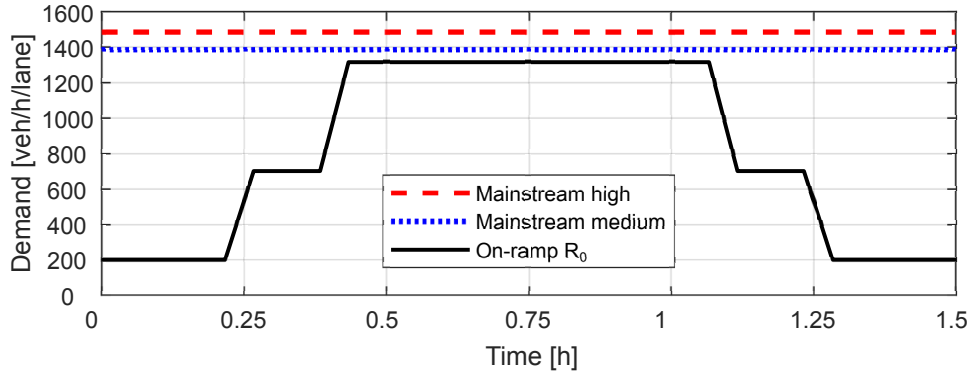


Figure 6.8: Tested traffic scenarios

Baselines of SPSC and WL-VSL

In the case of baselines, the best static VSL zone configuration $S_{2,(2L)} + S_{1,(1R)}$ (see Figure 6.6) and parameters were selected from several tests conducted for the medium traffic load. In the case of SPSC [109], the gain $K_v = 4.5$ and activation threshold of 23 [veh/km/lane] were selected from several tests.

The same best static VSL configuration is also used for the WL-VSL case. In WL-VSL, two local policies were used. Local policy LP_1 aims to maintain a higher speed on controlled motorway sections, while LP_2 aims to reduce congestion in the presence of an active bottleneck.

The observed state variables for LP_1 are densities within sections S_1 and $S_{2,(2L)}$, and for LP_2 densities within sections S_0 and S_1 , the *bottleneck region*. Each element of the action set contains two variables (the section S_1 and $S_{2,(2L)}$ speed limits). In this way, the winning policy will set speed limits for both sections [71]. The two rewards associated with the mentioned policies were modeled as follows:

$$r_{LP1,t+1} = \begin{cases} 0, & \text{if } \min\{\bar{v}_{n,t+1} \mid n = 0, 1, 2\} \geq 100 \\ -0.4(TTS_{1,(1R),t+1} + 2TTS_{2,(2L),t+1}) & \text{otherwise} \end{cases}, \quad (6.14)$$

$$r_{LP2,t+1} = -TTS_{0,t+1}. \quad (6.15)$$

DWL-ST-VSL parameters

For both DWL2 and DWL4-ST-VSL and WL-VSL the *learning Q (somewhat) before learning W* scheme [55] is used, controlled by $\alpha_W(1 - \alpha_Q)^\omega$ part in (6.1), where $\alpha_Q = \frac{1}{n(s,a)}$ and $\alpha_W = \frac{1}{n(s)}$ depend on the number of visits to $Q_i(s,a)$. Thus, the weight is larger when an agent is sure of what it does in a given state. This is indicated by a higher frequency of nominating a particular action based on the highest Q-value. The parameter $\omega = 1.5$ controls how fast W converges and was selected from multiple tests. The author of WL [55], in his demonstrated example, used $\omega = 3$. The parameter $\gamma = 0.8$ was chosen from [71]. The exploration probability is decreased by the parameter $\varepsilon = \exp \frac{-\ln(20)N}{6000}$, which diminishes with the number of simulation runs N [26]. In the DWL-ST-VSL nomination process (6.2), the cooperation between agents is controlled by remote policies (RP_i) via a cooperation coefficient C . The cooperation levels tested are $C \in \{0, 0.25, 0.5, 0.75, 1\}$.

DWL2-ST-VSL parameters

To keep the W-values of the local policies comparable to the W-values of the remote policies, the reward function (6.13) is scaled by the factor $\beta = 0.75$ in the case of agents A_2 and for agent A_1 $\beta = 1.25$. This is necessary because sections ($S_n, n = 1, 2$) are longer than S_0 , which affects the final comparison in choosing the winning action since the W-values are bounded by Q_{max} and Q_{min} . The bounds on the Q-values depend on the reward values r_{min} and r_{max} [55].

DWL4-ST-VSL parameters

Similarly, to keep the W-values of local policies comparable to the W-values of remote policies in the case of DWL4-ST-VSL, the reward function (6.13) is scaled by a factor $\beta = 0.75$ for the case of agents ($A_i, i = 2, 3, 4$) and for agent A_1 $\beta = 1.25$.

6.3.4 Simulation results

The VSL strategies are evaluated using the overall TTS measured on the entire simulated motorway segment (including ramps). Traffic parameters, average speed and density are measured in the bottleneck area (section S_0). The results presented in Figures 6.9–6.12 are from the exploitation phase. The specific response behavior of the allocation of dynamic VSL zones compared to the case of static zones was analyzed. The space-time congestion analysis is used to analyze the spatiotemporal behavior of dynamic VSL zone allocation and its impact on traffic flow control. To assess the benefits of cooperation between agents using DWL's remote policies, the impact of the cooperation coefficient on agent performance is also evaluated. As a measure of the learning rate of proposed agent-based learning VSL approaches, the convergence curves of overall motorway TTS during the training (learning) process are shown in Figure 6.13.

It is important to note that the purpose of this experiment is not to show the extent to which DWL4-ST-VSL can improve traffic but to investigate how the dynamic (spatiotemporal) adaptation of VSL zone configurations and the increased number of learning agents affect the traffic control optimization problem. Thus, an improvement over baseline should be considered primarily as a comparative measure between two different VSL approaches, the commonly used static VSL zones and the new paradigm with the dynamic adaptation of VSL zones, rather than as an absolute measure of performance.

Comparison of dynamic VSL zones allocation and static VSL zones

Note that the baselines use the best static VSL zone configuration found for a medium traffic load. Using a medium and a high traffic load in the experimental setup enables a simulation of significant differences in the spatial displacement of the congestion tail. In this way, the benefits and necessity of continuous spatiotemporal adaptive VSL control with dynamic VSL zones adaptation are demonstrated. Different VSL zone configurations per traffic scenario are learned without requiring manual setup, i.e., they are dynamically assigned using DWL4-ST-VSL to better respond to the different spatial distribution of traffic congestion caused by the bottleneck. At the same time, the experiment has highlighted the weaknesses of the static VSL zone configuration, which performs suboptimally under high traffic load. Therefore, VSL zones in VSL with the static VSL zone configuration need to be set up manually every time the traffic pattern changes, which is not practical and, therefore, justifies the foundation of a new concept of VSL control modeling based on dynamic VSL zone adaptation.

Medium traffic load

The performed simulations show that the best combination for establishing static VSL zones is $S_{2,(2L)} + S_{1,(1R)}$. In this case, VSL is able to control congestion in the case of medium traffic

load. In DWL2-ST-VSL, by additionally activating VSL zones within the S_2 section during the highest congestion peak (around $t = 1$ [h]), the agent A_2 helps its downstream neighbor A_1 , which contributes to an even more effective congestion resolution than the baselines (SPSC and WL-VSL) with static VSL zones. In DWL4-ST-VSL, the agents closest to the congestion (A_1 and A_2) are assisted by upstream agents (A_3 and A_4) that activate additional VSL zones within S_3 and S_4 just before the highest congestion peak ($t = 1$ [h]) (for a shorter period than DWL2-ST-VSL). In this way, agents A_3 and A_4 help their downstream neighbors.

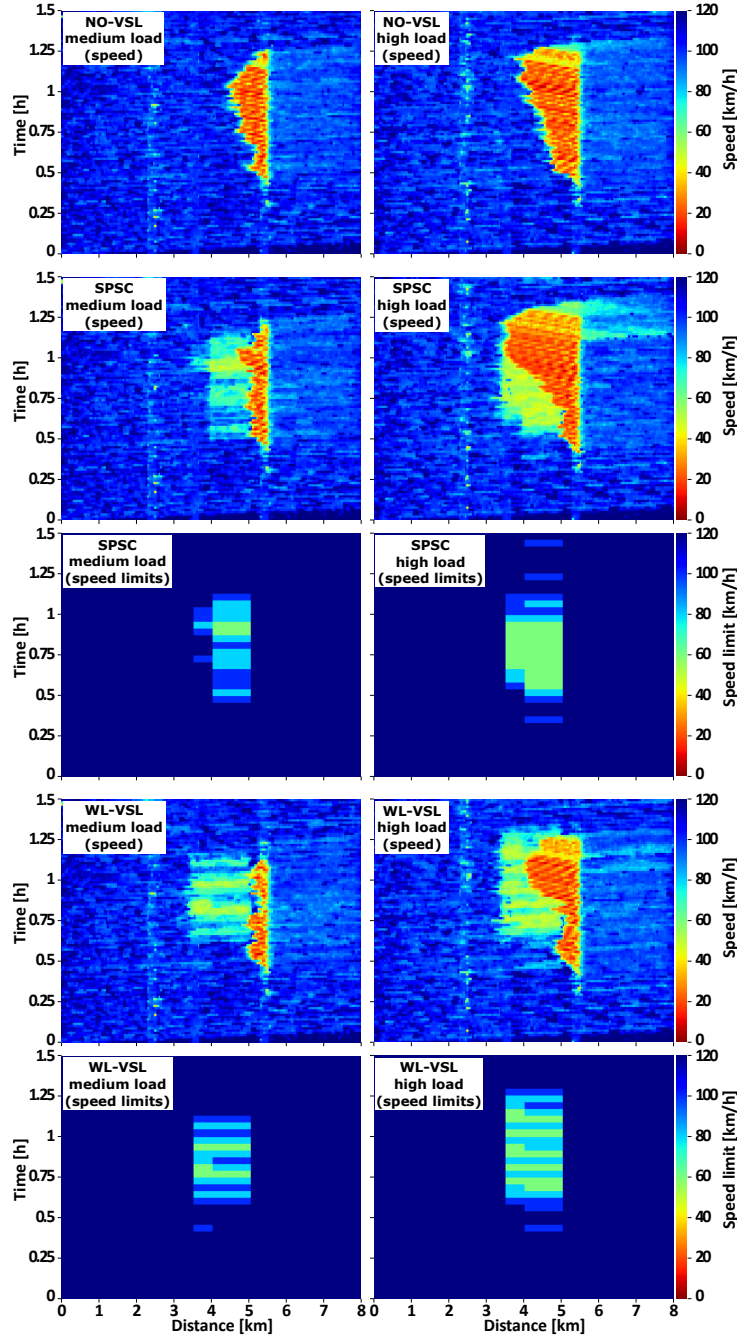


Figure 6.9: Space-time diagrams for simulated scenarios with static VSL zones

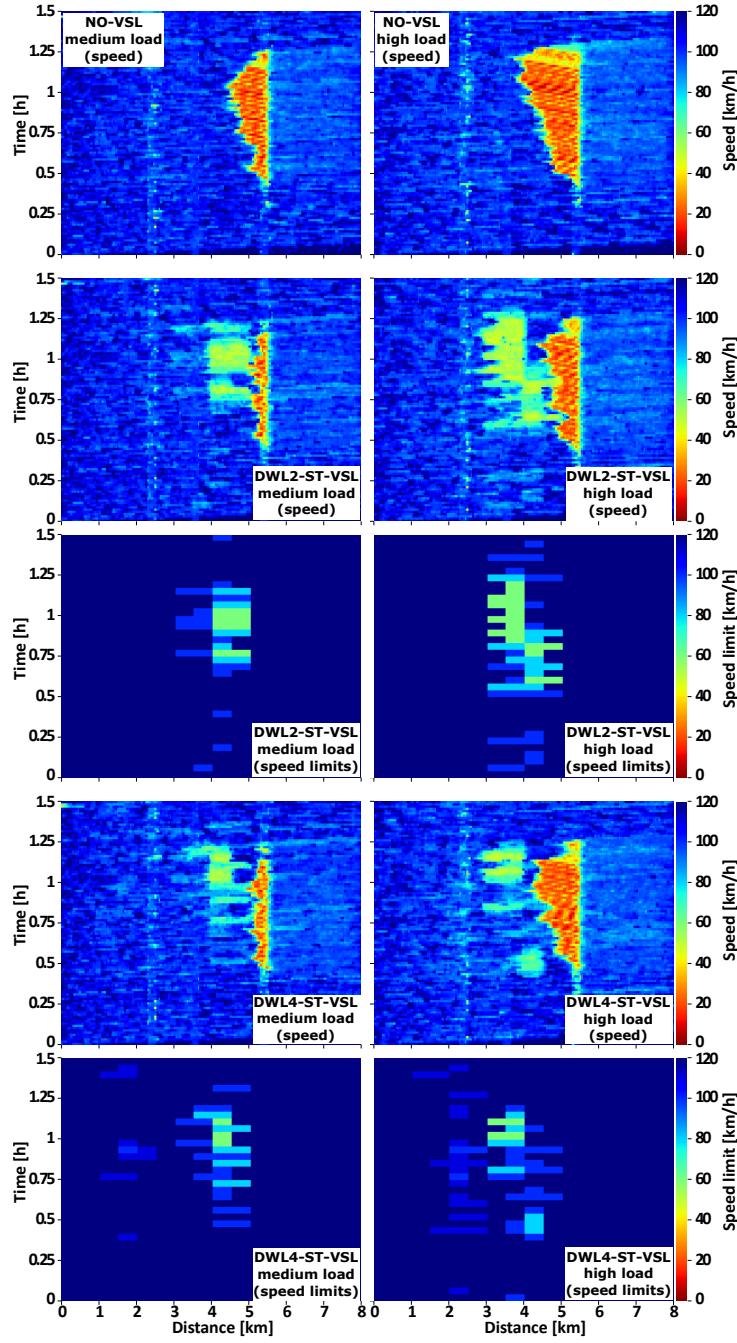


Figure 6.10: Space-time diagrams for simulated scenarios with multi-agent dynamic VSL zones

High traffic load

The performed simulations indicate that the static VSL zones perform suboptimally in a high traffic scenario. By applying different VSL zone configurations during the simulation within $S_n, n = 1, 2$ by DWL2-ST-VSL, and within sections $S_n, n = 1, 2, 3, 4$ in the case of DWL4-ST-VSL, they contribute more significantly to congestion clearing than baselines, which results from the gradual adjustment of the VSL application area. In the DWL2-ST-VSL case, agents started with stronger activation of the speed limits and VSL zones in section S_1 at the beginning

of the congestion. Over time, the congestion starts to propagate upstream through the motorway. The agents begin to use the VSL zones principally in sections S_1 and S_2 , while finally, for the highest congestion peak, the VSL zones are primarily activated in section S_2 .

In the case of DWL4-ST-VSL, VSL zones are activated mainly in all VSL sections at the onset of congestion (somewhat more sparsely for agent A_3 , while agent A_4 was almost not activated at all). Agents A_1 and A_2 preferred a shorter VSL zone configuration, while A_3 preferred a longer one. The application of shorter VSL zones in the downstream sections S_1 and S_2 could be due to the additional support provided by the upstream agents, particularly the speed limits applied by agent A_3 , which reduced the need for longer VSL zones and sudden decreases of the speed limit. As congestion increases, it can be seen in Figure 6.10 that the area of inactive VSL zones increases between upstream and downstream controlled sections, primarily due to the use of shorter VSL zones by agent A_3 and sparsely activated VSL zones by A_2 . After $t = 0.75$ [h], agent A_2 starts applying speed limits again in response to the sudden increase of the queue ahead of the bottleneck (faster propagation of the congestion upstream through the motorway). As congestion intensity approaches its peak, agent A_2 promotes a longer VSL zone, including lower speed limits. Agent A_1 is mostly inactive during this time period, thus, forming an additional valuable acceleration area (transition zone [29]) between the active VSL application area and the congestion tail. A somewhat unexpected behavior during the highest congestion peak is observed for agent A_4 , which did not apply speed limits below 120 [km/h] while A_3 was not active for 3 control steps (Figure 6.10). In the next section, some arguments that are believed to help explain this unexpected behavior of the agents will be given.

Nevertheless, both DWL2-ST-VSL and DWL4-ST-VSL adjusted the VSL zones to the spatially moving tail of the resulting congestion. This control strategy is more pronounced in the case of the high congestion scenario, in which agents attempt to create an additional artificial moving bottleneck to reduce the outflow from it and, thus, relieve the congested area. From Figure 6.10, it can be seen that the agents aim to create such a VSL configuration that ensures additional space (without speed limit) between the VSL zones and the congested tail. This can be viewed as an acceleration zone after the VSL zone, allowing vehicles to accelerate to v_{cr} (at which capacity is reached) before entering the congestion tail, as indicated in [28]. This feature of DWL-ST-VSL is very useful compared to the static VSL zone (fixed configuration) and confirms the findings that the higher the speed limit, the farther the VSL application zone should be from the bottleneck, which has been recently demonstrated in [29].

Space-time congestion analysis

Space-time diagrams are interesting for visualizing how traffic conditions evolve along the observed motorway segment. The on-ramp R_0 in S_0 is located at $x = 5.3$ [km]. DWL2-ST-VSL

ranges from $x = 3$ to $x = 5$ [km], while DWL4-ST-VSL ranges from $s = 1$ to $x = 5$ [km]. The best configuration of the static VSL zones (WL-VSL, SPSC) ranges from $x = 3.5$ to $x = 5$ [km]. The initial acceleration (transition) area [29] after the VSL zone starts at $x = 5$ [km] to the on-ramp R_0 and can be changed if the configuration of the VSL zones changes during agents' operations in DWL-ST-VSL.

Medium traffic load

In Figures 6.9 and 6.10, the mixed shades of red and orange correspond to congestion where vehicles are traveling at low speeds. The patterns of red stripes represent the propagation of the shock wave upstream through the motorway. Congestion begins at about $t = 0.4$ [h] in the bottleneck area and propagates upstream. After the demand on the on-ramp R_0 decreases, the congestion decreases and finally dissipates at $t = 1.25$ [h].

In both DWL-ST-VSL control strategies, the congestion (red) area is much smaller than in the baseline cases. The mixed shades of yellow-green-light blue in front of the congestion area correspond to the speed of vehicles obeying the speed limits (60 – 100 [km/h]) within active VSL zones. Such an artificially generated moving bottleneck (adaptive VSL application area) with a significantly higher average traveling speed than the one measured in the congestion area still reduces inflow into the congestion area, which helps to resolve congestion more efficiently than baselines. In response to spatially varying congestion, both DWL-ST-VSL produce more stable downstream flow than the best baselines with static VSL zones. In the medium load scenario, congestion propagates upstream from the bottleneck to location $x = 4.4$ [km]. In the case of DWL2-ST-VSL and DWL4-ST-VSL, the propagation is approximately reduced to $x = 5$ [km], which is an improvement of 66.7% compared to NO-VSL. Finally, the average density in the congested area (bottleneck S_0 and directly affected upstream section S_1) is reduced from 26.0 in NO-VSL to 20.0 [veh/km/lane] in the case of DWL2-ST-VSL, an improvement of 23.1%. The improvement for simulated DWL4-ST-VSL is 20.8%.

High traffic load

Again, for both DWL2-ST-VSL and DWL4-ST-VSL, the congestion area is smaller than for baselines. During the simulated scenario, different combinations of VSL zones were applied to respond to the changing congestion intensities and moving congestion tails. In this way, DWL2-ST-VSL and DWL4-ST-VSL are able to reduce the congestion area much more effectively than the baselines with static VSL zones. In the case of NO-VSL for the high-load scenario, the congestion spreads upstream from the bottleneck to the location $x = 3.8$ [km]. When DWL2-ST-VSL is applied, the propagation is reduced to near $x = 4.4$ [km], an improvement of 40%. Using the extended version with four agents (DWL4-ST-VSL), propagation is reduced to about

$x = 4.2$ [km], an improvement of 26.7%. Finally, the average traffic density in the congested area (S_0 and S_1) is reduced from the original 34.1 to 28.7 [veh/km/lane] by using DWL2-ST-VSL, an improvement of 15.8%. In the case of DWL4-ST-VSL, the improvement achieved is 14.7%. Just for comparison, in the case of WL-VSL with static zones, the congestion propagates near $x = 4$ [km], resulting in negligible improvement. Similar behavior is observed in the case of SPSC, eventually degrading the system performance.

Level of cooperation analysis

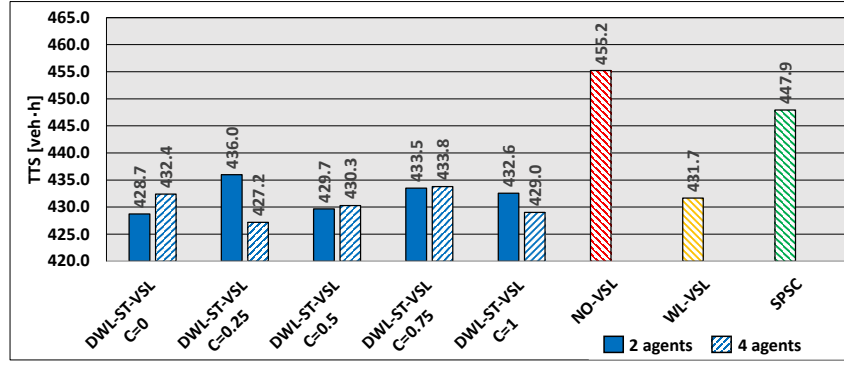
To evaluate the benefits of cooperation between agents using the DWL's concept of remote policies, the effects of the cooperation coefficient on agent performance are also assessed. The effects of different levels of agent cooperation on system performance are presented in Figures 6.11 and 6.12.

Medium traffic load

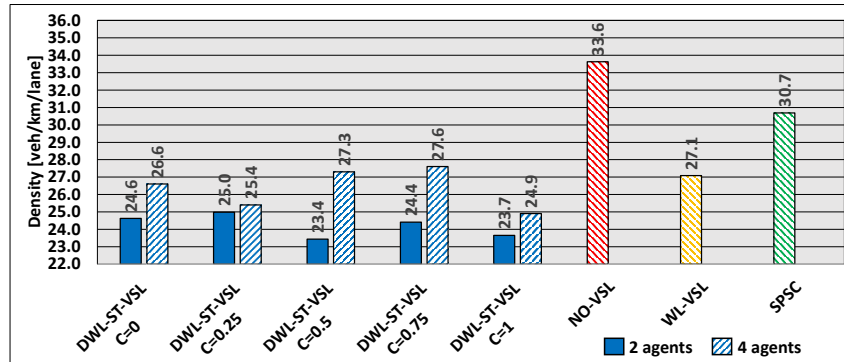
It can be seen that all DWL-based approaches outperform the baselines used in the experiment. The lowest TTS value is obtained with DWL4-ST-VSL and is 427.2 [veh · h] for $C = 0.25$. Compared to the NO-VSL case, ($TTS = 455.2$ [veh · h]), a reduction of 6.2% (Figure 6.11a). The best density is 23.4 [veh/km/lane] for $C = 0.5$ in the case of DWL2-ST-VSL, while it is 33.6 for the case of NO-VSL, an improvement of 30.4% (Figure 6.11b). In particular, the average vehicle speed for $C = 1$ in the case of DWL2-ST-VSL is 85.8 [km/h], while the speed in the case of NO-VSL is 73.3 [km/h], an improvement of 17.1% (Figure 6.11c).

High traffic load

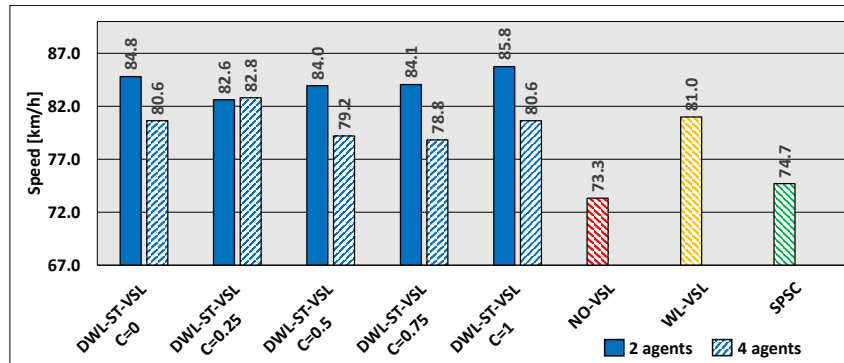
Similar results were obtained in the high traffic load experiment, where both DWL-ST-VSL configurations outperformed the baseline controllers. The lowest TTS value in the cooperative agent case in DWL4-ST-VSL is 501.0 [veh · h] for $C = 0.25$. Compared to the NO-VSL case, ($TTS = 524.8$ [veh · h]); this is an improvement of 4.5% (Figure 6.12a). The density is 34.0 [veh/km/lane] for DWL2-ST-VSL ($C = 1$), while in the case of NO-VSL it is 38.5, a reduction of 11.7%. The density is reduced by 7.8% by using DWL4-ST-VSL (Figure 6.12b). In particular, in case of DWL2-ST-VSL, the average vehicle speed for $C = 1$ is 73.8 [km/h], while in case of NO-VSL, the speed is 67.8 [km/h], an improvement of 8.8% (Figure 6.12c). In the case of DWL4-ST-VSL, the average speed is 10.9% higher (for $C = 1$).



(a) TTS in the overall network.

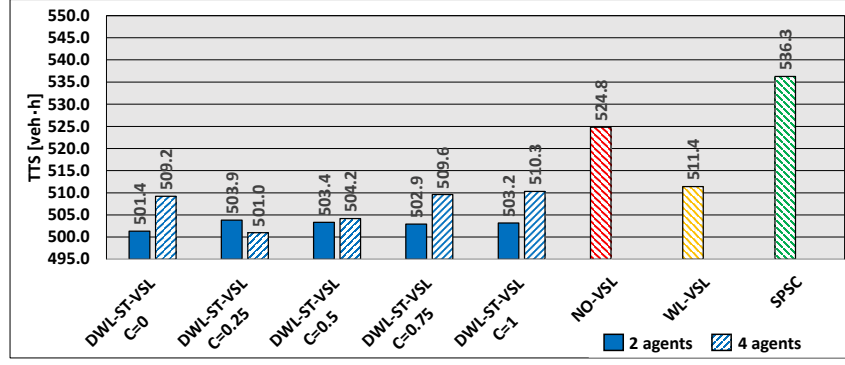


(b) Average traffic density in section S_0 .

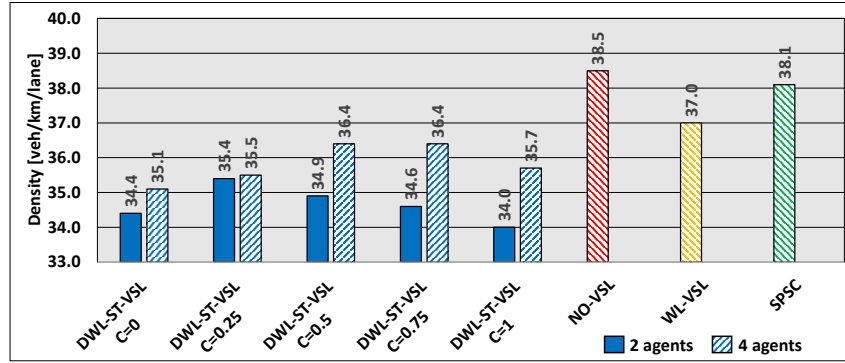


(c) Average vehicle speed in section S_0 .

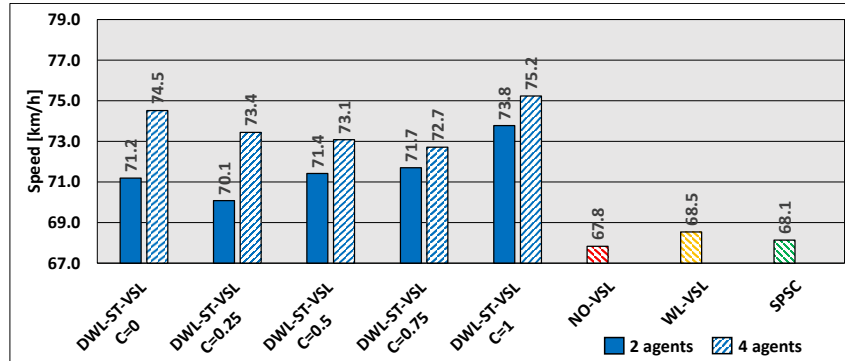
Figure 6.11: Traffic parameters for different levels of cooperation for the medium traffic load scenario



(a) TTS in the overall network.



(b) Average traffic density in section S_0 .



(c) Average vehicle speed in section S_0 .

Figure 6.12: Traffic parameters for different levels of cooperation for the high traffic load scenario

Convergence of TTS during the training process

A comparison of the convergence of TTS measured per training episode (episode \equiv one simulation) during the learning process is shown in Figure 6.13. The graphs are created using the moving average over 10 episodes, while TTS was measured in the entire motorway network (including all on- and off-ramps). At the beginning of the learning process, all agent-based VSL approaches performed inferiorly compared to NO-VSL, since agents explore the environment by executing random actions with high probability. As simulations progress, the number

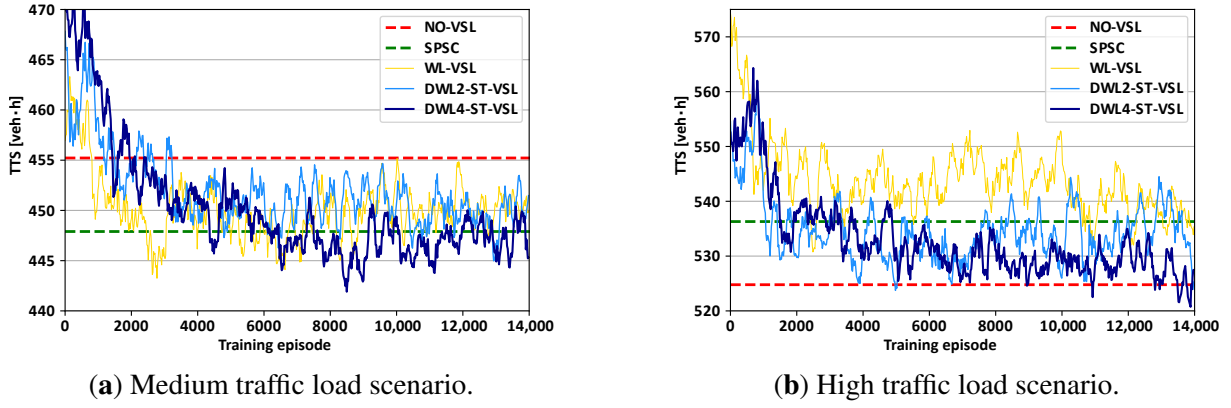


Figure 6.13: The convergence of TTS during the training process

of random actions taken reduces, and the exploitation of learned experiences increases. Consequently, TTS decreases, indicating progress in learning. Due to the different complexities of proposed MARL-based VSL controllers, the different decrease rates of TTS can be observed throughout the learning process. From Figure 6.13a, it can be seen that all approaches have stable decreasing learning curves; generally, DWL4-ST-VSL leads in TTS reduction over other strategies in the medium traffic scenario.

For the high traffic scenario (Figure 6.13b), the static VSL zones used in WL-VSL are prone to performing poorly compared to the dynamic VSL zones. Cases with dynamic VSL zone allocation via DWL2-ST-VSL and DWL4-ST-VSL need a higher number of training episodes to approach lower TTS values. As the learning process approaches 14,000 episodes, TTS in the case of DWL2-ST-VSL and DWL4-ST-VSL converges moderately towards and below the TTS value obtained in NO-VSL. Eventually, compared with the starting values, the overall TTS is gradually improved for all agent-based VSL strategies, favoring the learning rate of DWL4-ST-VSL in both traffic scenarios.

In the case of the high traffic scenario (Figure 6.13b), it can be seen that DWL4-ST-VSL needs a slightly longer time, i.e., higher number of training episodes (around 11,000) to reduce TTS below the value obtained by NO-VSL. Nevertheless, when converted in real time, it takes roughly 90 [h] of training in a simulator (on an Intel(R) Core(TM) i7-10750H CPU processor). Thus, if the traffic pattern changes significantly and, for example, in a real application the performance of DWL4-ST-VSL becomes poor, it can be retrained offline (on simulations) and deployed in a real application in a short time. Thus, DWL4-ST-VSL can be retrained offline to deal with traffic changes in the operating environment to ensure good performance in the newly observed traffic scenarios (similar to the continuous learning scheme for Q-Learning based VSL suggested in [20]).

The longer time needed to reach the favorable TTS level may be directly linked to the larger number of agents. They eventually need more training episodes to become aware of the inter-

ference they cause by their actions on their immediate neighbors and the controlled motorway system as a whole.

In the second half of the learning process, the oscillations in TTS are more pronounced. The possible contribution to this might be delaying W 's convergence until Q is well known (see section 6.3.3). Thus, W -values are more altered as Q -values are more learned. Consequently, this influences the policies' nomination (6.2) in the DWL process and eventually influences the cooperation strategies between agents. As a result, it might cause a change in a learned set of optimal policies, thus, resulting in different system responses during the second half of the learning process. The new policy can induce new rarely seen system states that have not been encountered before, thus, affecting agents' poor decisions. Nevertheless, the function approximation techniques can address this problem by ensuring better generalization (reasonable outputs) for rarely seen states, consequently, stabilizing the learning process.

6.3.5 Discussion

The outermost agents (A_3 and A_4) do not perceive congestion directly and, therefore, tend to exploit local stable traffic conditions by promoting higher speed limits and, in particular, favoring their local policy LP_{i1} . As a result, for small values of the cooperation coefficient C , they do not fully contribute to helping downstream agents to eliminate the congestion. This raises the question of whether C should be scaled differently depending on the spatial location of the agents rather than using uniformly distributed equal values for all agents. It might make sense to increase the coefficients of C the farther agents are from the location of the bottleneck so that they are more sensitive to the preferences of downstream agents and, therefore, give more priority to remote policies in the case of active congestion. The question then arises: to what extent?

The converse is also true since the actions of the downstream agents affect the state variables (in particular, the measured average vehicle speed) of the upstream agents. The upstream agents always observe the average speed in their immediate downstream area (in the case of local policy LP_{i1}) and, possibly, the actions performed by the downstream agents (lower speed limits) reduce the chance of winning the LP_{i1} . Therefore, a penalty by the measured TTS is more likely, even if the local environment is in free-flow conditions. This dependence is implicitly communicated to the downstream neighboring agent A_j in the form of a higher W -value for remote policy RP_{ji1} , which complements the local policy LP_{i1} of the upstream agents A_i .

The above observation shows the possible trade-off in choosing optimal values for C . A feasible solution to make the cooperation coefficient C adaptive is to use a scaling scheme as in (6.1) *learning Q (somewhat) before learning W* [55]. In this scheme, the updates of W -values are weighted differently. The weighting is higher when an agent is sure of what it does in a given

state. Given that the underlying DWL process (WL algorithm) is considered as a *fair* resolution of competition, this leads to the question: can the W-values of local policies, together with the probability of nominating a particular action in a given state, be communicated between neighbors and used as input to change the prior belief about good C , thus making C adaptive? Obviously, further research on the adaptive cooperation coefficient C is desirable.

Furthermore, the overlapping states of the environment, including the downstream neighborhood (see Figure 6.7), have positive and negative effects on the agent's learning behavior. The negative effect arises from the nonstationarity caused by the neighbors' actions, resulting in a moving learning target (particularly during the exploration phase in the training process) since agents are learning simultaneously. Thus, each time, A_i 's policy changes might cause other agents' policies to change, too [126]. The positive effect is the agent's ability to detect and respond to the early impulse of congestion in downstream traffic. All learning-based approaches were trained with the same number of simulations. However, due to nonstationarity, DWL4-ST-VSL may require more simulations to converge to better control policies for a given traffic scenario due to a higher number of agents. Therefore, DWL4-ST-VSL (and the final results) may be in a slightly unfavorable position compared to DWL2-ST-VSL.

In the performed experiments, it is assumed that all measurements (traffic data) are perfect. In reality, sensors are not ideal, and raw sensor data need to be analyzed and filtered before being used for traffic state estimation. Thus, accurate traffic states are important for real-time traffic control. Raw traffic flow data collected from sensors might be contaminated by different noises caused by the imperfection or damage of sensors. In [127], the authors introduced data denoising schemes to suppress the potential data outliers from raw traffic data for accurate traffic state estimation and prediction. This presents an open question for further research.

Additionally, the efficiency of DWL-ST-VSL is highly dependent on the learning process performed in traffic simulations. Since simulations themselves depend on the given initial parameters, not all possible relevant traffic conditions can be covered. A possible direction to improve the training process of DWL-ST-VSL by ensuring that all relevant traffic scenarios are covered is to use the idea of structured simulations. Originally proposed in [128], structured simulations are intended for testing the behavior of complex adaptive systems in general by changing the inputs into the simulations in a structured way. Such a framework might augment existing traffic scenarios (real or synthetic) with unprecedented scenarios that evoke or replicate important aspects of real traffic, such as rarely-seen traffic states in which VSL agents performed poorly. Thus, a structured simulations approach can enrich the training data set and consequently minimize unexpected behavior of the RL-VSL controller in practice.

Even under a medium load scenario, the resulting congestion on the motorway can be classified as a serious traffic problem. However, it has been shown that DWL2-ST-VSL and DWL4-

ST-VSL can effectively resolve the congestion in this scenario due to their added ability to dynamically adjust the VSL zone configurations. Since the DWL agents could not fully handle the congestion in the high load scenario (even when using four agents), it might be useful to extend the DWL4-ST-VSL control, e.g., by integrating it with the merge control (ramp metering) using the DWL multi-agent framework.

Experimental results confirmed the usefulness of using dynamic VSL zone allocation (the capability to adapt the VSL application area) while optimizing speed limits in traffic conditions with varying congestion. Similarly, in [56], a VSL strategy able to adjust each traffic controller cycle's length (duration) online, given the changes in traffic conditions, was shown to be superior compared to a fixed cycle length. Thus, integrating dynamic VSL zone allocation and dynamic control cycles can make VSL more adaptive, making VSL's performance more robust when operating in a nonstationary environment like a motorway. To accomplish the full benefits of adaptivity, the principal time constants of the system should be long enough for the system to ignore false disturbances and yet short enough to respond to indicative changes in the environment (the *stability-plasticity dilemma*) [129].

The VSL control approaches with static VSL zone configuration performed more poorly in high traffic scenarios than those with dynamic VSL zone allocation. Thus, results strongly indicate the need for an adaptive speed limit system in the speed limit, length, and position of VSL zones to efficiently cope with the unpredictable spatiotemporal varying congestion, which is more likely to be the case in real traffic scenarios on motorways.

6.4 Concluding remarks

This chapter presents DWL-ST-VSL, a MARL-based VSL control approach for the dynamic adaptation of VSL zones and speed limits. It has also been proven that an extended version with four cooperative agents is suitable to control four segments on a simulated motorway network in front of a congested area using the DWL4-ST-VSL algorithm. The simulations show that DWL4-ST-VSL and the DWL2-ST-VSL consistently perform better than the used baselines WL-VSL and SPSC. The results do not differ significantly between DWL2-ST-VSL and DWL4-ST-VSL in terms of bottleneck parameters. In terms of system travel time, DWL4-ST-VSL gives better results. VSL control is improved by simultaneously adjusting speed limit values and VSL zone configuration in response to spatiotemporal changes in congestion intensity and congestion's moving tails. In addition, the performance is improved by DWL's ability to implement multiple different policies simultaneously and to use two sets of actions with different speed limit granularity, as well as by the cooperation between agents via remote policies.

Accordingly, the results have confirmed the hypothesis defined in the introductory chapter re-

garding the expected benefits of using adaptive VSL zones along with speed limit adjustment and validated the ability of the learning, multi-agent controller to be used for VSL control on a wider motorway section. Therefore, in the next chapter DWL-ST-VSL concept will be integrated with the DT-GM to verify its performance in real-traffic scenarios.

Chapter 7

Testing VSL using the Geneva motorway digital twin

In this chapter, the functional integration of DT-GM and DWL-ST-VSL is described. DT-GM as a real-time virtual representation of physical motorway dynamics is used to test the previously proposed DWL-ST-VSL and evaluate its ability to set speed limits and adjust VSL zones in such a realistic environment. Thus, the main objective of this experiment is to provide insight into the use of DT-GM and its potential foundation in the development and evaluation of adaptive VSL controllers. Therefore, in the following part of the paper, it will be demonstrated how the DT technology can be used to create, update, evaluate, and deploy a digital twin-based VSL controller on a motorway.

7.1 Traffic and geometry characteristics of DT-DWL-ST-VSL experiment

Geometric characteristics

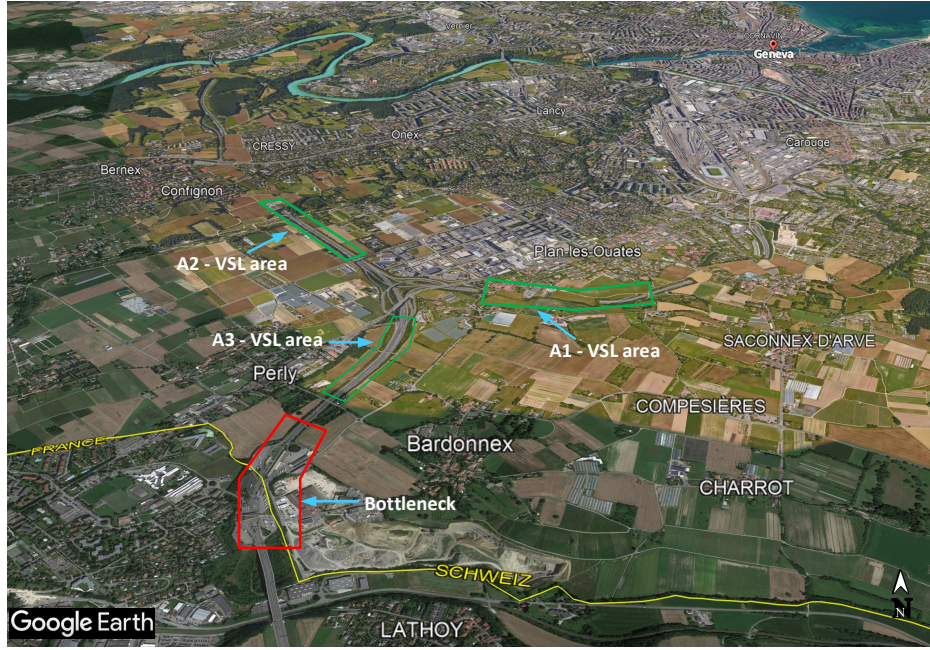
Taking into account the *Y* shape of the motorway geometry and mainly the similar traffic intensity from the direction of areas $A1$ and $A2$ (see Figure 7.1a), it can be assumed that the problem at hand possesses a symmetry property. Therefore, unlike the experiment in section 6.3.2, this experiment, based on symmetry property, uses a ring-cooperation structure (see Figure 7.2a) in modeling DWL-ST-VSL agents. For the sake of unambiguity, an agent A_i is abbreviated with a subscript letter, as usual, while the areas acted upon by the agents are designated without subscript letters to make the difference in the designation of agents or areas clear. The presented cooperative structure could be useful, in cases when, for example, agent A_3 is not able to solve the bottleneck problem alone. Thus, it is likely that the weaknesses of A_3 will be communicated

implicitly (via RPs) to the upstream agents A_1 and A_2 . Accordingly, the agents A_1 and A_2 can simultaneously help the agent A_3 . Moreover, by the ring-cooperation scheme, they can understand its mutual reactions even better and respond to spatiotemporal traffic changes efficiently by the ability to adapt (*activate*) a higher number of different VSL zones combinations.

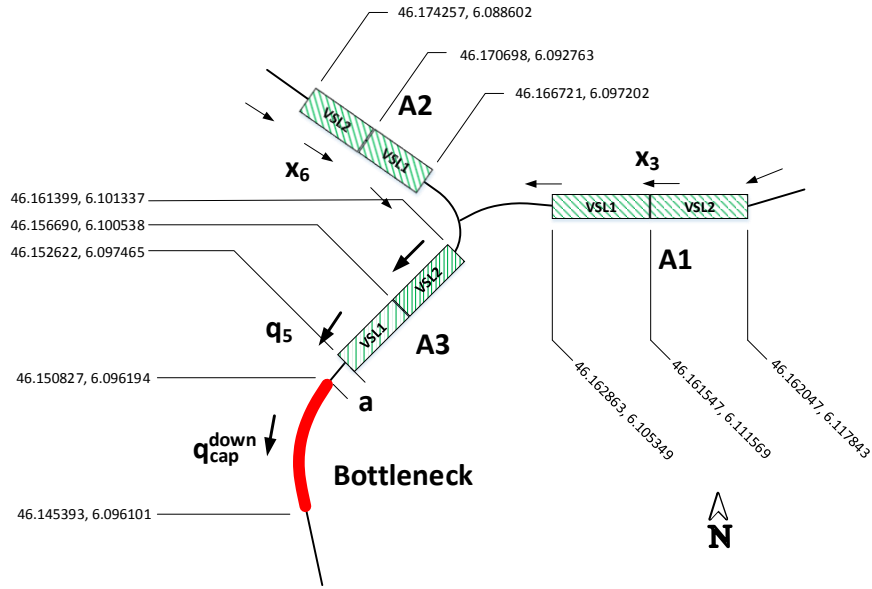
As noticed from Figure 7.1a, the VSL between areas A_1 , A_2 , and A_3 has not been applied due to the fact that this region mainly consists of the network topology with grade-separated intersections with noticeable curvature. Thus, lower fixed speed limits in this region are assumed. As can be seen, the scope of VSL is to optimize the bottleneck region (mainly counting for segments a and A_3) that are in proximity to the border crossing. In order to optimize the bottleneck, VSL application areas labeled as A_1 , A_2 , and A_3 , within which different VSL zones (VSL1 and VSL2) can be activated, are located upstream of the bottleneck, thus, controlling the flow rate going in the direction to south (Figures 7.1a and 7.1b).

Traffic characteristics

As pointed out in chapter 3, the most prominent occurrence of congestion is present during the morning and the afternoon rush hours (weekday traffic characteristic). For the evaluation process of DWL-ST-VSL and its performance analysis in this experiment, the focus is on the southern region between Switzerland and the French border. Despite the fact that the passage is free (no stopping), vehicles need to slow down when passing this area. In peak periods, this presents a severe capacity drop. Consequently, q_{cap}^{down} becomes lower than the peak hour traffic volume q_5 arriving in the bottleneck area. Therefore, the main objective of VSL is to optimize the traffic flow during the afternoon period going in the direction of the Swiss-France border.



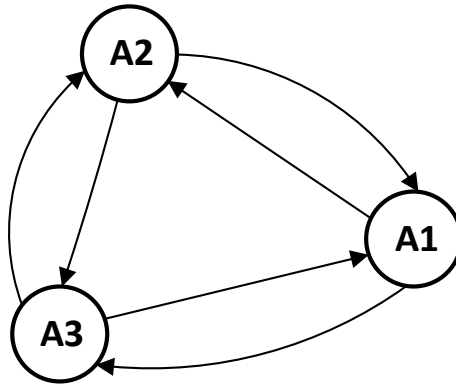
(a) Illustration of study motorway area



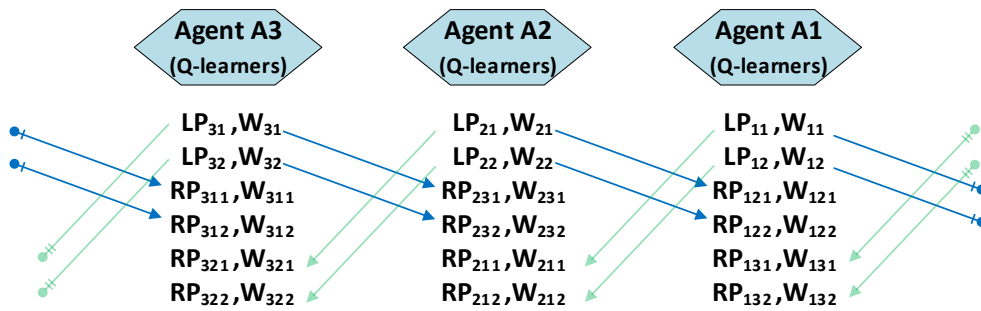
(b) Abstraction of the motorway model and DWL-ST-VSL agents locations

Figure 7.1: Real motorway and its abstract representation

Figure 7.2b provides a more detailed view (i.e., one level deeper) of the agents' cooperation scheme. It, thus, provides the cooperation aspect of agents and the relations between LPs and RPs. This part is particularly on the scale of technicality for those who would like to use the model in the future and, thus, achieve a better understanding of the VSL controller's program code.



(a) A ring-cooperation scheme model for DWL-ST-VSL agents



(b) A closer look on LPs and RPs information exchange

Figure 7.2: Agents cooperation scheme

7.2 Modeling DT-GM as DWL-ST-VSL control problem

Here, necessary definitions of agent environments are introduced, e.g., what part of the network is observed by them and what actions and signals (rewards) are used to optimize traffic in DT-GM. To distinguish between experiments and setups already performed, the integration of DT-GM with DWL-ST-VSL is abbreviated as DT-DWL-ST-VSL.

7.2.1 State definition

To make the agents' state space definition more explainable four regions of interest are defined. Following labels in Figure 7.1b the bottleneck occurs at the border crossing area, as previously described. Thus, by applying VSL control in shaded areas A1, A2, and A3 additional storage is created that will reduce the flow rate q_5 and release the bottleneck. Particularly, the classical approach here is not suitable since congestion becomes too strong and propagates much more upstream through the network (see Figure 3.9). Thus, region A3 (VSL application areas: VSL1

and VSL2) is not sufficient. Therefore, two additional VSL agents A_1 and A_2 are added to help agent A_3 in bottleneck resolving.

Similar to agents from the previous chapter, in this experiment A_1 's local policy LP_{11} at time t senses state $s = (a_{1,t-1}, V_{a+A3}, \rho_{A1}, \rho_{a+A3})$, while LP_{12} $s = (a_{1,t-1}, V_{A1}, \rho_{A1}, \rho_{a+A3})$. Here $a_{1,t-1}$ is the previous action executed by an agent. V_{a+A3} is the normalized speed measured in areas a and $A3$. Similarly, ρ_{A1} is density in area $A1$ while ρ_{a+A3} is density in areas $a + A3$. Unlike in LP_{11} , the normalized speed is measured in area $A1$ and labeled as V_{A1} . A_2 's LP_{21} senses state $s = (a_{2,t-1}, V_{a+A3}, \rho_{A2}, \rho_{a+A3})$, while LP_{22} senses state $s = (a_{2,t-1}, V_{A2}, \rho_{A2}, \rho_{a+A3})$. In case of A_3 , LP_{31} senses state $s = (a_{3,t-1}, V_a, \rho_{A3}, \rho_a)$, while LP_{32} senses $s = (a_{3,t-1}, V_{A3}, \rho_{A3}, \rho_a)$. Due to agent A_3 proximity to the bottleneck, the small area a is chosen as an early indicator for the formation of congestion that signals the agent that speed limits and/or VSL zones configuration might require recomputation. This warning and A_3 activation can also trigger the A_1 and A_2 activation via RPs. Therefore, the general conceptual state definition slightly differs from the concept for example used in the sections 6.3.1 and 6.3.2. The state parameters measured in the downstream section $a+A3$ are spatially dislocated to agents A_1 and A_2 .

Traffic flow speed V_n is described as the normalized mean speed value due to the fact that each area may contain multiple edges (motorway sections) and some of them may have different nominal speed limits. Thus, if the flow speed on each section is maintained close to the nominal speed limits, V_n will be closer to the value of one while during the congested period or under the VSL application with imposed lower speed limits, V_n would have a value lower than one. Assuming m edges the equation for computing the normalized average speed is as follows:

$$V_n = \frac{\frac{v_1}{v_{1,nominal}} + \dots + \frac{v_m}{v_{m,nominal}}}{m}. \quad (7.1)$$

Speeds are encoded in the variable V_n , which corresponds to the measured average vehicle speed $\bar{v}_{n,t}$ at time t in motorway section S_n ($n = A1, A2, A3, a + A3$), as shown in Figure 7.1b. Similarly to bounds used in section 6.3.2, in this experiment, each normalized speed measurement can fall into one of four normalized intervals defined with boundary points $v_{1,norm} = \frac{50.4}{v_{free}}$, $v_{2,norm} = \frac{75.6}{v_{free}}$, $v_{3,norm} = \frac{97.2}{v_{free}}$, where $v_{free} = 120$ [km/h], while speeds 50.4, 75.6 and 97.2 [km/h] are used as bounds. In this way, the lowest speed interval is associated with strong congestion. As the speed increases, the closer the traffic is to free flow conditions that are characterized by traffic flow speed above 97.2 [km/h]. In this way, the state space variable V_n across the areas $A1, A2, A3$, and a is uniquely described.

Similarly, due to different lane numbers of particular sections for traffic density ρ , a lane-weighted number of vehicles expressed as an equivalent number of vehicles per kilometer per

lane computed by the formula below is used:

$$\rho = \frac{1000}{\sum e_{i,length}} \sum \frac{e_{i,N}}{e_{i,lanes}}, \quad (7.2)$$

where $e_{i,length}$ stands for length of edge e_i , $e_{i,lanes}$ is the number of lanes on particular edge, while $e_{i,N}$ represents the number of vehicles on a particular edge. All edges' IDs can be found in the model on the GitHub repository mentioned in chapter 1.

Current traffic density $\bar{\rho}_{n,t}$ measured in the motorway section n , is stored in the variable ρ_n . Each density measurement can fall into one of twelve intervals defined by the boundary points 15, 20, 23, 26, 29, 32, 35, 37, 45, 55, 65 [veh/km/lane].

7.2.2 Action space

Again, the same rules defined and explained in chapter 6 apply to this experiment. Thus, each element in the action sets (7.3) and (7.4) consists of two variables. The upper numeric value represents the speed limit [km/h], while the lower represents an active VSL zone in which the speed limit should be deployed according to the given index (see Figure 7.1b). Agent A_1 controls the speed limit and the length of the VSL zone in its VSL area A1, while A_2 controls area A2, and A_3 in area A3. The denoted variables $NOM12$ and $NOM3$ encode the nominal speed limit on corresponding VSL application areas. Accordingly, $NOM12$ defines nominal speed limits of 120 [km/h] across A1 and A2 areas while $NOM3$ defines speed limit in area A3 as follows. In VSL1, the nominal speed limit is 100 [km/h] due to the proximity of the border and, thus, vehicles are obliged to gradually decrease the traveling speed, while VSL2's nominal speed limit is 120 [km/h].

$$\mathcal{A}_{12} = \left\{ \left\{ \begin{matrix} 80 \\ VSL1 \end{matrix} \right\}, \left\{ \begin{matrix} 80 \\ VSL2 \end{matrix} \right\}, \left\{ \begin{matrix} 80 \\ VSL12 \end{matrix} \right\}, \left\{ \begin{matrix} 100 \\ VSL1 \end{matrix} \right\}, \left\{ \begin{matrix} 100 \\ VSL2 \end{matrix} \right\}, \left\{ \begin{matrix} 100 \\ VSL12 \end{matrix} \right\}, \left\{ \begin{matrix} NOM12 \\ VSL12 \end{matrix} \right\} \right\} \quad (7.3)$$

$$\mathcal{A}_3 = \left\{ \left\{ \begin{matrix} 60 \\ VSL1 \end{matrix} \right\}, \left\{ \begin{matrix} 60 \\ VSL2 \end{matrix} \right\}, \left\{ \begin{matrix} 60 \\ VSL12 \end{matrix} \right\}, \left\{ \begin{matrix} 80 \\ VSL1 \end{matrix} \right\}, \left\{ \begin{matrix} 80 \\ VSL2 \end{matrix} \right\}, \left\{ \begin{matrix} 80 \\ VSL12 \end{matrix} \right\}, \left\{ \begin{matrix} 100 \\ VSL2 \end{matrix} \right\}, \left\{ \begin{matrix} NOM3 \\ VSL12 \end{matrix} \right\} \right\} \quad (7.4)$$

It can be noticed that the value sets used by agents A_1 and A_2 differ from the action set of agent A_3 . The agent A_3 may reduce the speed limit to lower values (minimum speed limit is 60 [km/h]). This is done under the assumption of agent A_3 proximity to the bottleneck that allows for stronger flow reduction in case of bottleneck activation.

7.2.3 Reward function

The same concept from section 6.3.2 is used in the LPs modeling and, therefore, in the modeling of reward functions. Thus, some notations are not explained twice to keep the focus on the newly introduced ones. Similarly to the state definition of this experiment, the concept of normalized speeds introduced above as well as in the definition of reward functions is used so as to ensure a better agent interpretation of state and rewards signals correlations. Therefore, the reward function is defined as explained in continuation.

Agent A_1 and A_2 reward function

For LP_{i1} agent's policy, the next reward function is defined:

$$r_{A1,LPi1,t+1} = \begin{cases} 0, & \text{if } \min\{\bar{v}_{e,t+1} \mid e \in \bar{A}_{1,a}\} \geq \frac{114}{v_{free}} = 0.95 \\ -\alpha_{A12} \sum_{e \in A1,A3,a} TTS(e, t+1) & \text{otherwise} \end{cases}, \quad (7.5)$$

where $\bar{A}_{1,a}$ defines all motorway segments, including the starting point of area $A1$ and including area a , as well as all segments between them. In this way, notation in summation is maintained clear for reading. In the case of $e \in A1, A3, a$, the term e represents motorway segments in areas $A1, A3, a$. Particularly the term edge e is used since SUMO internal road network modeling contains edges and nodes that connect them to the structure of the road network and they follow mainly the topology attributes of the real roads. For example, $A3$ area is represented by multiple connected edges where not all edges have the same number of traffic lanes. Nonetheless, this is a rather technical detail yet it is defined to make the controllers' program code logic and developed DT-GM model understandable for further users. The number 0.95 represents the 95 % of free-flow speed as a threshold for maximal reward.

For LP_{i2} policy the reward is defined as follows:

$$r_{A1,LPi2,t+1} = -\beta_{A12} \sum_{e \in A3,a} TTS(e, t+1). \quad (7.6)$$

In both cases, α_{A12} and β_{A12} are scaling factors that define how sensitive the agent is to congestion. Particularly, terms $A12$ means that both agents A_1 , and A_2 are using the same parameters due to the symmetry property of DT-GM in the observed region.

A similar methodology for agent $A2$ is used to set up its rewarding function.

Agent A_3 reward function

Similarly, for agent A_3 the reward function in case of $LPi1$ is as follows:

$$r_{A3,LPi1,t+1} = \begin{cases} 0, & \text{if } \min\{\bar{v}_{e,t+1} \mid e \in \bar{A}_{3,a}\} \geq \frac{114}{v_{free}} = 0.95 \\ -\alpha_{A3} \sum_{e \in A_{3,a}} TTS(e, t+1) & \text{otherwise} \end{cases}, \quad (7.7)$$

while LP_2 's policy reward is defined as follows:

$$r_{A3,LPi2,t+1} = -\beta_{A3} \sum_{e \in a} TTS(e, t+1). \quad (7.8)$$

7.3 Functional integration of DT-GM and DWL-ST-VSL

As already pointed out, two parts of the framework must be integrated and run in parallel. Technically, the main part of integrating DT-GM and DWL-ST-VSL is to provide connectivity and parallel execution of both programs, as well as synchronization at run-time. The explanation of this functionality can be found in Algorithm 7.1. First, all parameters of the two algorithms must be initialized. Then, DT-GM starts processing the traffic data streams received from the motorway traffic counters, i.e., DT-GM is calibrated online (see Algorithm 3.1). For every $t = 150$ [s], the program part of DWL-ST-VSL is executed (for details see Algorithm 6.1). In this part of the code, the VSL agents are trained in the initial phase, while in the final phase, their acquired knowledge is used for traffic control in the running DT-GM replica of the physical motorway system. The training process is generally repeated until the expected satisfactory control objective level is reached. However, this process is generally limited by execution time, as agents perform training on simulations that run in real time as traffic evolves on a motorway. This can potentially be overcome by introducing the concept of parallelism in simulations e.g., the DTI concept introduced in chapter 3, which can accelerate the training process by having DWL-ST-VSL train on multiple preferably randomized simulation partitions at each stage, and, thus, allowing for training the DWL-ST-VSL in real time on actual traffic data using the DT-GM.

Algorithm 7.1 DT-DWL-ST-VSL at each simulation step

```

// Set DWL-ST-VSL parameters and run DT-GM model
Init  $\vec{S}, \vec{C}, \vec{P}, \vec{q}, \vec{v}, \vec{v}_{type}, \vec{X}, \vec{X}_{free}, \vec{X}_{free-des}$ 
for each simulation step
  if simulation time % 60 [s] == 0 then
    Get new actual traffic data via ODPMS:  $\vec{q}, \vec{v}, \vec{v}_{type}$ 
    // DFC computations
    For given  $\vec{X}_{free-des}$  and  $\vec{q}$  calculate  $\vec{X}_{free}$  and  $\vec{X}$  (see (3.2))
    Update calibrators  $\vec{C}$  and routes distributions  $\vec{P}$  using  $\vec{X}$ 
    // DWL-ST-VSL traffic control during DT-GM run-time
    if simulation time % 150 [s] == 0 then
      // Update  $Q$ 's and  $W$ 's and execute actions
      for each  $A_i$  in  $A$ 
        for each  $LP_{il}$  in  $LP_i$ 
          Get  $LP_{il}$ 's traffic states,  $s_{il,t}$ 
          Get rewards from  $A_i$ 's environment
          Update  $Q_{LP_{il}}(s_{il,t-1}, a_{il,t-1})$  for  $LP_{il}$ 
          Update  $W_{LP_{il}}(s_{il,t-1})$ 
          Nominate action  $a_{il,t}$  with max  $Q_{LP_{il}}$  for  $s_{il,t}$ 
          Get  $W_{LP_{il}}(s_{il,t})$ 
        end for
        // Get nomination by remote policies
        for each  $RP_{ijr}$  in  $RP_i$ 
          Get  $RP_{ijr}$ 's traffic states,  $s_{ijr,t}$ 
          Get rewards for  $s_{ijr}$  from  $A_j$ 's environment
          Update  $Q_{RP_{ijr}}(s_{ijr,t-1}, a_{ijr,t-1})$  for  $RP_{ijr}$ 
          Update  $W_{RP_{ijr}}(s_{ijr,t-1})$ 
          Nominate action  $a_{ijr,t}$  with max  $Q_{RP_{ijr}}$  for  $s_{ijr,t}$ 
          Get  $W_{RP_{ijr}}(s_{ijr,t})$ 
        end for
        // Select and perform action
        Compute and execute winning action  $a_k$  (using formula (6.2))
      end for
    end if
  end if
  Calibrate and update simulation scenario
end for
Save DT-GM simulation state, DWL-ST-VSL's parameters and close simulation

```

7.4 Experimental setting

This sections mainly distinguish between parameters set up among agents and the baselines' configurations used in comparative analysis. Additionally, it presents the traffic flow characteristic that is used in the testing phase comprising real-time actual traffic from the motorway in the region of Geneva that is being used for testing of the proposed DT-DWL-ST-VSL controller.

Since the DWL-ST-VSL performances are verified in detail earlier in this thesis on the synthetic motorway model, a comparison with all baselines presented in chapter 6 is not performed again. Instead, the goal is to assess whether the dynamic adaptation of VSL zones and cooperation between agents have an advantage over static VSL zones and non-cooperative agents and to which extent in the context of the DT microsimulation environment. For that, the proposed system is tested in realistic motorway traffic conditions that are reflected very accurately by the proposed DT-GM. Thus, for comparison and analysis purposes, the following baselines are used: No-VSL case, cooperating agents with different cooperation levels, non-cooperating agents (cases with a level of cooperation $C = 0$ that may be seen as WL-VSL introduced earlier), and VSL with fixed zones. Therefore, the training and evaluation processes are performed on the real traffic scenario analyzed in chapter 3.

DT-GM parameters

In the training phase, all DT-GM parameters are kept as described in the introduction part of DT-GM. However, an additional noise in simulations is introduced by random simulation seeds.

DT-DWL-ST-VSL parameters

With regard to different time scope analyses (24-hour long simulation scenario) and different network topologies, some parameters were adjusted following mostly the same logic that is explained in chapter 6. The ones that are not mentioned remain the same as in chapter 6. Since the simulation (one episode) in this analysis lasts for 24 [h], the exploration probability is re-computed using the modified exponential decrease as follows: $\varepsilon = \exp \frac{-\ln(20)N}{50}$. This approach decreases the exploration probability with the number of simulation runs N . However, one simulation takes approximately 1.5 [h] in a simulator. Since the longer duration of one simulation run in testing the DT-DWL-ST-VSL nomination process, the cooperation levels tested are $C \in \{0, 0.5, 1\}$.

Comparability of LPs and RPs is ensured by properly scaling the rewards functions. In the case of agents A_1 and A_2 , the reward function (7.5) is scaled by the factor $\alpha_{A12} = 0.6$ and reward function (7.6) by the factor $\beta_{A12} = 0.245$. Similarly, for agent A_3 the scaling factors are $\alpha_{A3} = 1$ and $\beta_{A3} = 1$, regarding the reward functions (7.7) and (7.8) respectively. The same procedure is used as before, i.e., scaling down rewards reasonably to make policies in DWL comparable. Particularly scaling is based on the section lengths with which agents' reward functions are associated. One has to bear in mind that in this experiment, the mainstream model is not uniform in the number of lanes. Thus, an absolute measure of lanes-length of the section is used instead of just taking section length to scale properly the TTS part of rewards. Thus, the scaling factors mainly maintain the proportions as in DWL-ST-VSL in section 6.3.2.

7.5 Results and analysis

The analysis is divided into two parts. The first part refers to the training phase, where the results are presented and analyzed in terms of the controller's learning behavior from the point of view of nonstationarity introduced by the DWL and DT simulation environment.

The second part focuses more on the possibility of using the new DT-GM microscopic simulation paradigm in the development and evaluation of VSL controllers by evaluating the proposed adaptive VSL controller using different DWL settings. In general, it provides data-based evidence for the use of DT-GM in the training and evaluation process of the DWL-ST-VSL controller and how such an environment, which is not necessarily predictable but reflects real motorway traffic well, affects the controller's policies. This is evaluated against the performance and behavior of the adaptive VSL controller by analyzing the dynamic adaptation of VSL zones and speed limits in the context of optimizing the examined bottleneck and motorway network performance.

7.5.1 Learning performances of DWL-ST-VSL controller

To evaluate the learning performance of DT-DWL-ST-VSL, *TTS* (particularly *TTS* convergence rate) is analyzed as the main objective of the VSL controllers. To get a complete picture of VSL performance, it is important to note that this analysis represents the *TTS* of the system measured in the whole simulated motorway network. Each measurement in Figure 7.3 corresponds to *TTS* accumulated during 24-hour-long traffic simulation (episode). In this way, the global impact of the agents on traffic performance is captured.

TTS convergence

Before commenting on *TTS* results, the sources of randomness that are present in the performed simulation experiments are also worth mentioning. The first one is related to MARL nonstationarity, i.e., the interference that agents cause when executing actions. This is particularly pronounced in a decentralized MARL, especially in the case of the underlying WL algorithm (by the nature of policy selection) that is at the core of DWL. In addition, the randomness is produced by RL itself due to the exploration/exploitation mechanism. It is also used in the initialization of objective function coefficients in Simplex while solving flow balance equations in the DT-GM model (see section 3.4.1). Moreover, it has been decided to introduce randomness in simulations by the author. Particularly, the use of random seeds in each episode (simulation) is allowed, which makes things closer to reality rather than using fixed simulation seed as was the case with the deterministic environment in the analysis of DWL-ST-VSL in chapter 6. Under that assumption, the weekday traffic scenarios (randomized by different seeds) based on

the weekday traffic presented in chapter 3 are used as a basis for the training and evaluation process. Accordingly, better generality and robustness of the controller can be expected in the final application in controlling stochastic processes, i.e., while applied to control new weekday traffic scenarios that were not seen by the controller before.

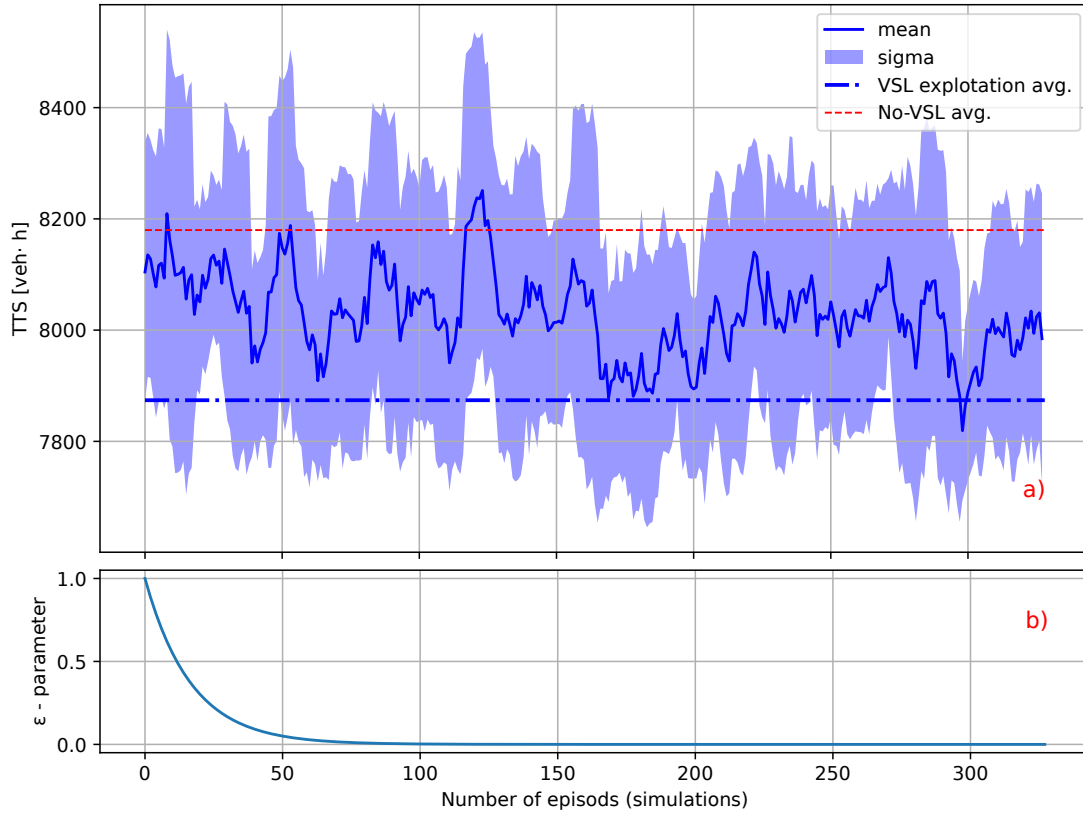


Figure 7.3: Overall network's TTS behavior during the training phase in fully randomized DT-GM simulations

Accordingly, the above-mentioned *randomness* may be seen as the reason for the less notable convergence of TTS in Figure 7.3 compared to the four and two agents DWL-ST-VSL design in chapter 6. However, to bound the experiment, one can, for example, select a few seeds and run each of them per episode and, thus, use their average result per episode for which similar convergence performances are expected to be achieved. Since simulational proof of DWL-ST-VSL convergence has been given earlier in this thesis, this chapter is oriented to generate an experiment that closely reflects real traffic and its inherent stochastic characteristics using DT technology. As such, the goal is to emphasize the ability to use DT in the training process of adaptive control systems that mimic an environment that closely reflects real traffic nature rather than using a simulation environment with deterministic properties. Such simulations, thus, provide an adequate test for dynamic VSL zones adaptation controlled by DWL-ST-VSL. Nonetheless, it is obvious from Figure 7.3a that DT-DWL-ST-VSL extended in the ring-oriented cooperation structure learns and, thus, improves (reduces) the objectives over time as episodes run (337

episodes in total) despite the mentioned nonstationarity. In order to have a reference point for comparison of TTS convergence rate, in Figure 7.3a, two additional measurements (averages obtained without VSL control and one obtained under VSL control) are shown together with the TTS curve. As can be seen, the DT-DWL-ST-VSL outperforms the case without VSL control (No-VSL). $TTS = 8,179.86 [veh \cdot h]$ for No-VSL is obtained as an average of several simulation runs using random seeds. The same is done for DT-DWL-ST-VSL (with $C = 0.5$) that gives an average $TTS = 7,873.97 [veh \cdot h]$, while the TTS curve and associated standard deviation (the shaded region around TTS curve) are computed using the moving average over 10 episodes. According to the presented results, DT-DWL-ST-VSL, after being trained, is able to (on average) reduce TTS for 3.74 % per simulated daily traffic scenario with a noticeable reduction in variation of TTS .

It is worth mentioning that the ϵ -parameter used for the exploration/exploitation phase is illustrated in Figure 7.3b. Accordingly, a strong exploration phase is at the beginning of the training process. With the increase in simulation runs, it exponentially decays. After approximately 100 episodes, ϵ leans toward the zero value. Therefore, after a certain number of simulation episodes, randomness in choosing actions is reduced. Therefore, further DWL-ST-VSL's action selection relies on gained knowledge. However, it should be pointed out that the learning process is not completely done. This is the case due to the nature of the WL algorithm (see section 6.3.3) that uses a mechanism that allows for learning Q -function somewhat before W -values. Thus, during the entire training process agents are allowed to learn W -values. This mechanism, as presented by the author of the WL algorithm, ensures that W s will settle down with the agent's visits of state-action pairs. Thus, among other randomness listed above, changes in the TTS curve might be influenced by switching between the agent's winner policies determined by W s. However, the updates of W s might be, for example, stopped by simply defining the thresholds of TTS in the learning phase but it is not considered in this experiment.

7.5.2 Spatiotemporal analysis of dynamic VSL zones adaptation

Space-time speed limit diagrams are useful for visualizing traffic-responsive VSL operation regarding the speed limit computation together with VSL zone adaptation. Particularly in the context of the performed experiment, it helps to interpret how the VSL control process develops along with the evolving traffic in the proximity of the bottleneck (area a+A3). According to the best achieved performance, in Figure 7.4 the representative traffic scenario controlled by DT-DWL-ST-VSL with cooperation coefficient $C = 0.5$ is used for analysis. Figures 7.4 a, b, and c show three VSL areas A1, A2, A3, and the corresponding VSL1 and VSL2 zones activation and their speed limit values. As an indicator for traffic state, traffic density and traffic speed, along with controller responses are depicted in Figures 7.4 d and e.

The mixed colors of red to yellow and blue in Figures 7.4 a, b, and c correspond to speed limits (60 – 100 [km/h]) within active VSL zones (in VSL1 or VSL2 or in both zones). It is worth noting that light blue in the case of VSL1 in area A3 corresponds to the nominal speed limit on this section, which is set at 100 [km/h] due to the proximity of the border crossing and the need to smoothly slow down vehicles to safely adjust speed for passing the border.

Notably, the entire 24 [h] simulation results are shown to highlight the controller’s ability not only to respond appropriately during the congestion period (between 5 and 9 pm) but also to allow vehicles to travel at maximum speed when there is no congestion. This is desirable controller behavior, especially for adaptive controllers that are designed to handle and adapt control policies to different traffic situations. A closer look at traffic-responsive DT-DWL-ST-VSL performance is given in Figure 7.5.

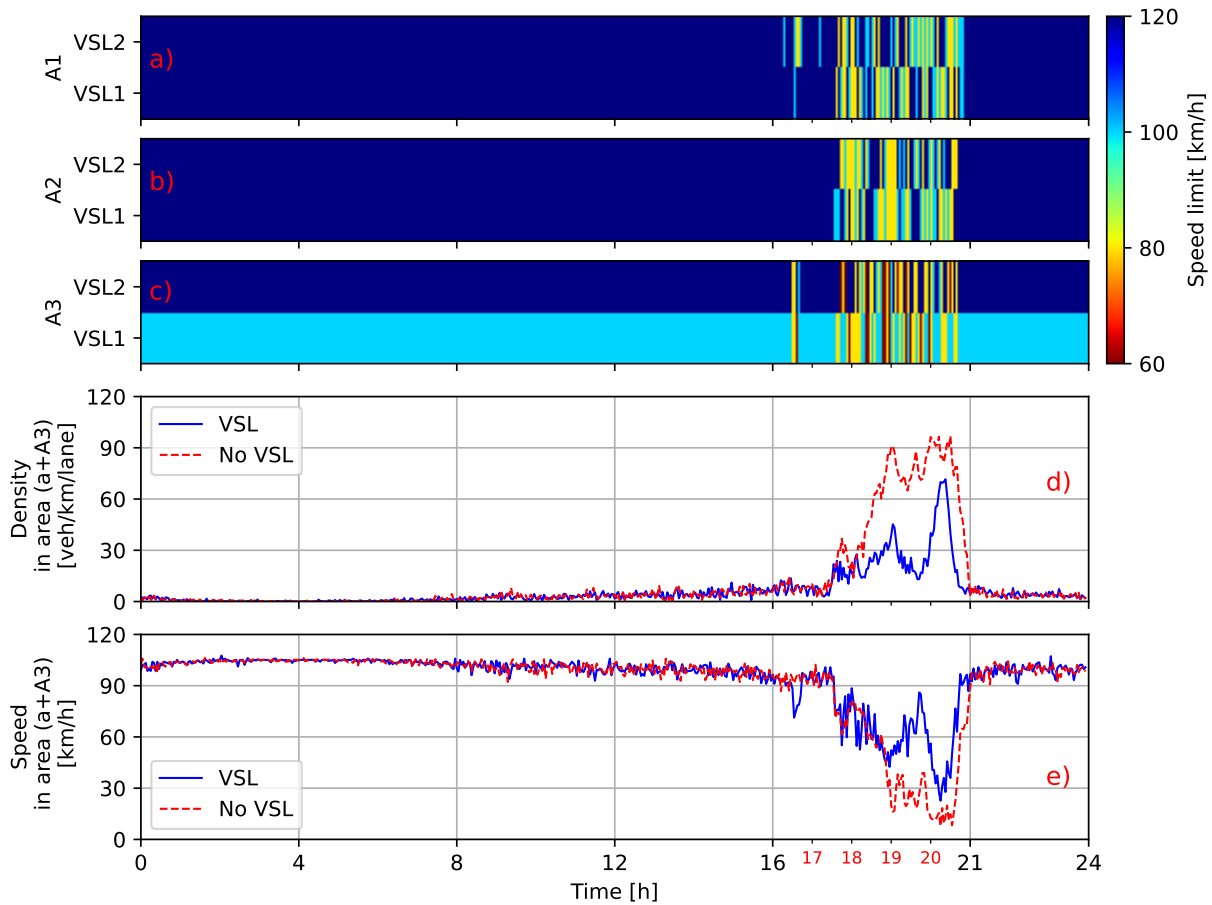


Figure 7.4: Speed limit space-time diagram (a) of area A1, (b) area A2, and (c) area A3 and (d) traffic density, and (e) traffic speed measured in the proximity of the bottleneck (area a+A3)

The first activation of the VSL occurs in VSL areas A1 and A3. This can be seen as an early proactive control response aimed at preventing possible adverse traffic conditions to evolve further that could lead to severe congestion since the VSL agents have acquired knowledge about traffic characteristics that are likely to lead to unstable traffic through training on real-time fine-

grained traffic data with randomized traffic simulation scenarios that comprise a variety of traffic patterns and traffic state modalities. Thus, VSL agents are able to recognize the predecessors that indicate the bottleneck activation and proactively react. The reason for the earlier proactive activation of agent A_1 could be that agent A_1 is facing an increased traffic flow from the east, which may not bother it locally, but it has been learned that the approaching traffic wave may be critical for downstream agent A_3 , which is operating near the bottleneck, and, therefore, agent A_3 needs additional help from upstream agents. Consequently, agent A_3 may implicitly ask for help via RPs, which is detected by agent A_1 , who activates VSL and helps reduce the flow to the south. Noticeable VSL activation begins approximately at 5:30 pm when traffic density increases near the bottleneck. It is evident that during the peak traffic period, different VSL zones configurations have been implemented along with different speed limits. As a result, the traffic density is kept at much lower values compared to the case without control (No VSL). At the same time, the traffic speed is also higher.

The simulation-based DT-DWL-ST-VSL performance evaluation in DT-GM on actual fine-grained real traffic data shows that agents are able to activate additional VSL zones within all available VSL application areas (A_1 , A_2 , and A_3) during the highest congestion peak (between 7:00 and 8:30 pm). It is evident that agents A_1 and A_2 assist their downstream neighbor A_3 . This contributes to an even more effective congestion resolution than a case with noncooperative agents. Moreover, with dynamic VSL zones adaptation, an improvement is generally achieved over static VSL zones as summarized by average results presented in Table 7.1.

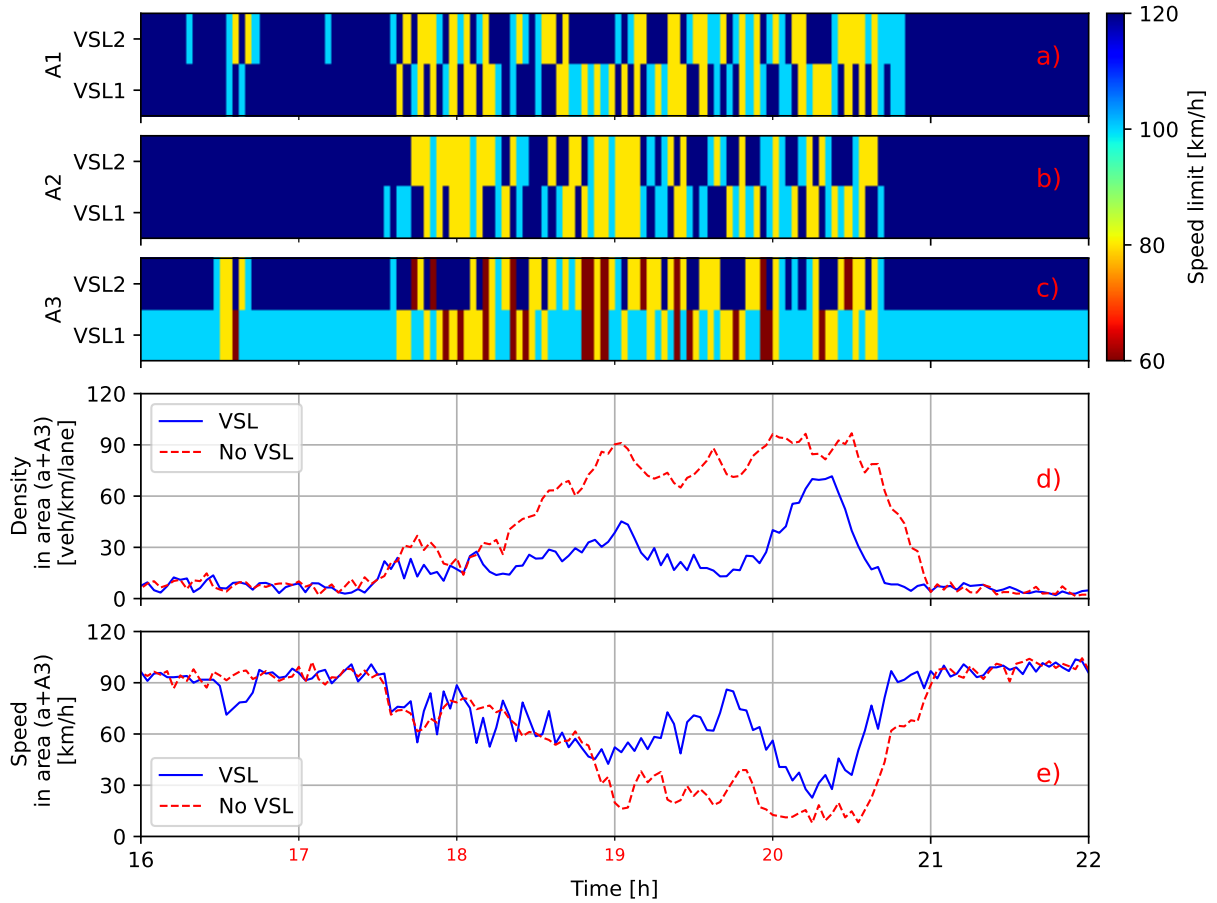


Figure 7.5: Closer look on dynamic VSL zones adaptation

7.5.3 Level of cooperation analysis

To evaluate the benefits of cooperation between agents using the DWL concept of remote policies, the effects of the cooperation coefficient on agent performance are also assessed. The influence of different levels of agent cooperation on system performance are shown in Table 7.1. The absolute improvement or deterioration in traffic performance is expressed based on the value obtained without VSL control. The measurements for speed and density are expressed as the 24-hour average of several random simulation runs. The *TTS* is obtained using the procedure explained in section 7.5.1.

Note that for a comparison between static and dynamic VSL zones, the initial VSL application domains are the same in both cases. It goes without saying that in the VSL experiment with static zones, the speed limit computed by VSL is implemented in both zones (VSL1 and VSL2). Consequently, the static VSL zone is guaranteed. However, with dynamic VSL zones, the agents can switch between activating the VSL zones during operation and thus adjust the different VSL zones with possibly different speed limits over time. In this way, agents can collectively define multiple VSL zone configurations that result in significant differences in VSL performance compared to static VSL zones. Accordingly, agents with dynamic VSL zone adaptation are

Table 7.1: AVERAGE PERFORMANCE STATISTICS OF TESTED APPROACHES IN DT-GM ENVIRONMENT

Case	TTS [veh·h]	Imp. %	Speed [km/h]	Imp %	ρ [veh/km/lane]	Imp. %
No VSL	8179.9		94.12		9.57	
DT-DWL-ST-VSL $C=0$	8039.2	1.72	93.50	-0.65	8.31	13.11
DT-DWL-ST-VSL $C=0.5$	7874.0	3.74	95.02	0.96	6.85	28.43
DT-DWL-ST-VSL $C=1$	7961.6	2.67	94.11	-0.01	7.42	22.46
DT-DWL-ST-VSL-static $C=0$	8018.2	1.98	94.85	0.78	7.63	20.25
DT-DWL-ST-VSL-static $C=0.5$	7968.8	2.58	94.43	0.33	7.78	18.62
DT-DWL-ST-VSL-static $C=1$	7905.7	3.35	94.74	0.66	7.10	25.77

generally able to respond to spatial shifts in traffic congestion. The benefits and necessity of adaptive spatiotemporal VSL control are summarized in the results presented in Table 7.1 and previously visualized in Figures 7.5 a, b and c. Thus, different VSL zone configurations per traffic scenario are learned (without requiring manual setup) and dynamically assigned to better respond to spatially propagating traffic congestion on the motorway in the Geneva region.

It can be seen that the DT-DWL-ST-VSL approach with dynamic VSL zones adaptation outperforms other approaches. The lowest TTS value is 7,874.0 [veh·h] for $C = 0.5$. Compared to the No VSL case, ($TTS = 8,179.9$ [veh·h]), representing a reduction of 3.74%. The best average density in the area (a+A3) is 6.85 [veh/km/lane] for $C = 0.5$, while it is 9.57 for the case of No VSL, which is an improvement of 28.43%. In particular, the average vehicle speed for $C = 0.5$ in the case of DT-DWL-ST-VSL is 95.02 [km/h], while the speed in the case of No VSL is 94.12 [km/h], which is an improvement of 0.96%. Although the improvements may seem small at first, they represent a significant cumulative improvement in traffic performance over the year.

7.6 Discussion

The presented results of the above experiment prove that the proposed DT-GM, calibrated in real time, can be integrated into learning-based traffic control systems and used for their training/evaluation. Moreover, it is shown that DWL-ST-VSL with a dynamic VSL zones adaptation model is able to cope with the uncertainties introduced into the control process by using fine-grained actual traffic data from the real motorway in the Geneva region in DT-GM.

With respect to the results presented, it is a reasonable assumption that, for example, after a longer period of time (e.g., 24 [h]) all vehicles will have passed the measurement point and, thus, the same cumulative throughput will be achieved, either under the VSL application or without control measures. However, improving the operating capacity at critical moments over a shorter period implies the elimination or reduction of congestion, i.e., the minimization of

TTS, which implies a reduction of traffic density and an increase of speed in critical motorway areas on finer time scale resolutions, which is confirmed by the obtained simulation results.

In general, the results summarize all relevant aspects of the presented controller designs and their impact on traffic. Accordingly, the capability of the proposed dynamic VSL zone adaptation model driven by cooperative learning-based DWL agents is able to provide an improvement over the VSL control case with static VSL zones and, of course, over the case without VSL control.

However, the best improvement in VSL control performance with dynamic VSL zones is noticeable for the cooperation coefficient $C = 0.5$. This indicates that the main advantage of cooperation in the presence of high traffic load is that the actions nominated by the neighbors are performed only when the action performed is really important for them. It is also observed that for other tested levels, the non-cooperative ($C = 0$) scenario is the worst cooperation level, while for the full cooperation $C = 1$ the results differ from those obtained with the $C = 0.5$ level. This highlights the importance of DWL's ability to enable agents to engage in different levels of cooperation, as controlled by the cooperation coefficient since neither fully cooperative nor fully non-cooperative scenarios yield the best results.

However, due to the sparse granularity of the cooperation coefficient, it is generally not possible to definitively answer whether one cooperation setup is better than the other. Instead, in the context of this experiment, this comparison aims to shed light on the similarities and differences between the two VSL control designs tested (static and dynamic zones), especially when considering the thresholds for cooperation ($C = 0$ and $C = 1$). Thus, it may be that finer control of the cooperation level is needed.

7.7 Concluding remarks

This chapter presented the functional integration of DT-GM and DWL-ST-VSL called DT-DWL-ST-VSL. It, thus, provided a simulation-based proof of the practical use and potential of DT simulations in the development and rigorous evaluation of learning traffic control systems. The results of the conducted experiment proved that the proposed DT-GM, which was calibrated in real time with fine-grained real traffic data, can be integrated with DWL-ST-VSL and used as a framework for training and evaluating DWL-ST-VSL. Moreover, it was shown that DWL-ST-VSL is able to deal with uncertainties in real traffic by appropriately controlling real traffic on the Geneva motorway simulated via DT-GM.

Chapter 8

Conclusion and future work

This thesis has addressed the problem of advancing motorway traffic flow optimization through the new concept of dynamic adaptation of VSL zones based on cooperative learning agents, thus, dynamic VSL zones configuration selection is learned rather than manually designed. Moreover, the concept of a microsimulation-based DT for motorways has been proposed, which extends the online *simulation-based* learning process and analytics of MARL-VSL agents by introducing traffic uncertainty mapped to the running simulation process directly from the real motorway in real time. Contributions of this thesis along with the perspectives for future research are, thus, summarized in continuation.

Contributions of the thesis

The efficiency of learning traffic controllers depends heavily on the training process performed in the simulations. Therefore, this thesis proposed for the first time the DT paradigm for motorway traffic simulation modeling. That is, methodological steps were presented for the functional integration of actual fine-grained traffic data received in real time from the physical sensors on the Geneva motorway with an ongoing microscopic simulation, i.e., an online calibration process of the microscopic traffic simulation using actual traffic data streams is proposed. As shown, the proposed DT-GM model is able to accurately reflect the traffic dynamics of its physical motorway counterpart in real time.

Moreover, the necessity of spatiotemporal dynamic adjustment of VSL zones was analyzed. This assumption is based on the fact that the intensity and area of congestion on motorways are not stationary, i.e., they can change in time and space, and, therefore, it is necessary to adjust the positions of VSL zones to respond appropriately to congestion. To explain the need for dynamic VSL zones and for a model to dynamically adjust VSL zones, the nondeterministic properties of motorway capacity are considered, as well as the fact that capacity is not constant

but changes with position and time. This phenomenon is used to justify the need for dynamic VSL zones. On this basis, a model for dynamic spatiotemporal adjustment of VSL zones is proposed as a new concept for VSL on motorways. It is shown that dynamic adaptation of VSL zones, together with speed limit adaptation, outperforms the prevailing VSL with static VSL zones. Specifically, it has been shown that dynamic adaptation of the VSL application area upstream of the congestion region can further increase the operational capacity of active motorway bottlenecks without violating upstream traffic conditions.

To enable spatiotemporal adaptation of VSL zones and simultaneous control of speed limit, the DWL-ST-VSL, a distributed multi-agent multi-policy RL-based VSL controller has been proposed. First, the DWL-ST-VSL approach was analyzed in a traffic simulation scenario on an urban motorway, where cooperative agents learned to jointly control multiple adaptive VSL zones ahead of a congested area using the DWL-based control algorithm. The simulations show that DWL-ST-VSL consistently performs better than the baseline solutions used with static VSL zones. The results show that DWL-ST-VSL improves the VSL control by simultaneously adjusting the speed limit values and the VSL zones configurations in response to spatiotemporal changes in the prevailing congestion characteristics. The VSL control performance is enhanced by DWL's inherent ability to implement multiple heterogeneous control policies simultaneously, along with heterogeneous action sets (e.g., different speed limit granularity per agent), which enables smoother spatial speed control of vehicles on motorway sections. In addition, VSL control performance (and overall motorway performance) is improved by cooperation between VSL agents along the motorway corridor via DWL's remote policies.

Finally, this thesis has also provided simulation-based evidence of the practical use and potential of DT simulations in the development and rigorous evaluation of learning-based traffic control systems. For this purpose, the functional integration of DT-GM and DWL-ST-VSL has been proposed. The presented results of this experiment have proved that the proposed DT-GM, calibrated in real-time with actual traffic data, can be integrated with and used as a framework for the training and evaluation of DWL-ST-VSL. By doing so, it has been shown that DWL-ST-VSL is able to cope with uncertainties in real traffic by appropriately controlling real traffic on the Geneva motorway simulated by DT-GM.

Although this thesis is a fairly comprehensive study, the work is based on some assumptions that might affect the run-time DT-GM microscopic simulation results and tested DWL-ST-VSL. One important factor is the limited number of traffic counters on the observed real motorway segment used for modeling DT-GM. This is partially overcome with the traffic flow conservation model explained in the digital-twin chapter, which in general approximates missing traffic data for particular locations in the DT-GM model. Another important factor appears to be the evolving traffic on the motorway network, which is likely to be affected by the introduction

of VSL. Since no information is available on unregulated or regulated traffic using the VSL system, the DT-DWL-ST-VSL experiment is based on the assumption that nominal (maximum allowable) speed limits are used on Swiss motorways. Another important simplification is that the DWL-ST-VSL optimization process, i.e., the objectives of the agent's control policies, and the presented multi-objective analysis of the optimal allocation of VSL zones do not explicitly consider safety aspects as optimization criteria. Thus, further studies taking into account both traffic performance and safety criteria as objectives are desired.

Perspectives for future research

One of the main research questions related to the dynamic adaptation of VSL zones is to provide a formal model that can be used as a guideline for the design of VSL zones in general. Despite the fact that DWL-ST-VSL is able to extract useful features about the controlled process from sequences of traffic states (and their patterns), and, accordingly, make proactive traffic-responsive decisions about optimal VSL zones adjustment, it would be useful to perform inductive backward reasoning based on the results in order to obtain more explainable conclusions and arguments on the optimal adaptation of VSL zones and to provide a more formal procedure for the *a priori* design of VSL zones in general. Thus, this thesis posits the research direction that needs to be explored in order to provide a formal proof of the proposed conjecture about the dynamic adaptation of VSL zones as a function of probabilistic capacity, i.e., based on the fact that capacity is not constant along the motorway and taking into account the spatial and temporal variations of traffic congestion characteristics on motorways. Therefore, it would be beneficial to provide a general physics-based model for optimal VSL zone setup, rather than using only the data-driven machine learning-based approach as in DWL-ST-VSL. This is desirable because it could further simplify and accelerate the learning process in DWL-ST-VSL (making it more scalable) and generally enable the improvement of other VSL control approaches, especially those based on, e.g., classical feedback control. It would also help practitioners in implementing VSL control strategies.

Besides, the perspective of heterogeneity of DWL policies with different spatial and temporal scopes makes it possible to complement DWL-ST-VSL by simultaneously controlling traffic flow entering the motorway via ramp metering control. Before doing so, however, it would be beneficial to investigate how and to what extent ramp metering control affects the optimal allocation of the VSL application area.

Given changing traffic conditions, it may be useful to also consider the adaptive control cycle of DWL-ST-VSL. Thus, by including dynamic VSL zones and dynamic control cycles, the VSL can become more adaptive, which makes the performance of the VSL more robust when operating in a nonstationary environment. However, to take full advantage of adaptability, the

system's time constants should be long enough for it to ignore false disturbances but short enough to respond to indicative changes in the environment. Particularly, the introduction of an adaptive control cycle in DWL-ST-VSL will require a different paradigm for modeling the control process. For example, semi-Markov decision processes that allow the use of actions that are not of equal duration can be used for this purpose. Thus, further research in this direction is desirable, as is a sensitivity analysis of the DWL-ST-VSL control cycle duration. This will allow for further advancement of the VSL system towards instantaneous speed control in the presence of emerging vehicle-to-infrastructure technologies.

Due to RL-VSL agents interacting with a stateful environment via their actions, another interesting equation arises. For a good control system, one must accurately estimate the current controller process state in order to come up with an appropriate decision. In the case of applying DWL-ST-VSL on motorways, the motorway segment is divided into smaller homogeneous sections where agents observe the traffic state. This raises a new question about the optimal section length when considering, for example, the density field. If the sample length is too short, a so-called microscopic uncertainty can occur, and if the length is too long, a macroscopic uncertainty will appear in the measurements (e.g., an analogy from fluid mechanics). Thus, the general belief that the shorter the observed motorway sections, the more accurate the state estimation and the better the decision-making of the learning controller, may not be the case. As a result, increasing the resolution of the state description may potentially lead to an exponential increase in the computational complexity of agents rather than significantly improving the control process. To preserve the continuum property of traffic flow measurements between agents and, thus, between dynamic VSL zones, the range of appropriate section lengths within which sample measurements are desirable should be systematically defined. Further research on this topic is, therefore, desirable.

In general, the analysis of DWL-ST-VSL suggests that there may be multiple local optima for different levels of cooperation coefficients, which requires further sensitivity analysis of this parameter. The question of how resilient the learning system is to the loss of information exchange when one or more agents fail, which is common in a real-world scenario where sensors and equipment are imperfect and can fail, also highlights open research directions. This raises the question of whether VSL agents can learn the most appropriate cooperation coefficient per pair of agents and learn the size and scope of the group of agents with whom it is beneficial to cooperate. This can be particularly useful when the control process is not symmetric, e.g., in the case of asymmetric road topology and/or traffic demand, to avoid manual planning of agent cooperation level. For example, in the case of traffic on an urban motorway, agent cooperation level during morning rush hour may not be appropriate for afternoon rush traffic, etc.

In addition, to ensure wider applicability of DWL-ST-VSL and better generalization of the con-

troller in nonlinear dynamical systems, an extension of the underlying RL algorithm in DWL-ST-VSL to the continuous state-action space along with the nonlinear function approximation technique for the Q-function approximation also represent an interesting future research contribution.

As for DT-GM, a further granularity of traffic data, which may be feasible in the foreseeable future, would provide the basis for generating instantaneous (ideally event-based) traffic demand inputs to the running microscopic traffic simulation. This would, theoretically, allow accurate quantification of traffic uncertainties, i.e., estimation of the natural distribution of traffic flow during system run-time, so that the DT-GM would behave as an accurate mirror of its physical correspondence. In addition, this information could enable the online calculation (and even calibration) of a variety of parameters, such as the distribution of headways, which are important parameters in microscopic traffic models but have not yet been estimated/calibrated online during system run-time.

The current implementation of DT-GM allows for the parallelization of simulations, i.e., the creation of multiple virtual environments that can be used for confident simulation-based evaluation of control strategies during system run-time. While this has been integrated into the conceptual DTI framework and partially demonstrated by wiring DT-GM and DWL-ST-VSL, it has yet to be implemented and tested in real time. Future work will, therefore, explore the potential of applying DT-GM to traffic optimization in a safety-critical decision context. The focus will be on conceptual modeling of the bidirectional interaction between the DT simulations and the physical traffic system, i.e., closing the information loop between DT and the physical world to achieve digital-physical convergence. Future plans also include adding real-time analytics to the DT-DWL-ST-VSL control system to detect significant changes in the environment e.g., changes in the mean and variance of traffic data streams. Accordingly, DT-GM (and DTI) can serve as a virtual framework in which DWL-ST-VSL can be retrained during system run-time to provide a continuous learning scheme and enable DWL-ST-VSL continuously adapt to changes in a nonstationary environment. Besides, transfer learning, where a previously trained machine learning model can be reused for a new problem, and enhancing *social skills* of VSL agents taking into account coordination and negotiation between agents are also open research areas in the MARL-VSL environment in general.

Therefore, further research on the application and potential of DT in the field of adaptive transport systems and learning-based traffic control, but also for real-time traffic prediction and analytics in general, is desirable.

Bibliography

- [1] Khondaker, B., Kattan, L., “Variable speed limit: an overview”, *Transportation Letters*, Vol. 7, No. 5, 2015, pp. 264-278.
- [2] Strömgren, P., Lind, G., “Harmonization with variable speed limits on motorways”, *Transportation Research Procedia*, Vol. 15, 2016, pp. 664-675.
- [3] Carlson, R. C., Papamichail, I., Papageorgiou, M., “Comparison of local feedback controllers for the mainstream traffic flow on freeways using variable speed limits”, in 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2011, pp. 2160-2167.
- [4] Papageorgiou, M., Kosmatopoulos, E., Papamichail, I., “Effects of variable speed limits on motorway traffic flow”, *Transportation Research Record: Journal of the Transportation Research Board*, 2008, pp. 37–48.
- [5] Hegyi, A., Hoogendoorn, S., “Dynamic speed limit control to resolve shock waves on freeways - field test results of the SPECIALIST algorithm”, in 13th International IEEE Conference on Intelligent Transportation Systems, 2010, pp. 519-524.
- [6] Grumert, E., Tapani, A., Ma, X., “Characteristics of variable speed limit systems”, *European Transport Research Review*, Vol. 10, 2018.
- [7] Shao-long, G., Jun, M., Jun-li, W., Xiao-qing, S., Yan, L., “Methodology for variable speed limit activation in active traffic management”, *Procedia - Social and Behavioral Sciences*, Vol. 96, 2013, pp. 2129-2137.
- [8] Li, D., Ranjitkar, P., “A fuzzy logic-based variable speed limit controller”, *Journal of advanced transportation*, Vol. 49, 2015, pp. 913–927.
- [9] Li, D., Ranjitkar, P., Zhao, Y., “Mitigating recurrent congestion via particle swarm optimization variable speed limit controllers”, *KSCE Journal of Civil Engineering*, Vol. 23, No. 7, 2019, pp. 3174-3179.
- [10] Como, G., Lovisari, E., Savla, K., “Convexity and robustness of dynamic traffic assign-

- ment and freeway network control”, *Transportation Research Part B: Methodological*, Vol. 91, 2016, pp. 446-465.
- [11] Kušić, K., Ivanjko, E., Gregurić, M., Miletić, M., “An overview of reinforcement learning methods for variable speed limit control”, *Applied Sciences*, Vol. 10, No. 14, 2020.
 - [12] Lu, X. Y., Varaiya, P., Horowitz, R., Su, D., Shladover, S. E., “A new approach for combined freeway variable speed limits and coordinated ramp metering”, *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2010, pp. 491–498.
 - [13] Zhang, Y., Sirmatel, I. I., Alasiri, F., Ioannou, P. A., Geroliminis, N., “Comparison of feedback linearization and model predictive techniques for variable speed limit control”, in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018.
 - [14] LA, P., Bhatnagar, S., “Reinforcement learning with function approximation for traffic signal control”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, No. 2, 2011, pp. 412-421.
 - [15] Lu, C., Huang, J., Gong, J., “Reinforcement learning for ramp control: An analysis of learning parameters”, *PROMET - Traffic&Transportation*, Vol. 28, No. 4, 2016, pp. 371-381, available at: <https://traffic.fpz.hr/index.php/PROMTT/article/view/1830>
 - [16] Gong, I., Oh, S., Min, Y., “Train Scheduling with Deep Q-Network: A Feasibility Test”, *Applied Sciences*, Vol. 10, No. 23, 2020.
 - [17] Gueriau, M., Cugurullo, F., Acheampong, R. A., Dusparic, I., “Shared autonomous mobility on demand: A learning-based approach and its performance in the presence of traffic congestion”, *IEEE Intelligent Transportation Systems Magazine*, Vol. 12, No. 4, 2020, pp. 208-218.
 - [18] Gosavi, A., *Parametric Optimization Techniques and Reinforcement Learning*, 2nd Edition. Springer, 2015.
 - [19] Zhu, F., Ukkusuri, S. V., “Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach”, *Transportation Research Part C: Emerging Technologies*, Vol. 41, 2014, pp. 30 - 47.
 - [20] Li, Z., Liu, P., Xu, C., Duan, H., Wang, W., “Reinforcement Learning-Based Variable Speed Limit Control Strategy to Reduce Traffic Congestion at Freeway Recurrent Bottlenecks”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, 2017, pp. 3204–3217.
 - [21] Walraven, E., Spaan, M. T., Bakker, B., “Traffic flow optimization: A reinforcement

- learning approach”, *Engineering Applications of Artificial Intelligence*, Vol. 52, 2016, pp. 203–212.
- [22] Zhou, W., Yang, M., Lee, M., Zhang, L., “Q-learning-based coordinated variable speed limit and hard shoulder running control strategy to reduce travel time at freeway corridor”, *Transportation Research Record Journal of the Transportation Research Board*, Vol. 2674, No. 11, 2020, pp. 915-925.
- [23] Gregurić, M., Kušić, K., Vrbanić, F., Ivanjko, E., “Variable speed limit control based on deep reinforcement learning: A possible implementation”, in *2020 International Symposium ELMAR*, 2020.
- [24] Schmidt-Dumont, T., van Vuuren, J. H., “A case for the adoption of decentralised reinforcement learning for the control of traffic flow on South African highways”, *Journal of the South African Institution of Civil Engineering*, Vol. 61, 2019, pp. 7 - 19.
- [25] Wang, C., Zhang, J., Xu, L., Li, L., Ran, B., “A new solution for freeway congestion: Cooperative speed limit control using distributed reinforcement learning”, *IEEE Access*, Vol. 7, 2019, pp. 41 947-41 957.
- [26] Kušić, K., Ivanjko, E., Vrbanić, F., Gregurić, M., Dusparic, I., “Dynamic variable speed limit zones allocation using distributed multi-agent reinforcement learning”, in *In Proceedings of the 2021 IEEE 24th International Conference on Intelligent Transportation Systems (ITSC)*, 2021, pp. 1-8.
- [27] Dusparic, I., Cahill, V., “Distributed W-Learning: Multi-Policy Optimization in Self-Organizing Systems”, in *2009 Third IEEE International Conference on Self-Adaptive and Self-Organizing Systems*, Sep. 2009, pp. 20-29.
- [28] Müller, E. R., Carlson, R. C., Kraus, W., Papageorgiou, M., “Microsimulation Analysis of Practical Aspects of Traffic Control With Variable Speed Limits”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 1, 2015, pp. 512-523.
- [29] Martínez, I., Jin, W.-L., “Optimal location problem for variable speed limit application areas”, *Transportation Research Part B: Methodological*, Vol. 138, 2020, pp. 221-246.
- [30] Ni, D., “Limitations of current traffic models and strategies to address them”, *Simulation Modelling Practice and Theory*, Vol. 104, 2020, pp. 102137.
- [31] Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., Wiessner, E., “Microscopic Traffic Simulation using SUMO”, in *2018 21st International Conference on Intelligent Transportation Systems*, 2018, pp. 2575-2582.

-
- [32] Lighthill, M. J., Whitham, G. B., “On kinematic waves I. Flood movement in long rivers. II: A Theory of traffic flow on long crowded roads.”, *Proc. Royal. Soc. Lond.*, Vol. A229, 1955, pp. 281–316.
- [33] Lorenz, M. R., Elefteriadou, L., “Defining freeway capacity as function of breakdown probability”, *Transportation Research Record*, Vol. 1776, No. 1, 2001, pp. 43-51.
- [34] Lai, F., Carsten, O., Tate, F., “How much benefit does intelligent speed adaptation deliver: An analysis of its potential contribution to safety and environment”, *Accident Analysis & Prevention*, Vol. 48, 2012, pp. 63-72.
- [35] Vrbanić, F., Ivanjko, E., Kušić, K., Čakija, D., “Variable Speed Limit and Ramp Metering for Mixed Traffic Flows: A Review and Open Questions”, *Applied Sciences*, Vol. 11, No. 6, 2021.
- [36] HCM 2010, *Highway Capacity Manual*. Washington, D.C., Transportation Research Board, 2010.
- [37] Carlson, R. C., Papamichail, I., Papageorgiou, M., Messmer, A., “Optimal mainstream traffic flow control of large-scale motorway networks”, *IFAC Proceedings Volumes (IFAC-PapersOnline)*, Vol. 42, No. 15, 2010, pp. 1–6.
- [38] Carlson, R., Papamichail, I., Papageorgiou, M., “Local feedback-based mainstream traffic flow control on motorways using variable speed limits”, *Intelligent Transportation Systems, IEEE Transactions on*, Vol. 12, 2011, pp. 1261 - 1276.
- [39] Chung, K., Rudjanakanoknad, J., Cassidy, M. J., “Relation between traffic density and capacity drop at three freeway bottlenecks”, *Transportation Research Part B: Methodological*, Vol. 41, No. 1, 2007, pp. 82-95.
- [40] Vizioli, H. T., Kušić, K., Ivanjko, E., Cunha, A. L., “A method to calibrate variable speed limit control on high-truck share roads”, in *Proceedings of the 6th Brazilian Technology Symposium (BTSym'20)*. Cham: Springer International Publishing, 2021, pp. 204–211.
- [41] Soriguera, F., Martínez, I., Sala, M., Menéndez, M., “Effects of low speed limits on freeway traffic flow”, *Transportation Research Part C: Emerging Technologies*, Vol. 77, 2017, pp. 257-274.
- [42] Gao, C., Xu, J., Li, Q., Yang, J., “The effect of posted speed limit on the dispersion of traffic flow speed”, *Sustainability*, Vol. 11, No. 13, 2019, available at: <https://www.mdpi.com/2071-1050/11/13/3594>
- [43] van den Hoogen, E., Smulders, S., “Control by variable speed signs: results of the Dutch

- experiment”, in Seventh International Conference on Road Traffic Monitoring and Control, 1994, pp. 145-149.
- [44] Yang, Y., Yuan, Z.Z., Sun, D.Y., Wen, X.L., “Analysis of the factors influencing highway crash risk in different regional types based on improved apriori algorithm”, *Advances in Transportation Studies*, Vol. 49, 2019, pp. 165-178.
- [45] Ivanjko, E., Kušić, K., Gregurić, M., “Simulational analysis of two controllers for variable speed limit control”, *Proceedings of the Institution of Civil Engineers - Transport*, Vol. 175, No. 7, 2022, pp. 413-425.
- [46] Baldasano, J. M., Gonçalves, M., Soret, A., Jiménez-Guerrero, P., “Air pollution impacts of speed limitation measures in large cities: The need for improving traffic data in a metropolitan area”, *Atmospheric Environment*, Vol. 44, No. 25, 2010, pp. 2997–3006.
- [47] Soriguera, F., Torné, J. M., Rosas, D., “Assessment of dynamic speed limit management on metropolitan freeways”, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, Vol. 17, No. 1, 2013, pp. 78–90.
- [48] Hegyi, A., “Model predictive control for integrating traffic control measures”, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2004.
- [49] Hegyi, A., De Schutter, B., Hellendoorn, J., “Optimal coordination of variable speed limits to suppress shock waves”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 6, No. 1, 2005, pp. 102–112.
- [50] Papamichail, I., Kampitaki, K., Papageorgiou, M., Messmer, A., “Integrated ramp metering and variable speed limit control of motorway traffic flow”, *IFAC Proceedings Volumes*, Vol. 41, No. 2, 2008, pp. 14 084-14 089, 17th IFAC World Congress.
- [51] Zackor, H., “Speed limitation on freeways: Traffic-responsive strategies”, Papageorgiou, M. (Ed.), *Concise Encyclopedia of Traffic and Transportation Systems*, Vol. 14, 1991, pp. 507–511.
- [52] Hegyi, A., Hoogendoorn, S. P., Schreuder, M., Stoelhorst, H., Viti, F., “Specialist: A dynamic speed limit control algorithm based on shock wave theory”, in 2008 11th International IEEE Conference on Intelligent Transportation Systems, 2008, pp. 827-832.
- [53] Kušić, K., Ivanjko, E., Gregurić, M., “A Comparison of Different State Representations for Reinforcement Learning Based Variable Speed Limit Control”, *MED 2018 - 26th Mediterranean Conference on Control and Automation*, 2018, pp. 266–271.
- [54] Vinitzky, E., Parvate, K., Kreidieh, A., Wu, C., Bayen, A., “Lagrangian Control through

- Deep-RL: Applications to Bottleneck Decongestion”, in 2018 21st International Conference on Intelligent Transportation Systems, Nov 2018, pp. 759-765.
- [55] Humphrys, M., “Action selection methods using reinforcement learning”, Ph.D. thesis, University of Cambridge, 1996.
- [56] Zhang, Y., Ma, M., Liang, S., “Dynamic control cycle speed limit strategy for improving traffic operation at freeway bottlenecks”, *KSCE Journal of Civil Engineering*, Vol. 25, No. 2, 2021, pp. 692-704.
- [57] Carlson, R. C., Papamichail, I., Papageorgiou, M., Messmer, A., “Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering”, *Transportation Science*, Vol. 44, No. 2, 2010, pp. 238–253.
- [58] Tympakianaki, A., Spiliopoulou, A., Kouvelas, A., Papamichail, I., Papageorgiou, M., Wang, Y., “Real-time merging traffic control for throughput maximization at motorway work zones”, *Transportation Research Part C: Emerging Technologies*, Vol. 44, 2014, pp. 242-252.
- [59] Richards, P. I., “Shock waves on the highway”, *Operations Research*, Vol. 4, No. 1, 1956, pp. 42–51.
- [60] Daganzo, C. F., “Requiem for second-order fluid approximations of traffic flow”, *Transportation Research Part B: Methodological*, Vol. 29, No. 4, 1995, pp. 277-286, available at: <https://www.sciencedirect.com/science/article/pii/019126159500007Z>
- [61] Lu, X.-Y., Varaiya, P., Horowitz, R., Su, D., Shladover, S. E., “Novel freeway traffic control with variable speed limit and coordinated ramp metering”, *Transportation Research Record*, Vol. 2229, No. 1, 2011, pp. 55-65.
- [62] Jeon, S., Park, C., Seo, D., “The multi-station based variable speed limit model for realization on urban highway”, *Electronics*, Vol. 9, No. 5, 2020, available at: <https://www.mdpi.com/2079-9292/9/5/801>
- [63] Wang, W., Cheng, Z., “Variable speed limit signs: Control and setting locations in freeway work zones”, *Journal of Advanced Transportation*, Vol. 2017, 2017, pp. 1-13.
- [64] Banerjee, S., Jelihani, M., Khadem, N. K., Brown, D. D., “Units of information on dynamic message signs: a speed pattern analysis”, *European Transport Research Review*, Vol. 11, No. 15, 2019.
- [65] Huzjan, B., Šoštarić, M., Mandžuka, S., “Analysis of speed limit obedience on croatian highways - technical report”, available at: <https://cordis.europa.eu/project/id/317671>
Znanstveno-istraživački projekt 7. okvirnog programa EU komisije „Intelligent

- Cooperative Sensing for Improved Traffic Efficiency“ 2012-2013., EC-FP7-317671. 2013.
- [66] Matowicki, M., Pribyl, O., “A driving simulation study on drivers speed compliance with respect to variable message signs”, *Journal of Intelligent Transportation Systems*, 2022.
- [67] Codeca, L., Härrä, J., “Towards multimodal mobility simulation of C-ITS: The Monaco SUMO traffic scenario”, in *VNC 2017, IEEE Vehicular Networking Conference*, Torino, Italy, 2017.
- [68] Guériaux, M., Dusparic, I., “Quantifying the impact of connected and autonomous vehicles on traffic efficiency and safety in mixed traffic”, in *23rd IEEE International Conference on Intelligent Transportation Systems*, 2020.
- [69] Chu, T., Wang, J., Codecà, L., Li, Z., “Multi-agent deep reinforcement learning for large-scale traffic signal control”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 3, 2020, pp. 1086-1095.
- [70] Yuan, T., Alasiri, F., Ioannou, P. A., “Selection of the Speed Command Distance for Improved Performance of a Rule-Based VSL and Lane Change Control”, *IEEE Transactions on Intelligent Transportation Systems*, 2022, pp. 1-10 .
- [71] Kušić, K., Dusparic, I., Guériaux, M., Gregurić, M., Ivanjko, E., “Extended variable speed limit control using multi-agent reinforcement learning”, in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems*, 2020, pp. 1-8.
- [72] Koch, L., Buse, D. S., Wegener, M., Schoenberg, S., Badalian, K., Dressler, F., Andert, J., “Accurate physics-based modeling of electric vehicle energy consumption in the SUMO traffic microsimulator”, in *2021 IEEE International Intelligent Transportation Systems Conference*, 2021, pp. 1650-1657.
- [73] Troullinos, D., Chalkiadakis, G., Manolis, D., Papamichail, I., Papageorgiou, M., “Lane-Free Microscopic Simulation for Connected and Automated Vehicles”, in *2021 IEEE International Intelligent Transportation Systems Conference*, 2021, pp. 3292-3299.
- [74] Collins, A. J., Robinson, R. M., Jordan, C. A., Khattak, A., “Development of a traffic incident model involving multiple municipalities for inclusion in large microscopic evacuation simulations”, *International Journal of Disaster Risk Reduction*, Vol. 31, 2018, pp. 1223-1230.
- [75] Madni, A. M., Madni, C. C., Lucero, S. D., “Leveraging Digital Twin Technology in Model-Based Systems Engineering”, *Systems*, Vol. 7, No. 1, 2019.
- [76] Liu, M., Fang, S., Dong, H., Xu, C., “Review of digital twin about concepts, technolo-

- gies, and industrial applications”, *Journal of Manufacturing Systems*, Vol. 58, 2021, pp. 346-361.
- [77] Šemanjski, I., *Smart Urban Mobility: Transport Planning in the Age of Big Data and Digital Twins*, ser. 1st Edition. Elsevier Science, 2022.
- [78] Pau, P., Kastner, K.-H., Keber, R., Samal, M., “Real-Time Traffic Conditions with SUMO for ITS Austria West”, 2013.
- [79] Argota Sánchez-Vaquerizo, J., “Getting Real: The Challenge of Building and Validating a Large-Scale Digital Twin of Barcelona’s Traffic with Empirical Data”, *ISPRS International Journal of Geo-Information*, Vol. 11, No. 1, 2022.
- [80] Dasgupta, S., Rahman, M., Lidbe, A. D., Lu, W., Jones, S., “A Transportation Digital-Twin Approach for Adaptive Traffic Control Systems”, *arXiv preprint arXiv:2109.10863*, 2021.
- [81] White, G., Zink, A., Codecá, L., Clarke, S., “A digital twin smart city for citizen feedback”, *Cities*, Vol. 110, 2021.
- [82] Kumar, S., Madhumathi, R., Chelliah, P. R., Tao, L., Wang, S., “A novel digital twin-centric approach for driver intention prediction and traffic congestion avoidance”, *Journal of Reliable Intelligent Environments*, Vol. 4, 2018, pp. 199–209.
- [83] Bhattacharya, P., Shukla, A., Tanwar, S., Kumar, N., Sharma, R., “6Blocks: 6G-enabled trust management scheme for decentralized autonomous vehicles”, *Computer Communications*, Vol. 191, 2022, pp. 53–68.
- [84] Liao, X., Zhao, X., Wang, Z., Han, K., Tiwari, P., Barth, M. J., Wu, G., “Game Theory-Based Ramp Merging for Mixed Traffic With Unity-SUMO Co-Simulation”, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, pp. 1-12 .
- [85] Veledar, O., Damjanovic-Behrendt, V., Macher, G., “Digital Twins for Dependability Improvement of Autonomous Driving”, in *Systems, Software and Services Process Improvement*, Walker, A., O’Connor, R. V., Messnarz, R., (ed.). Cham: Springer International Publishing, 2019, pp. 415–426.
- [86] Wang, Z., Han, K., Tiwari, P., “Digital Twin Simulation of Connected and Automated Vehicles with the Unity Game Engine”, in *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence*, 2021, pp. 1-4.
- [87] Wang, Z., Gupta, R., Han, K., Wang, H., Ganlath, A., Ammar, N., Tiwari, P., “Mobility Digital Twin: Concept, Architecture, Case Study, and Future Challenges”, *IEEE Internet of Things Journal*, 2022, pp. 1-1 .

- [88] Liao, X., Wang, Z., Zhao, X., Han, K., Tiwari, P., Barth, M. J., Wu, G., “Cooperative Ramp Merging Design and Field Implementation: A Digital Twin Approach Based on Vehicle-to-Cloud Communication”, *IEEE Transactions on Intelligent Transportation Systems*, 2021, pp. 1-11 .
- [89] Chen, L., Lopez, A. J., Semanjski, I., Gautama, S., Ochoa, D., “Assessment of Smartphone Positioning Data Quality in the Scope of Citizen Science Contributions”, *Mobile Information Systems*, Hindawi, Vol. 2017, 2017.
- [90] Cárdenas-Benítez, N., Aquino-Santos, R., Magaña-Espinoza, P., Aguilar-Velazco, J., Edwards-Block, A., Medina Cass, A., “Traffic congestion detection system through connected vehicles and big data”, *Sensors*, Vol. 16, No. 5, 2016.
- [91] Khoshkhah, K., Pourmoradnasseri, M., Hadachi, A., Tera, H., Mass, J., Keshi, E., Wu, S., “Real-Time System for Daily Modal Split Estimation and OD Matrices Generation Using IoT Data: A Case Study of Tartu City”, *Sensors*, Vol. 22, No. 8, 2022.
- [92] [dataset] ODPMS, “Open data platform mobility Switzerland”, available at: <https://opentransportdata.swiss/en/> Accessed: 14 March 2022. 2021.
- [93] Durand, A., Zijlstra, T., van Oort, N., Hoogendoorn-Lanser, S., Hoogendoorn, S., “Access denied? Digital inequality in transport services”, *Transport Reviews*, Vol. 42, No. 1, 2022, pp. 32-57.
- [94] Schumann, R., *Performance Maintenance of ARTS Systems*. Cham: Springer International Publishing, 2016, pp. 181–195.
- [95] Maimaris, A., Papageorgiou, G., “A review of Intelligent Transportation Systems from a communications technology perspective”, in *2016 IEEE 19th International Conference on Intelligent Transportation Systems*, 2016, pp. 54-59.
- [96] Wegener, A., Piórkowski, M., Raya, M., Hellbrück, H., Fischer, S., Hubaux, J.-P., “TraCI: An Interface for Coupling Road Traffic and Network Simulators”, in *Proceedings of the 11th Communications and Networking Simulation Symposium*, ser. CNS '08. New York, NY, USA: Association for Computing Machinery, 2008, pp. 155–163.
- [97] Li, D., Lasenby, J., “Mitigating urban motorway congestion and emissions via active traffic management”, *Research in Transportation Business & Management*, 2022.
- [98] Kušić, K., Schumann, R., Ivanjko, E., “A digital twin in transportation: Real-time synergy of traffic data streams and simulation for virtualizing motorway dynamics”, *Advanced Engineering Informatics*, Vol. 55, 2023, pp. 101858.

-
- [99] Google Earth, “Region of Geneva”, available at: <https://earth.google.com/web/> Accessed: 24 March 2022. 2022.
- [100] Iordanidou, G., Roncoli, C., Papamichail, I., Papageorgiou, M., “Feedback-Based Mainstream Traffic Flow Control for Multiple Bottlenecks on Motorways”, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 2, 2015, pp. 610-621.
- [101] Kušić, K., Schumann, R., Ivanjko, E., “Building a Motorway Digital Twin in SUMO: Real-Time Simulation of Continuous Data Stream from Traffic Counters”, in *Proc. of 64th International Symposium ELMAR*, 2022.
- [102] OSM, “OpenStreetMap”, available at: <http://download.geofabrik.de/> Accessed: 10 October 2021. 2021.
- [103] Salles, D., Kaufmann, S., Reuss, H.-C., “Extending the Intelligent Driver Model in SUMO and Verifying the Drive Off Trajectories with Aerial Measurements”, in *SUMO User Conference 2020*, 2020.
- [104] Erdmann, J., “SUMO’s Lane-Changing Model”, in *Modeling Mobility with Open Data. Lecture Notes in Mobility.*, Behrisch, M., Weber, M., (ed.). Springer International Publishing, 2015, pp. 105–123.
- [105] Ha, D.-H., Aron, M., Cohen, S., “Time headway variable and probabilistic modeling”, *Transportation Research Part C: Emerging Technologies*, Vol. 25, 2012, pp. 181-201.
- [106] Roy, R., Saha, P., “Headway distribution models of two-lane roads under mixed traffic conditions: a case study from India”, *European Transport Research Review*, Vol. 10, No. 3, 2017.
- [107] Qasim, G., Jameel, A., Abdulwahab, A., Rajaa, A., “Estimating a congested road capacity-headway relationship of a multi-lane highway in an urban area based on lane position”, *Periodicals of Engineering and Natural Sciences (PEN)*, Vol. 8, No. 3, 2020, pp. 1263-1279.
- [108] Babić, D., Babić, D., Fiolić, M., Eichberger, A., Magosi, Z. F., “A Comparison of Lane Marking Detection Quality and View Range between Daytime and Night-Time Conditions by Machine Vision”, *Energies*, Vol. 14, No. 15, 2021.
- [109] Wang, Y., “Dynamic variable speed limit control: Design, analysis and benefits”, Ph.D. thesis, University of Southern California, 2011.
- [110] Miloš, J., Hršak, P., Topić, N., Jakšić, L., Kušić, K., Vrbanić, F., Ivanjko, E., “Influence of spatial placement of variable speed limit zones on urban motorway traffic control”, *Promet - Traffic&Transportation*, Vol. 34, No. 4, 2022, pp. 511-522.

- [111] Sutton, R., Barto, A., Reinforcement Learning: An Introduction, ser. A Bradford book. Bradford Book, 1998.
- [112] Watkins, C. J. C. H., Dayan, P., “Q-learning”, Machine Learning, Vol. 8, No. 3-4, 1992, pp. 279–292.
- [113] Watkins, C. J. C. H., Dayan, P., “Q-learning”, in Machine Learning, 1992, pp. 279–292.
- [114] Gregurić, M., Kušić, K., Ivanjko, E., “Impact of Deep Reinforcement Learning on Variable Speed Limit strategies in connected vehicles environments”, Engineering Applications of Artificial Intelligence, Vol. 112, 2022, pp. 104850.
- [115] Li, Z., Liu, P., Xu, C., Duan, H., Wang, W., “Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks”, IEEE Transactions on Intelligent Transportation Systems, Vol. 18, No. 11, Nov 2017, pp. 3204-3217.
- [116] Walraven, E., “Traffic Flow Optimization using Reinforcement Learning”, Master’s thesis, Faculty EEMCS, Delft University of Technology, the Netherlands, 2014.
- [117] Kušić, K., “Framework for Simulation of Variable Speed Limit Control Systems on Urban Motorways Based on Learning”, Master’s thesis, University of Zagreb, Faculty of Transport and Traffic Science, Croatia, 2017.
- [118] Sherstov, A. A., Stone, P., “Function approximation via tile coding: Automating parameter choice”, in Proceedings of the 6th International Conference on Abstraction, Reformulation and Approximation, ser. SARA’05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 194–205.
- [119] Prashanth, L. A., Bhatnagar, S., “Reinforcement learning with function approximation for traffic signal control”, IEEE Transactions on Intelligent Transportation Systems, Vol. 12, No. 2, 2011, pp. 412-421.
- [120] Kretchmar, R. M., Anderson, C. W., “Comparison of CMACs and radial basis functions for local function approximators in reinforcement learning”, in International Conference on Neural Networks, Vol. 2, Jun 1997, pp. 834-837.
- [121] Kušić, K., Korent, N., Gregurić, M., Ivanjko, E., “Comparison of two controllers for variable speed limit control”, in 2016 International Symposium ELMAR, Zadar, Croatia, Sept 2016, pp. 101-106.
- [122] Papamichail, I., Kampitaki, K., Papageorgiou, M., Messmer, A., “Integrated ramp metering and variable speed limit control of motorway traffic flow”, IFAC Proceedings Volumes, Vol. 41, No. 2, 2008, pp. 14 084 - 14 089.

-
- [123] Schumann, R., *Engineering Coordination: A Methodology for the Coordination of Planning Systems*. IOS Press, 2011.
- [124] Wegener, A., Piorkowski, M., Raya, M., Hellbrück, H., Fischer, S., Hubaux, J.-P., “TraCI: An Interface for Coupling Road Traffic and Network Simulators”, *Proceedings of the 11th Communications and Networking Simulation Symposium, CNS’08*, 2008.
- [125] Kušić, K., Ivanjko, E., Vrbanić, F., Gregurić, M., Dusparic, I., “Spatial-Temporal Traffic Flow Control on Motorways Using Distributed Multi-Agent Reinforcement Learning”, *Mathematics*, Vol. 9, No. 23, 2021.
- [126] Busoniu, L., Babuska, R., De Schutter, B., “A comprehensive survey of multiagent reinforcement learning”, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 38, No. 2, 2008, pp. 156-172.
- [127] Chen, X., Chen, H., Yang, Y., Wu, H., Zhang, W., Zhao, J., Xiong, Y., “Traffic flow prediction by an ensemble framework with data denoising and deep learning model”, *Physica A: Statistical Mechanics and its Applications*, Vol. 565, 2021, pp. 125574.
- [128] Schumann, R., Tamarcaz, C., “Towards systematic testing of complex interacting systems”, *Proceedings of the First Workshop on Systemic Risks in Global Networks co-located with 14. Internationale Tagung Wirtschaftsinformatik (WI 2019)*, 2019, pp. 55-63.
- [129] Haykin, S., *Neural Networks and Learning Machines*, ser. Third Edition. Prentice Hall, 2008.

List of Figures

2.1. Fundamental diagrams of functional dependencies between flow rate, traffic density, and speed [36]	12
2.2. VSL concept for mainstream traffic control on motorway [37]	14
2.3. Flow-density fundamental diagram with capacity drop at critical density ρ_{cr} [40]	14
2.4. VSL concept for mainstream traffic control on motorway [37]	15
2.5. Temporary reduction of traffic flow due to the transition of density (from a low value to a higher one) in the VSL application area	18
2.6. Impact of a new imposed speed limit v_2 on simplified flow-density triangular fundamental diagram [28]	19
2.7. Fundamental diagram for different VSL rates [38]	20
2.8. Application of VSL to increase throughput in areas with bottlenecks	21
2.9. Feedback VSL control scheme	22
2.10. Variation of TTS with the length of (a) application and (b) acceleration areas [28]	25
3.1. Physical motorway (left) and abstract corresponding model (right)	42
3.2. Scheme of run-time synchronized DT-GM and concept of bidirectional interaction between virtual and physical motorway using DTI-GM in real-time TM . .	43
3.3. Predictive analytics with DTI	44
3.4. Dynamic Flow Calibration principle in SUMO	50
3.5. Lognormal headway distribution model	54
3.6. Illustration of study motorway area	55
3.7. Comparison of actual and simulated daily flow with minute resolution	57
3.8. Box-Whistler diagram for actual and calibrated traffic at minute resolution . . .	58
3.9. Comparison of spatiotemporal speed distribution between actual traffic and run-time DT-GM	61
4.1. Set up of VSL zones for the feedback VSL controller	69
4.2. Simulated traffic demand	70
4.3. Variation of TTS with the lengths of VSL zones and acceleration zones	71

4.4. Variation of average speed in L3 with the lengths of VSL zones and acceleration zones	71
4.5. Variation of traffic density in L3 with the lengths of VSL zones and acceleration zones	71
4.6. Spatiotemporal diagrams for best obtained VSL configurations	74
4.7. Average speed for the best VSL configuration 300 – 300 – 100	75
4.8. Speed limits for the best VSL configuration 300 – 300 – 100	75
4.9. Scheme for VSL application zone design	78
4.10. Pareto front of optimal non-dominated solutions for different configurations of VSL zones for an active speed limit of 80 [km/h] during the congestion period .	78
4.11. Pareto front of optimal non-dominated solutions for different configurations of VSL zones and speed limits	80
4.12. Impact of capacity as a function of position and time on VSL zones positioning	82
4.13. Scheme for dynamic VSL zones adaptation	83
5.1. Concept of applying RL in VSL control	87
5.2. Controlled motorway stretch divided into four sections	89
5.3. An example of a binary feature vector corresponding to 2D state space	94
5.4. An example of one-dimensional radial basis functions [111]	95
5.5. Convergence of the normalized <i>TTS</i> during the training process	97
5.6. Traffic parameters for tile coding in sections L_3 and L_4	98
6.1. Motorway model	103
6.2. Tested traffic scenarios	106
6.3. Results in the dynamic traffic scenario	109
6.4. Convergence of <i>TTS</i> during training process	111
6.5. Results in the static traffic scenario	112
6.6. DWL2-ST-VSL configuration scheme	114
6.7. DWL4-ST-VSL scheme and configuration of motorway segment	118
6.8. Tested traffic scenarios	124
6.9. Space-time diagrams for simulated scenarios with static VSL zones	127
6.10. Space-time diagrams for simulated scenarios with multi-agent dynamic VSL zones	128
6.11. Traffic parameters for different levels of cooperation for the medium traffic load scenario	132
6.12. Traffic parameters for different levels of cooperation for the high traffic load scenario	133
6.13. The convergence of <i>TTS</i> during the training process	134

7.1. Real motorway and its abstract representation	141
7.2. Agents cooperation scheme	142
7.3. Overall network's TTS behavior during the training phase in fully randomized DT-GM simulations	150
7.4. Speed limit space-time diagram (a) of area A1, (b) area A2, and (c) area A3 and (d) traffic density, and (e) traffic speed measured in the proximity of the bottleneck (area a+A3)	152
7.5. Closer look on dynamic VSL zones adaptation	154

List of Tables

3.1. DT IN THE TRANSPORT SYSTEMS AND URBAN MOBILITY	38
3.2. GEH STATISTIC FOR FLOW (OUT) EAST	59
3.3. GEH STATISTIC FOR FLOW (OUT) NORTH	59
3.4. GEH STATISTIC FOR FLOW (OUT) SOUTH	60
4.1. OBTAINED MoEs IN CASE OF VSL ZONES LENGTH 100-100, AND DIFFER- ENT ACCELERATION ZONE LENGTHS	72
4.2. OBTAINED MoEs IN CASE OF VSL ZONES LENGTH 300-300, AND DIFFER- ENT ACCELERATION ZONE LENGTHS	72
4.3. OBTAINED MoEs IN CASE OF VSL ZONES LENGTH 500-500, AND DIFFER- ENT ACCELERATION ZONE LENGTHS	73
4.4. Characteristics of non-dominated optimal solutions (<i>Pareto front</i>) for different configurations of VSL zones with a speed limit of 80 [km/h]	79
4.5. Characteristics of non-dominated optimal solutions (<i>Pareto front</i>) for different configurations of VSL zones and speed limits	80
5.1. VSLs PERFORMANCE RESULTS	97
6.1. VSLs PERFORMANCE RESULTS	110
7.1. AVERAGE PERFORMANCE STATISTICS OF TESTED APPROACHES IN DT-GM ENVIRONMENT	155

List of Algorithms

3.1. DT-GM at each simulation step	52
6.1. DWL2-ST-VSL at each learning step	117
7.1. DT-DWL-ST-VSL at each simulation step	147

Nomenclature

AI	Artificial Intelligence
API	Application Programming Interface
ATSC	Adaptive Traffic Signal Control
AWS	Amazon Web Services
BA-LWR	bounded acceleration LWR
CAVs	Connected Automated Vehicles
CVs	Connected Vehicles
DFC	Dynamic Flow Calibrator
DT	digital twin
DT-GM	digital twin model of the Geneva motorway
DTI	Digital Twin Instances
DWL	distributed W-learning
DWL-ST-VSL	distributed spatial-temporal multi-agent VSL
EIDM	Enhanced Intelligent Driver Model
FB-VSL	feedback-based VSL controller
FEDRO	Swiss Federal Roads Office
IND-VSL	two independent VSL agents
IoT	Internet of Things
ISA	intelligent speed assistance
ITS	intelligent transportation systems
LC	lane change

LoS level of service

LWR Lighthill-Whitham-Richards

MARL multi-agent RL

MDPs Markov decision processes

MoEs Measures of Effectiveness

MTFC mainstream traffic flow control

OD origin-destination

OPDMS Open Data Platform Mobility Switzerland

OSM Open Street Map

P-ACC personalized adaptive cruise control

QVSL Q-learning VSL control

QVSL-FA QVSL with function approximation

QVSL-FS QVSL with full state representation

RBF radial basis function

RL reinforcement learning

RL-VSL single agent-based RL controller for VSL

RM ramp metering

SGL-VSL single VSL agent

SPSC-VSL simple proportional speed controller

SUMO Simulation of Urban MObility

TM traffic management

TraCI Traffic Control Interface

TT Travel Time

TTS total time spent

VMS variable message signs

VSL variable speed limit

VV Virtual Vehicle

WL W-learning

Biography

Krešimir Kušić was born in Zagreb on November 30, 1987, where he also completed his primary and secondary education. In 2012, he started studying Traffic Engineering at the Faculty of Transport and Traffic Sciences, University of Zagreb. He obtained his Master's degree in traffic engineering in 2017. At the same faculty, in 2017 he enrolled in a Ph.D. program (Technological systems in traffic and transport) and he was employed as a Research assistant at the Department of Intelligent Transportation Systems. He is currently working as a teacher in the following courses: Computer Sciences, Intelligent Transportation Systems, Virtual Reality in Traffic Systems, and Operations Research. His research interest is related to traffic flow optimization on motorways and microsimulation traffic modeling. In 2016, he received the Rector's Award from the University of Zagreb and the Dean's Award as the best graduate student in the program: Intelligent Transportation Systems. He perfected his skills at Trinity College Dublin in Ireland in the scope of three international mobilities. He also received the Swiss Government Excellence Scholarship for the academic year 2021/2022, which he spent as a scientific researcher in Switzerland at the Swiss University of Applied Sciences Western Switzerland.

List of published works

Papers in journals

1. Kušić, Krešimir; Schumann, René; Ivanjko, Edouard, A digital twin in transportation: Real-time synergy of traffic data streams and simulation for virtualizing motorway dynamics. *Advanced Engineering Informatics*, 55 (2023), 101858, 17
doi:10.1016/j.aei.2022.101858
2. Gregurić, Martin; Kušić, Krešimir; Ivanjko, Edouard, Impact of Deep Reinforcement Learning on Variable Speed Limit strategies in connected vehicles environments. *Engineering applications of artificial intelligence*, 112 (2022), 104850, 17
doi:10.1016/j.engappai.2022.104850
3. Ivanjko, Edouard; Kušić, Krešimir; Gregurić, Martin, Simulational analysis of two con-

- trollers for variable speed limit control. Proceedings of the institution of civil engineers-transport, 175 (2022), 7; 413-425 doi:10.1680/jtran.19.00069
4. Miletić, Mladen; Ivanjko, Edouard; Gregurić, Martin; Kušić, Krešimir, A review of reinforcement learning applications in adaptive traffic signal control. IET Intelligent Transport Systems, 16 (2022), 10; 1269-1285 doi:10.1049/itr2.12208
 5. Miloš, Josip; Hršak, Patrik; Topić, Nikola; Jakšić, Leon; Kušić, Krešimir; Vrbanić, Filip; Ivanjko, Edouard, Influence of Spatial Placement of Variable Speed Limit Zones on Urban Motorway Traffic Control. Promet - Traffic&Transportation, 34 (2022), 4; 511-522 doi:10.7307/ptt.v34i4.4073
 6. Vrbanić, Filip; Ivanjko, Edouard; Kušić, Krešimir; Čakija, Dino, Variable Speed Limit and Ramp Metering for Mixed Traffic Flows: A Review and Open Questions. Applied Sciences-Basel, 11 (2021), 6; 2574, 26 doi:10.3390/app11062574
 7. Kušić, Krešimir; Ivanjko, Edouard; Vrbanić, Filip; Gregurić, Martin; Dusparic, Ivana, Spatial-Temporal Traffic Flow Control on Motorways Using Distributed Multi-Agent Reinforcement Learning. Mathematics, 9 (2021), 23; 3081, 28 doi:10.3390/math9233081
 8. Kušić, Krešimir; Ivanjko, Edouard; Gregurić, Martin; Miletić, Mladen, An Overview of Reinforcement Learning Methods for Variable Speed Limit Control. Applied Sciences, 10 (2020), 14; 4917, 14 doi:10.3390/app10144917

Papers at international scientific conferences

1. Kušić, Krešimir; Schumann, René; Gregurić, Martin; Ivanjko, Edouard; Šoštarić, Marko, Reliable Learning-based Controllers and How Structured Simulation is a Path towards Them Proceedings of the 5th International Conference on Advances in Signal Processing and Artificial Intelligence Yurish, Sergey Y. (ur.). Tenerife, Španjolska, 2023. pp. 268-274
2. Kusic, Kresimir; Schumann, Rene; Ivanjko, Edouard, Building a Motorway Digital Twin in SUMO: Real-Time Simulation of Continuous Data Stream from Traffic Counters. Proceedings of ELMAR-2022 / Muštra, Mario; Zovko-Cihlar, Branka; Vuković, Josip (ur.). Zagreb: IEEE, 2022. pp. 171-176 doi:10.1109/elmar55880.2022.9899796
3. Kušić, Krešimir; Ivanjko, Edouard; Vrbanić, Filip; Gregurić, Martin; Dusparic, Ivana, Dynamic Variable Speed Limit Zones Allocation Using Distributed Multi-Agent Reinforcement Learning. 2021 IEEE International Intelligent Transportation Systems / Conference Indianapolis (IN), Sjedinjene Američke Države: Institute of Electrical and Electronics Engineers (IEEE), 2021. pp. 3238-3245 doi:10.1109/itsc48978.2021.9564739

4. Vrbanić, Filip; Čakija, Dino; Kušić, Krešimir; Ivanjko, Edouard, Traffic Flow Simulators with Connected and Autonomous Vehicles: A Short Review. Transformation of Transportation / Petrović, Marjana; Novačko, Luka (ur.). Cham: Springer International Publishing, 2021. pp. 15-30 doi:10.1007/978-3-030-66464-0_2
5. Vizioli, Helena Tanoue; Kušić, Krešimir; Ivanjko, Edouard; Cunha, André Luiz, A Method to Calibrate Variable Speed Limit Control on High-Truck Share Roads. Smart Innovation, Systems and Technologies book series (SIST, volume 233): Proceedings of the 6th Brazilian Technology Symposium (BTSym'20) / Iano, Y.; Saotome, O.; Kemper, G.; de Seixas, Ana C.M.; de Oliveira, Gabriel G. (ur.). Cham: Springer International Publishing, 2021. pp. 204-211 doi:10.1007/978-3-030-75680-2_24
6. Miletić, Mladen; Kušić, Krešimir; Gregurić, Martin; Koltovska Nečoska, Daniela; Ivanjko, Edouard; Kalinić, Hrvoje, Creating A Data-Set For Sustainable Urban Mobility Analysis: Lessons Learned. ELMAR-2020: 62nd International Symposium: Proceedings / Muštra, Mario; Vuković, Josip; Zovko-Cihlar, Branka (ur.). Zagreb: Croatian Society Electronics in Marine - ELMAR, 2020. pp. 73-78
doi:10.1109/ELMAR49956.2020.9219038
7. Gregurić, Martin; Kušić, Krešimir; Vrbanić, Filip; Ivanjko, Edouard, Variable Speed Limit Control Based on Deep Reinforcement Learning: A Possible Implementation. Proceedings of ELMAR-2020 / Muštra, Mario; Vuković, Vuković; Zovko-Cihlar, Branka (ur.). Zagreb, 2020. pp. 67-72 doi:10.1109/ELMAR49956.2020.9219031
8. Miletić, Mladen; Kušić, Krešimir; Gregurić, Martin; Ivanjko, Edouard, State Complexity Reduction in Reinforcement Learning based Adaptive Traffic Signal Control. Proceedings of ELMAR-2020 / Muštra, Mario; Vuković, Vuković; Zovko-Cihlar, Branka (ur.). Zagreb: Faculty of Electrical Engineering and Computing, 2020. pp. 61-66
doi:10.1109/ELMAR49956.2020.9219024
9. Kušić, Krešimir; Dusparic, Ivana; Guériau, Maxime; Gregurić, Martin; Ivanjko, Edouard, Extended Variable Speed Limit control using Multi-agent Reinforcement Learning. Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC 2020) / Lu, Meng; Wang, Yibing; Barth, Matthew (ur.). Piscataway, NJ USA: IEEE, 2020. 9294639, 8 doi:10.1109/itsc45102.2020.9294639
10. Ivanjko, Edouard; Oliveira Melo, Gabriel; Kušić, Krešimir; Gregurić, Martin, Comparison of Controllers for Variable Speed Limit Using Realistic Traffic Scenarios. Proceedings of 60th International Symposium ELMAR-2018 / Muštra, Mario; Grgić, Mislav; Zovko-Cihlar, Branka; Vitas, Dijana (ur.). Zagreb: Faculty of Electrical Engineering and Computing, University of Zagreb, 2018. pp. 39-42 doi:10.23919/ELMAR.2018.8534633

11. Kušić, Krešimir; Ivanjko, Edouard; Gregurić, Martin, A Comparison of Different State Representations for Reinforcement Learning Based Variable Speed Limit Control. Proceedings of MED-2018 Zadar, Hrvatska: Mediterranean Control Association, 2018. pp. 266-271 doi:10.1109/MED.2018.8442986
12. Gregurić, Martin; Ivanjko, Edouard; Korent, Nino; Kušić, Krešimir, Short Review of Approaches for Variable Speed Limit Control. Proceedings of the International Scientific Conference PERSPECTIVES ON CROATIAN 3PL INDUSTRY IN ACQUIRING INTERNATIONAL CARGO FLOWS ZIRP2016 Ščukanec, Anđelko; Babić, Darko (ur.). Zagreb: Faculty of transport and traffic sciences, 2016. pp. 41-52
13. Kušić, Krešimir; Korent, Nino; Gregurić, Martin; Ivanjko, Edouard, Comparison of Two Controllers for Variable Speed Limit Control. Proceedings ELMAR-2016 / Muštra, Mario; Tralić, Dijana; Zovko-Cihlar, Branka (ur.). Zadar: Faculty of Electrical Engineering and Computing, University of Zagreb, 2016. pp. 101-106
doi:10.1109/ELMAR.2016.7731764

Životopis

Krešimir Kušić rođen je u Zagrebu 30. studenog 1987., gdje je također završio osnovnu i srednju školu. 2012. godine upisuje studij prometnog inženjerstva na Fakultetu prometnih znanosti Sveučilišta u Zagrebu. Titulu magistar inženjer prometa stekao je 2017. godine. Na istom fakultetu 2017. godine upisao je doktorski studijski program (Tehnološki sustavi u prometu i transportu) te se zapošljava kao asistent na Zavodu za inteligentne prometne sustave. Trenutno radi kao nastavnik na sljedećim kolegijima: Računalstvo, Inteligentni transportni sustavi, Sustavi prividne stvarnosti u prometu i Operacijska istraživanja. Njegov znanstveni interes vezan je za optimizaciju prometnih tokova na autocestama i mikrosimulacijsko modeliranje prometa. Godine 2016. dobio je Rektorovu nagradu Sveučilišta u Zagrebu te Dekanovu nagradu kao najbolji student diplomskog studija na studiju: Inteligentni transportni sustavi. Usavršavao se na Trinity College Dublin u Irskoj u sklopu tri međunarodne mobilnosti. Također, dobitnik je Stipendije švicarske vlade za izvrsnost za akademsku godinu 2021./2022., tijekom koje je proveo godinu dana kao znanstveni istraživač u Švicarskoj na Sveučilištu Swiss University of Applied Sciences Western Switzerland.

Popis objavljenih djela

Radovi u časopisima

1. Kušić, Krešimir; Schumann, René; Ivanjko, Edouard, A digital twin in transportation: Real-time synergy of traffic data streams and simulation for virtualizing motorway dynamics. *Advanced Engineering Informatics*, 55 (2023), 101858, 17
doi:10.1016/j.aei.2022.101858
2. Gregurić, Martin; Kušić, Krešimir; Ivanjko, Edouard, Impact of Deep Reinforcement Learning on Variable Speed Limit strategies in connected vehicles environments. *Engineering applications of artificial intelligence*, 112 (2022), 104850, 17
doi:10.1016/j.engappai.2022.104850
3. Ivanjko, Edouard; Kušić, Krešimir; Gregurić, Martin, Simulation analysis of two con-

- trollers for variable speed limit control. Proceedings of the institution of civil engineers-transport, 175 (2022), 7; 413-425 doi:10.1680/jtran.19.00069
4. Miletić, Mladen; Ivanjko, Edouard; Gregurić, Martin; Kušić, Krešimir, A review of reinforcement learning applications in adaptive traffic signal control. IET Intelligent Transport Systems, 16 (2022), 10; 1269-1285 doi:10.1049/itr2.12208
 5. Miloš, Josip; Hršak, Patrik; Topić, Nikola; Jakšić, Leon; Kušić, Krešimir; Vrbanić, Filip; Ivanjko, Edouard, Influence of Spatial Placement of Variable Speed Limit Zones on Urban Motorway Traffic Control. Promet - Traffic&Transportation, 34 (2022), 4; 511-522 doi:10.7307/ptt.v34i4.4073
 6. Vrbanić, Filip; Ivanjko, Edouard; Kušić, Krešimir; Čakija, Dino, Variable Speed Limit and Ramp Metering for Mixed Traffic Flows: A Review and Open Questions. Applied Sciences-Basel, 11 (2021), 6; 2574, 26 doi:10.3390/app11062574
 7. Kušić, Krešimir; Ivanjko, Edouard; Vrbanić, Filip; Gregurić, Martin; Dusparic, Ivana, Spatial-Temporal Traffic Flow Control on Motorways Using Distributed Multi-Agent Reinforcement Learning. Mathematics, 9 (2021), 23; 3081, 28 doi:10.3390/math9233081
 8. Kušić, Krešimir; Ivanjko, Edouard; Gregurić, Martin; Miletić, Mladen, An Overview of Reinforcement Learning Methods for Variable Speed Limit Control. Applied Sciences, 10 (2020), 14; 4917, 14 doi:10.3390/app10144917

Članci objavljeni na konferencijama

1. Kušić, Krešimir; Schumann, René; Gregurić, Martin; Ivanjko, Edouard; Šoštarić, Marko, Reliable Learning-based Controllers and How Structured Simulation is a Path towards Them Proceedings of the 5th International Conference on Advances in Signal Processing and Artificial Intelligence Yurish, Sergey Y. (ur.). Tenerife, Španjolska, 2023. str. 268-274
2. Kusic, Kresimir; Schumann, Rene; Ivanjko, Edouard, Building a Motorway Digital Twin in SUMO: Real-Time Simulation of Continuous Data Stream from Traffic Counters. Proceedings of ELMAR-2022 / Muštra, Mario; Zovko-Cihlar, Branka; Vuković, Josip (ur.). Zagreb: IEEE, 2022. str. 171-176 doi:10.1109/elmar55880.2022.9899796
3. Kušić, Krešimir; Ivanjko, Edouard; Vrbanić, Filip; Gregurić, Martin; Dusparic, Ivana, Dynamic Variable Speed Limit Zones Allocation Using Distributed Multi-Agent Reinforcement Learning. 2021 IEEE International Intelligent Transportation Systems / Conference Indianapolis (IN), Sjedinjene Američke Države: Institute of Electrical and Electronics Engineers (IEEE), 2021. str. 3238-3245 doi:10.1109/itsc48978.2021.9564739

4. Vrbanić, Filip; Čakija, Dino; Kušić, Krešimir; Ivanjko, Edouard, Traffic Flow Simulators with Connected and Autonomous Vehicles: A Short Review. Transformation of Transportation / Petrović, Marjana; Novačko, Luka (ur.). Cham: Springer International Publishing, 2021. str. 15-30 doi:10.1007/978-3-030-66464-0_2
5. Vizioli, Helena Tanoue; Kušić, Krešimir; Ivanjko, Edouard; Cunha, André Luiz, A Method to Calibrate Variable Speed Limit Control on High-Truck Share Roads. Smart Innovation, Systems and Technologies book series (SIST, volume 233): Proceedings of the 6th Brazilian Technology Symposium (BTSym'20) / Iano, Y.; Saotome, O.; Kemper, G.; de Seixas, Ana C.M.; de Oliveira, Gabriel G. (ur.). Cham: Springer International Publishing, 2021. str. 204-211 doi:10.1007/978-3-030-75680-2_24
6. Miletić, Mladen; Kušić, Krešimir; Gregurić, Martin; Koltovska Nečoska, Daniela; Ivanjko, Edouard; Kalinić, Hrvoje, Creating A Data-Set For Sustainable Urban Mobility Analysis: Lessons Learned. ELMAR-2020: 62nd International Symposium: Proceedings / Muštra, Mario; Vuković, Josip; Zovko-Cihlar, Branka (ur.). Zagreb: Croatian Society Electronics in Marine - ELMAR, 2020. str. 73-78
doi:10.1109/ELMAR49956.2020.9219038
7. Gregurić, Martin; Kušić, Krešimir; Vrbanić, Filip; Ivanjko, Edouard, Variable Speed Limit Control Based on Deep Reinforcement Learning: A Possible Implementation. Proceedings of ELMAR-2020 / Muštra, Mario; Vuković, Vuković; Zovko-Cihlar, Branka (ur.). Zagreb, 2020. str. 67-72 doi:10.1109/ELMAR49956.2020.9219031
8. Miletić, Mladen; Kušić, Krešimir; Gregurić, Martin; Ivanjko, Edouard, State Complexity Reduction in Reinforcement Learning based Adaptive Traffic Signal Control. Proceedings of ELMAR-2020 / Muštra, Mario; Vuković, Vuković; Zovko-Cihlar, Branka (ur.). Zagreb: Faculty of Electrical Engineering and Computing, 2020. str. 61-66
doi:10.1109/ELMAR49956.2020.9219024
9. Kušić, Krešimir; Dusparic, Ivana; Guériau, Maxime; Gregurić, Martin; Ivanjko, Edouard, Extended Variable Speed Limit control using Multi-agent Reinforcement Learning. Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC 2020) / Lu, Meng; Wang, Yibing; Barth, Matthew (ur.). Piscataway, NJ USA: IEEE, 2020. 9294639, 8 doi:10.1109/itsc45102.2020.9294639
10. Ivanjko, Edouard; Oliveira Melo, Gabriel; Kušić, Krešimir; Gregurić, Martin, Comparison of Controllers for Variable Speed Limit Using Realistic Traffic Scenarios. Proceedings of 60th International Symposium ELMAR-2018 / Muštra, Mario; Grgić, Mislav; Zovko-Cihlar, Branka; Vitas, Dijana (ur.). Zagreb: Faculty of Electrical Engineering and Computing, University of Zagreb, 2018. str. 39-42 doi:10.23919/ELMAR.2018.8534633

11. Kušić, Krešimir; Ivanjko, Edouard; Gregurić, Martin, A Comparison of Different State Representations for Reinforcement Learning Based Variable Speed Limit Control. Proceedings of MED-2018 Zadar, Hrvatska: Mediterranean Control Association, 2018. str. 266-271 doi:10.1109/MED.2018.8442986
12. Gregurić, Martin; Ivanjko, Edouard; Korent, Nino; Kušić, Krešimir, Short Review of Approaches for Variable Speed Limit Control. Proceedings of the International Scientific Conference PERSPECTIVES ON CROATIAN 3PL INDUSTRY IN ACQUIRING INTERNATIONAL CARGO FLOWS ZIRP2016 Ščukanec, Anđelko; Babić, Darko (ur.). Zagreb: Faculty of transport and traffic sciences, 2016. str. 41-52
13. Kušić, Krešimir; Korent, Nino; Gregurić, Martin; Ivanjko, Edouard, Comparison of Two Controllers for Variable Speed Limit Control. Proceedings ELMAR-2016 / Muštra, Mario; Tralić, Dijana; Zovko-Cihlar, Branka (ur.). Zadar: Faculty of Electrical Engineering and Computing, University of Zagreb, 2016. str. 101-106 doi:10.1109/ELMAR.2016.7731764