KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

TRAFFIC CONTROL ON MOTORWAYS

Promotoren:
Prof. dr. ir. B. De Moor
Prof. dr. ir. B. De Schutter

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Tom BELLEMANS

Mei 2003

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

TRAFFIC CONTROL ON MOTORWAYS

Jury:
Prof. dr. ir. J. Berlamont, voorzitter
Prof. dr. ir. B. De Moor, promotor
Prof. dr. ir. B. De Schutter (TU Delft), promotor
Prof. dr. ir. J.H. van Schuppen (CWI Amsterdam)
Prof. dr. ir. L. Immers
Prof. dr. ir. J. Vandewalle
Prof. dr. ir. R. Boel (RUG, Gent)
Prof. dr. ir. G. Campion (UCL, Louvain-la-Neuve)

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Tom BELLEMANS

U.D.C. 656.021 Mei 2003

D/2003/7515/25

ISBN 90-5682-413-9

Voorwoord

Graag maak ik van deze gelegenheid gebruik om een aantal mensen te bedanken voor hun gewaardeerde bijdrage bij het tot stand komen van dit proefschrift.

Eerst en vooral richt ik een welgemeend woord van dank aan mijn promotoren Prof. Bart De Moor en Prof. Bart De Schutter.

Prof. Bart De Moor wil ik bedanken voor de mogelijkheid die hij me bood om te doctoreren in de onderzoeksgroep SISTA. Door zijn aanstekelijk enthousiasme en het vertrouwen dat hij in me stelde, creëerde hij een sfeer waarin het aangenaam was om te werken. Bovendien wist hij met zijn creatieve ideeën steeds mijn nieuwsgierigheid te prikkelen.

Prof. Bart De Schutter wil ik bedanken voor al de tijd die hij in mijn onderzoek heeft geïnvesteerd ondanks zijn verplichtingen in Delft. De vele zaterdagse onderzoeksbesprekingen in Leuven waren een zeer gewaardeerde bron van inspiratie en motivatie.

De leden van het leescomité, Prof. Jan van Schuppen, Prof. Ben Immers en Prof. Joos Vandewalle ben ik erkentelijk voor hun interesse in mijn onderzoek en voor hun accurate opmerkingen op mijn proefschrift. Deze opmerkingen hebben ontegensprekelijk geleid tot een beter eindresultaat.

Verder bedank ik Prof. René Boel en Prof. Guy Campion voor hun onmiddellijke bereidheid om deel uit te maken van de jury ondanks hun drukke agenda's.

Prof. Jean Berlamont dank ik voor het waarnemen van het voorzitterschap van de jury.

I would like to express my gratitude to Prof. Eric Feron for inviting me to his research group at the Massachusetts Institute of Technology. It was an enriching experience to be able to discuss my research in this exciting environment.

I would also like to thank Prof. Moshe Ben-Akiva for inviting me to present my research at the Intelligent Transportation Systems group at the Massachusetts Institute of Technology.

Ik wil Andreas Hegyi bedanken voor de waardevolle hints die hij gaf. Ook Sven Maerivoet en Steven Logghe, bij wie ik steeds terecht kon voor een verkeerskundige babbel, ben ik dank verschuldigd.

De collega's op SISTA verdienen eveneens een woordje van dank voor de vele interessante discussies tijdens de voorbije jaren.

De DWTC wens ik te danken voor de financiering van mijn onderzoek.

Tot slot bedank ik mijn ouders, Bert, en de rest van mijn familie en vrienden voor hun steun en hun rotsvast vertrouwen dat dit proefschrift tot een goed einde zou worden gebracht. Een heel speciaal woord van dank gaat uit naar Sara, voor al haar lieve zorgen en de vele kleine dingen van elke dag.

Abstract

Due to the ever-increasing demand for road transportation there will be more and more traffic congestion on motorways unless some far-reaching measures are taken. Given the need for solutions in the short term, we discuss ramp metering control of motorway traffic, which can be implemented quickly and relatively cheaply. In this thesis we present a model predictive control (MPC)-based approach to ramp metering control.

First, we present a general overview of the most important building blocks of an MPC-based traffic control system. This overview includes traffic sensors and traffic measurements, traffic models and traffic model classification, and motorway traffic control measures.

We also discuss the METANET traffic flow model, which will be used later on as the prediction model in the MPC framework and to simulate the traffic situation in a real-life case study.

Next, we present the ramp metering controller design starting with the description of the ramp metering concept followed by a discussion of the MPC framework. We apply the MPC framework to the control of ramp metering set-ups. We also discuss the widely used ALINEA ramp metering algorithm to compare the MPC-based ramp metering controller with.

We consider a stretch of the E17 motorway Ghent–Antwerp in Belgium as a case study to investigate the performance of MPC-based ramp metering. Several ramp metering control scenarios for the case study motorway are simulated and evaluated: the no-control scenario, the ALINEA-based ramp metering control scenario, and the MPC-based ramp metering control scenario including both non-coordinated and coordinated ramp metering control.

Finally, we present two extensions to the MPC-based ramp metering framework: the combined MPC-identification approach to ramp metering, which rejects disturbances and the anticipative MPC-based traffic control strategy, which takes the re-routing effects of the drivers due to the control actions into account.

Korte inhoud

Als gevolg van de steeds groeiende vraag naar wegtransport zullen er meer en meer structurele files ontstaan op de autosnelwegen tenzij verstrekkende maatregelen worden genomen. Gelet op de nood aan oplossingen op korte termijn behandelen wij in dit proefschrift de regeling van autosnelwegverkeer met behulp van toeritdosering. Toeritdosering kan snel geïmplementeerd worden en tegen een relatief lage kostprijs. In dit proefschrift stellen wij een aanpak voor de optimale regeling van toeritdosering voor die gebaseerd is op modelgebaseerd voorspellend regelen (MVR).

Eerst geven we een overzicht van de belangrijkste bouwstenen van een MVR-gebaseerd verkeersregelsysteem. In dit overzicht behandelen we verkeersdetectoren en verkeersmetingen, verkeersmodellen en de classificatie van verkeersmodellen, en maatregelen voor de regeling van autosnelwegverkeer.

We bespreken het METANET-verkeersstroommodel dat later zal gebruikt worden als het voorspellend model in de MVR-context en om de verkeerssituatie te simuleren in een gevalsstudie van een bestaande autosnelweg.

Vervolgens behandelen we het ontwerp van een MVR-gebaseerde regelaar voor toeritdosering. We starten met de behandeling van het toeritdoseringsconcept gevolgd door de uiteenzetting van de MVR-theorie. Vervolgens passen we de MVR-aanpak toe op de regeling van toeritdoseringsopstellingen. Om de performantie van de MVR-gebaseerde regeling van toeritdosering te kunnen toetsen, stellen we het veel gebruikte toeritdoseringsalgoritme ALINEA voor.

We gebruiken een stuk van de E17-autosnelweg Gent–Antwerpen als een gevalsstudie om de performantie van de MVR-gebaseerde regelaar voor toeritdosering te onderzoeken. Verschillende regelscenario's voor de E17-autosnelweg werden gesimuleerd en geëvalueerd: een scenario zonder regeling, een scenario met ALINEA-gebaseerde toeritdosering, en een scenario met MVR-gebaseerde regeling van zowel niet-gecoördineerde als gecoördineerde toeritdoseringsopstellingen.

Tenslotte behandelen we twee uitbreidingen van modelgebaseerd voorspellend

regelen: de gecombineerde MVR-identificatie aanpak van toeritdosering die verstoringen onderdrukt, en de anticipatieve MVR-aanpak van toeritdosering die de herrouteringseffecten van de bestuurders in respons op de regelacties in rekening brengt.

Glossary

List of symbols

l	simulation step counter
k	control step counter
m	link index
i	segment index
p	measurement sample counter
ΔT_{sim}	simulation time step
ΔT_{ctrl}	control time step
ΔT_{ident}	time between traffic prediction model updates
$\Delta T_{\text{anticip}}$	time between traffic assignment updates
ΔT_{meas}	traffic measurement time step
$v_{\text{free},m}$	free flow speed in link m
t	continuous time variable
$\rho_{m,i}(l)$	traffic density in segment i of link m at simulation step l
$v_{m,i}(l)$	mean speed in segment i of link m at simulation step l
$q_{m,i}(l)$	traffic flow (intensity) in segment i of link m at simulation step l
n_m	number of lanes in link m
$l_{m,i}$	length of the i-th segment of link m
$q_{\text{in},m,i}(l)$	traffic inflow into segment i of link m at simulation step l
$q_{\text{out},m,i}(l)$	traffic outflow out of segment i of link m at simulation step l
τ_m	time constant of the METANET speed relaxation term
$V[\rho_{m,i}(l)]$	equilibrium speed in segment i of link m at simulation step l as a function of the density $\rho_{m,i}(l)$
α	fundamental diagram fitting parameter
β	fundamental diagram fitting parameter
a_m	fitting parameter equilibrium speed function
ν_m	METANET speed anticipation term parameter
κ_m	METANET speed anticipation term parameter
δ_{m_o}	merging term parameter
L_{m_o}	merging term parameter

$\rho_{\mathrm{crit},m}$	critical density of link m
$q_{\mathrm{on},o}(l)$	service rate of on-ramp o at simulation step l
θ	vector with the model parameters
N_m	highest segment index in link m
ϕ_m	METANET model weaving term parameter
$w_o(l)$	queue length at on-ramp o at simulation step l
$D_o(l)$	traffic demand at on-ramp o at simulation step l
$Q_{\mathrm{cap},o}$	on-ramp capacity of on-ramp o
$\rho_{\mathrm{jam},m}$	jam density of link m
$r_o(k)$	metering rate at on-ramp o at control step k
$q_{\mathrm{max},o}(k)$	maximal number of vehicles allowed to enter the motorway through on-ramp o at control step k
$Q_n(l)$	total traffic flow entering (leaving) node n at simulation step l
I_n	set of all links entering node n
$q_{m,0}(l)$	traffic flow leaving a node through link m
$\beta_n^m(l)$	turning rate (non-destination oriented operation) at node n to link m
O_n	set of links leaving node n
$\rho_{m,i,j}(l)$	partial density of the traffic with destination j in segment i of link m at simulation step l
$\gamma_{m,i,j}(l)$	composition rate for the traffic with destination j in segment i of link m at simulation step l
J_m	set of destinations reachable through link m
$q_{m,i,j}(l)$	partial flow with destination j in segment i of link j at simulation step l
$w_{o,j}(l)$	queue length for the traffic with destination j at on-ramp o at simulation step l
$\gamma_{0,j}(l)$	composition rate of the traffic demand at on-ramp o
$Q_{n,j}(l)$	total traffic with destination j entering node n at simulation step l
$\beta_{n,j}^m(l)$	splitting rate (destination oriented operation), fraction of the total traffic flow $Q_{n,j}(l)$ entering node n and destined for destination j that leaves the node through link m at simulation step l
$\hat{\rho}$	ALINEA controller traffic density setpoint
K_R	ALINEA controller gain
N_{p}	prediction horizon length
N_{c}	control horizon length
α_{queue}	weighting parameter for the TTS in on-ramp queues
α_{var}	weighting parameter for the penalty term on variations of the metering rate
P_{TTS}	TTS performance measure
$\mathcal{I}_{\mathrm{m}}$	set of all index pairs (m,i) corresponding to motorway segments

$\mathcal{I}_{\mathrm{o}}$	set of all on-ramp indices
S_m	set of all segments in link m
$\mathcal{L}_{\mathrm{meas}}$	set of simulation steps l for which measurements are available
c_m	link cost of link m

Acronyms and Abbreviations

ATMS	Advanced Traffic Management System
CPU	Central Processing Unit
CTA	Current Traffic Assignment
ERP	Electronic Road Pricing
ETA	Equilibrium Traffic Assignment
LWR	Lighthill, Whitham and Richards
MPC	Model Predictive Control
MSA	Method of the Successive Averages
TTS	Total Time Spent

Contents

Verkeersregeling op autosnelwegen

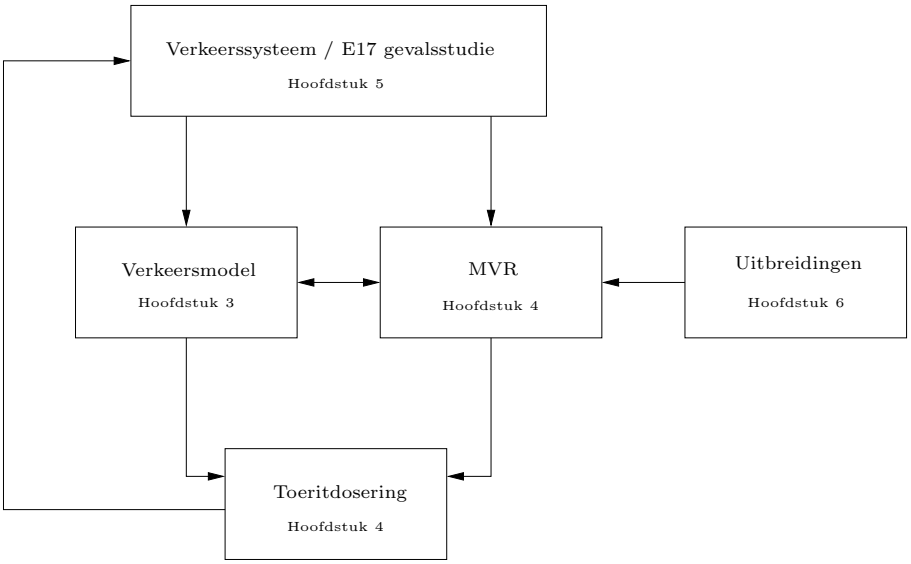
Nederlandse samenvatting

Hoofdstuk 1 Inleiding

Door de sterk groeiende vraag naar transport over de weg zullen steeds vaker files blijven ontstaan tenzij verstrekkende maatregelen worden genomen. Nu reeds wordt tijdens de ochtend- en de avondspits het saturatiepunt van een groot aantal autosnelwegen verspreid over heel de wereld bereikt, met files als gevolg. Deze files kosten de maatschappij handenvol geld en bijgevolg dringt zich een oplossing op korte termijn op.

In dit proefschrift beschouwen we verkeersregeling op autosnelwegen. Uit de verschillende strategieën die kunnen worden toegepast om files te bestrijden zoals daar bijvoorbeeld zijn: het bouwen van extra wegcapaciteit daar waar structurele files ontstaan, het nemen van maatregelen die het vervoer via andere transportmodi stimuleren (boni, ...), het nemen van maatregelen die de vraag naar wegtransport reduceren (belastingen, telewerk, ...), en zelfs het implementeren van intelligente voertuig/snelwegsystemen om de verkeersdoorstroming te verhogen, kiezen wij in dit proefschrift voor maatregelen die leiden tot een efficiënter gebruik van de bestaande infrastructuur. Een efficiënter gebruik van de bestaande infrastructuur kan bereikt worden door maatregelen zoals een verbeterde verkeersregeling met behulp van dynamische verkeersmanagementsystemen, het heffen van tol op bepaalde links, een betere planning van het transport van de industrie, ...

Gelet op de negatieve impact van files op de economie, de levenskwaliteit, de



Figuur 1: Overzicht van de onderwerpen die aan bod komen in dit proefschrift. In dit proefschrift beschouwen we de autosnelweg E17 Gent–Antwerpen als een gevalsstudie. We implementeren een verkeerssimulatiemodel voor deze gevalsstudie en we ontwerpen een toeritdoseringsregelaar gebaseerd op het concept van modelgebaseerd voorspellend regelen (MVR). In Hoofdstuk 6 ontwikkelen we twee uitbreidingen van de MVR-gebaseerde regelaar voor toeritdosering: een eerste uitbreiding moet de invloed van externe invloeden op de performantie van de regelaar onderdrukken terwijl de tweede uitbreiding rekening houdt met de herroutering die optreedt ten gevolge van o.a. de regelacties.

gezondheid, het milieu, ..., dringt zich een oplossing op korte termijn op. Daarom kiezen we in dit proefschrift voor verkeersregeling op snelwegen op basis van dynamische verkeersmanagementsystemen en onderzoeken we de optimale regeling van toeritdoseringsopstellingen. Immers, toeritdosering kan op korte termijn gerealiseerd worden en dit tegen een fractie van de kostprijs van het bouwen van nieuwe snelwegen aangezien toeritdosering gebruik maakt van de bestaande infrastructuur en slechts beperkte bijkomende investeringen vereist zijn. Toeritdosering kan lokaal op een enkele toerit toegepast worden, maar het kan ook ingepast worden in een dynamisch verkeersmanagementsysteem dat instaat voor de optimale verkeersregeling van een heel gebied. In een dergelijk dynamisch verkeersmanagementsysteem wordt gecoördineerde toeritdosering toegepast op meerdere toeritten en wordt toeritdosering ook vaak geïntegreerd met andere dynamische verkeersmaatregelen zoals bijvoorbeeld variabele snelheidslimieten en routegeleiding.

We geven nu een kort overzicht van de verschillende onderwerpen die in dit

proefschrift aan bod komen en lichten hun onderlinge samenhang toe.

In dit proefschrift beschouwen we een stuk van de autosnelweg E17 Gent–Antwerpen, voor het verkeer in de richting van Antwerpen, als een gevalsstudie van een verkeerssysteem. Dit stuk autosnelweg werd gekozen aangezien het een belangrijke verkeersader is die te kampen heeft met structurele congestie tijdens de ochtendspits. Bovendien zijn verkeersmetingen voor deze gevalsstudie beschikbaar. Deze verkeersmetingen vormen een belangrijke link tussen het verkeerssysteem enerzijds en het verkeersmodel en de MVR-gebaseerde regelaar anderzijds, zoals wordt voorgesteld in Figuur 1. Vandaar dat we in Hoofdstuk 2 een overzicht geven van welke verkeersmetingen beschikbaar zijn en hoe deze metingen verwerkt worden.

Aangezien het uit praktische en uit veiligheidsoverwegingen niet mogelijk is om de performantie van de ontwikkelde regelaars op een echte autosnelweg te testen, worden hiervoor in de praktijk verkeersmodellen gebruikt. Om de keuze van het verkeersmodel dat we in dit proefschrift gebruiken te onderbouwen, geven we in Hoofdstuk 2 een kort overzicht van enkele criteria op basis waarvan de verkeersmodellen kunnen worden geclassificeerd. In Hoofdstuk 3 bespreken we het tweede-orde METANET-verkeersstroommodel dat we in Hoofdstuk 5 en 6 gebruiken om de E17-gevalsstudie te simuleren.

In dit proefschrift gebruiken we modelgebaseerd voorspellend regelen (MVR) voor de optimale regeling van toeritdosering. De MVR-gebaseerde regeling is een model gebaseerde techniek die een predictiemodel en metingen van de toestand van het verkeerssysteem gebruikt om de regelsignalen die naar de toeritdoseringsopstelling gestuurd worden, te berekenen (zie Figuur 1). Toeritdosering als maatregel in een verkeersmanagementsysteem tracht de verkeerssituatie op de snelweg zo optimaal mogelijk te houden door slechts druppelsgewijze verkeer via de opritten tot de snelweg toe te laten. Het principe achter toeritdosering wordt in detail besproken in Hoofdstuk 4. In Hoofdstuk 5 illustreren we aan de hand van enkele simulatievoorbeelden voor de E17-gevalsstudie dat MVR-gebaseerde regeling kan toegepast worden voor één enkele toerit maar ook voor de gecoördineerde regeling van meerdere toeritten. Een belangrijk voordeel van MVR-gebaseerde regeling, dat aan de hand van de simulaties in Hoofdstuk 5 geïllustreerd wordt, is het feit dat MVR-gebaseerde regeling rekening kan houden met strikte beperkingen zoals bijvoorbeeld een beperking op de maximale lengte van de wachtrij aan de toeritten. Hoewel we ons in dit proefschrift beperken tot het ontwikkelen van MVR-gebaseerde regeling voor toeritdosering (zie Figuur 1) dient opgemerkt te worden dat MVR-gebaseerde regeling behalve voor toeritdosering ook kan toegepast worden voor andere maatregelen zoals bijvoorbeeld voor dynamische snelheidslimieten.

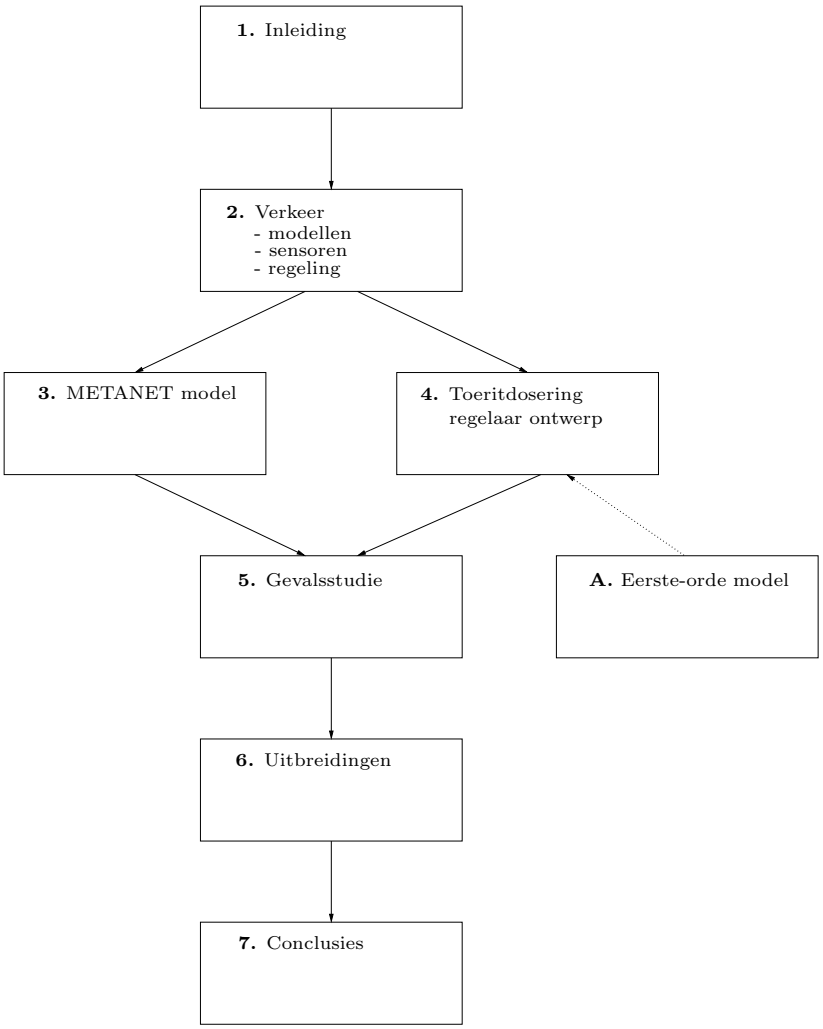
Tot slot werken we in Hoofdstuk 6 nog twee uitbreidingen uit van het MVR-concept zoals schematisch weergegeven wordt in Figuur 1. De eerste uitbreiding bestaat uit een regelmatige heridentificatie van het predictiemodel dat door de

MVR-gebaseerde regelaar gebruikt wordt. Op deze wijze wordt het verschil tussen het gedrag dat gemodelleerd wordt door het predictiemodel en het werkelijke gedrag van het verkeerssysteem zo klein mogelijk gehouden, wat een gunstige invloed heeft op de performantie van de regelaar. De tweede uitbreiding is anticipatieve MVR-gebaseerde toeritdosering. Anticipatieve MVR-gebaseerde toeritdosering houdt rekening met het feit dat de verkeersstromen in een verkeersnetwerk met meerdere routes zich spontaan zullen herorganiseren indien de verkeerstoestand langs bepaalde routes wijzigt ten gevolge van toeritdosering. Anticipatieve MVR-gebaseerde regeling anticipeert op dit gedrag en past de regeling aan zodanig dat toch een optimale toestand wordt bereikt ondanks de herroutering van de verkeersstromen.

De structuur van de hoofdstukken in dit proefschrift wordt in Figuur 2 weergegeven. In hoofdstuk 2 bespreken we verkeersmodellen, sensoren en maatregelen voor verkeersregeling. Op basis van de uiteenzetting van de modellen en hun classificatie in Hoofdstuk 2 wordt gekozen voor het METANET-model dat we in Hoofdstuk 3 bespreken. Uit de lijst met mogelijke maatregelen om het verkeer op autosnelwegen te regelen, kiezen we voor toeritdosering. In Hoofdstuk 4 gaan we dieper in op toeritdosering en ontwikkelen we een MVR-gebaseerde regelaar voor toeritdosering. De theoretische concepten uit Hoofdstukken 3 en 4 worden in Hoofdstuk 5 op de gevalsstudie van de E17-autosnelweg Gent–Antwerpen toegepast. Vervolgens bespreken we in Hoofdstuk 6 twee uitbreidingen van de MVR-gebaseerde regeling voor toeritdosering.

Hoofdstuk 2 Modellen, metingen en actuatoren in een verkeerscontext - Algemeen overzicht

De toepassing van MVR-gebaseerde regeling voor verkeer vereist de keuze van een verkeersmodel dat gebruikt wordt in de regelaar. In dit hoofdstuk bespreken we een aantal criteria volgens dewelke verkeersmodellen kunnen worden geclassificeerd zodat de keuze van een gepast verkeersmodel voor de regelaar vergemakkelijkt wordt. Daarnaast spelen ook metingen een belangrijke rol aangezien deze gebruikt worden als invoer voor de MVR-gebaseerde regelaar, maar ook omdat ze gebruikt worden voor de identificatie van de parameters van het verkeersmodel. In dit proefschrift gaan we dieper in op MVR-gebaseerde regeling van toeritdosering. Voor de volledigheid geven we echter ook een bondig overzicht van een aantal andere maatregelen die eveneens door een MVR-gebaseerde context kunnen worden aangewend.



Figuur 2: Schematische voorstelling van de inhoud van dit proefschrift per hoofdstuk.

Verkeersmodellen

De verkeersmodellen kunnen in verschillende klassen met gelijkaardige eigenschappen worden ingedeeld volgens verschillende criteria.

- **Fysische interpretatie:** Sommige modellen beschrijven de relaties tussen verschillende toestanden volgens fysische wetten die intuïtief interpreteerbaar zijn. Andere modellen gebruiken een meer inductieve benadering waarbij uitgegaan wordt van een generisch model met parameters die bepaald worden op basis van gemeten data. De relaties tussen de toestanden zijn voor deze modellen niet of moeilijk intuïtief interpreteerbaar. Tot slot is er nog een tussenvorm van bovenstaande aanpakken te onderscheiden waarbij vertrokken wordt van een geparameteriseerde intuïtieve beschrijving tussen de toestanden waarbij de parameters aan de hand van metingen worden bepaald.
- **Detailniveau:** Microscopische verkeersmodellen beschrijven het verkeer zeer gedetailleerd, namelijk op het niveau van individuele voertuigen. Macroscopische modellen daarentegen gebruiken geaggregeerde variabelen om de verkeerstoestand te beschrijven. Ook hier kunnen we een tussenvorm onderscheiden, namelijk de mesoscopische modellen die gedetailleerder zijn dan de macroscopische, maar minder gedetailleerd dan de microscopische modellen.
- **Deterministisch of stochastisch:** Stochastische modellen trachten het onregelmatige gedrag van verkeersstromen te beschrijven gebruik makend van ten minste één stochastische variabele. Deze stochastische variabelen worden typisch gekarakteriseerd door hun distributiefunctie. Deterministische verkeersmodellen daarentegen beschrijven de relaties tussen de toestanden van het model volledig op een deterministische wijze.
- **Discreet of continu:** In een verkeersmodel zijn er twee onafhankelijke variabelen, namelijk tijd en plaats. Afhankelijk van het type model kunnen deze onafhankelijke variabelen discreet of continu zijn.

Verkeersmetingen

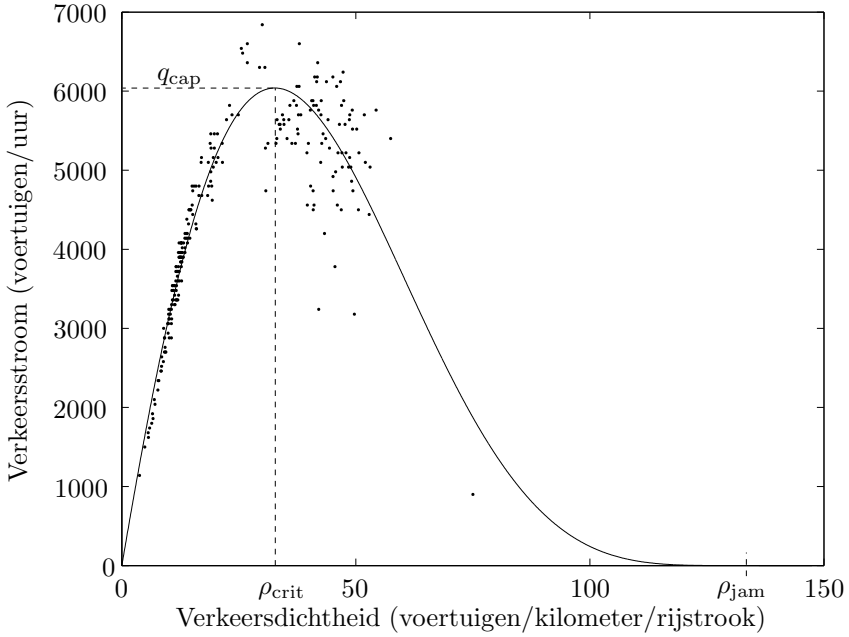
Verschillende variabelen kunnen gebruikt worden om de verkeerstoestand op een autosnelweg te beschrijven. De meest gangbare verkeersmetingen zijn het verkeersdebiet op de autosnelweg (voertuigen/uur), de gemiddelde snelheid op de autosnelweg (kilometer/uur) en de bezettingsgraad van de autosnelweg. De bezettingsgraad van de autosnelweg is de fractie van de totale tijd dat een sensor een voertuig detecteert. De keuze van deze parameters wordt deels bepaald door hun belang in de verkeerskundige theorieën, maar ook door het gemak waarmee ze kunnen gemeten worden met behulp van de huidig gangbare detectoren. Aangezien de bezettingsgraad makkelijk meetbaar is, wordt deze vaak

gemeten in plaats van de verkeersdichtheid (voertuigen/kilometer/rijstrook) die vaak gebruikt wordt in de verkeerskunde. De bezettingsgraad kan echter worden omgerekend naar de dichtheid.

De meest gangbare verkeerssensoren in België zijn inductieve lusdetectoren, camera's en slangdetectoren. De inductieve lusdetectoren zijn het meest in gebruik en bestaan uit een geleidende lus die in het wegdek is ingewerkt. Doorheen deze lus wordt een wisselspanning gestuurd. Een wagen die over de inductieve lusdetector rijdt, wordt gedetecteerd door een verschuiving van de inductantie van de meetlus. Daarnaast worden ook videocamera's gebruikt. Deze filmen de weg en de beelden worden gestuurd naar een algoritme dat de voertuigen telt, hun snelheid berekent en de voertuigen classificeert (personenwagen, vrachtwagen, ...). Tenslotte worden ook slangdetectoren gebruikt. Deze bestaan uit een rubber slang die over de weg wordt gelegd. De voertuigen worden gedetecteerd door de schokgolven die ze in de slang veroorzaken terwijl ze erover rijden.

De gegevens die worden opgemeten door de sensoren langs de autosnelwegen, worden verzameld in een centrale databank. Vooraleer deze gegevens kunnen gebruikt worden in allerlei toepassingen, dient er een voorbewerking op te gebeuren. Door allerlei externe factoren zoals bijvoorbeeld slecht functionerende sensoren, onderhoud van de sensoren, netwerkproblemen, ... kunnen immers gegevens in de databank ontbreken of kunnen onrealistische waarden opgeslagen worden. Aangezien het voor sommige toepassingen onaanvaardbaar is dat er gedurende een bepaalde periode geen gegevens beschikbaar zijn, dient er een schatting gemaakt te worden van de ontbrekende en de onrealistische gegevens. In het proefschrift stellen we een methode voor om een schatting te maken van onrealistische en/of ontbrekende waarden door interpolatie met behulp van een referentiedag. Deze methode is numeriek efficiënt en kan voor online toepassingen gebruikt worden. De referentiedag is een dag met gelijkaardige eigenschappen als de huidige. De meetwaarden in de referentiedag kunnen worden berekend als het gemiddelde van de meetwaarden over verschillende dagen met gelijkaardige eigenschappen als de huidige. Deze berekening kan offline plaatsvinden. Door de juiste keuze van de horizon waarover de referentiedag wordt berekend, kunnen bepaalde invloeden, zoals variaties ten gevolge van de seizoenen, in rekening gebracht worden in de referentiedag.

Een studie van de verzamelde meetgegevens van een snelweg kan heel wat informatie over de snelweg opleveren. In Figuur 3 worden de metingen van de verkeersdichtheid en het verkeersdebiet voor een ochtendspits op de E17-autosnelweg uitgezet. Een dergelijke grafiek wordt in de verkeerskunde een fundamenteel diagram genoemd omdat het voor alle wegen een gelijkaardige vorm heeft. We observeren in Figuur 3 dat voor een lage verkeersdichtheid het debiet stijgt met toenemende verkeersdichtheid. Er is echter een bepaalde verkeersdichtheid, de kritische dichtheid ρ_{crit}, waarbij het verkeersdebiet maximaal is. Dit maximale debiet noemen we de capaciteit van de snelweg q_{cap}. Eens



Figuur 3: Een plot van het verkeersdebiet en de verkeersdichtheid op de E17-autosnelweg in België (punten) gemeten op woensdag 23 februari 2000 van 5 uur tot 11 uur. Het getoonde verkeersdebiet is het debiet van de drie rijstroken van de snelweg samen. Een empirische verkeersdebiet-verkeersdichtheidsrelatie werd op de data gefit en voorgesteld in de volle lijn.

de dichtheid op de autosnelweg groter is dan de kritische dichtheid ρ_{crit}, neemt het verkeerdebiet af met toenemende verkeersdichtheid en ontstaat er congestie op de autosnelweg. Indien de verkeersdichtheid steeds verder blijft stijgen, wordt het verkeersdebiet uiteindelijk gelijk aan 0. De verkeersdichtheid waarbij het verkeersdebiet gelijk wordt aan 0, noemen we de opstoppingsdichtheid ρ_{jam}. Indien de opstoppingsdichtheid ρ_{jam} bereikt wordt, komt het verkeer op de snelweg tot stilstand. De opstoppingsdichtheid ρ_{jam} is de theoretisch maximaal bereikbare dichtheid op de snelweg. Naast het debiets-dichtheidsdiagram kunnen nog twee andere fundamentele diagrammen getekend worden, namelijk een snelheids-dichtheidsdiagram en een snelheids-debietdiagram. Deze twee diagrammen worden in Sectie 2.2.4 van het proefschrift besproken.

Verkeersmaatregelen

Ter afsluiting van dit hoofdstuk geven we een kort overzicht van de beschikbare maatregelen voor de regeling van verkeer op snelwegen. In regeltechnische termen kunnen deze verkeersmaatregelen gezien worden als de actuatoren die door de verkeersregelaar worden aangestuurd. Volgende alternatieve maatregelen zijn beschikbaar:

- extra rijstroken die gedurende pieken in de verkeersvraag kunnen worden opengesteld,
- extra rijstroken die voor een van de twee rijrichtingen kunnen worden opengesteld al naargelang de verkeersvraag op dat moment,
- snelheidsharmonisatie waarbij het verschil in snelheid tussen de voertuigen wordt beperkt om schokgolven te vermijden,
- intelligente snelheidsaanpassing waarbij de bestuurders een maximale snelheid krijgen opgelegd die kan afhangen van de verkeerssituatie, het weer, ...,
- elektronische tolheffing om overbelasting van bepaalde routes tegen te gaan,
- herroutering van bepaalde voertuigstromen om een optimaal gebruik van de alternatieve routes te bewerkstelligen, en
- toeritdosering om te voorkomen dat de verkeersvraag aan een toerit leidt tot congestie op een snelweg.

Hoofdstuk 3 Het METANET-model

In dit hoofdstuk bespreken we het METANET-verkeersstroommodel dat we gebruiken voor de simulaties in de Hoofdstukken 5 en 6. Het METANET-model bestaat in niet-bestemmingsgerichte vorm en in bestemmingsgerichte vorm. Het METANET-model is een macroscopisch, deterministisch en discreet verkeersstroommodel dat werd ontwikkeld door Papageorgiou [96, 110] gebaseerd op het tweede-orde verkeersstroommodel voorgesteld door Payne [117] en Whitham [150]. Het verkeersstroommodel voorgesteld door Payne en Whitham is geïnspireerd op de analogieën tussen verkeersstromen en stromen van vloeistoffen.

In de literatuur zijn zowel voorstanders [5,32,107] als tegenstanders [36,59] van het tweede-orde verkeersstroommodel van Payne en Whitham te vinden. In het proefschrift gaan we daar in Sectie 3.1 dieper op in. Vooraleer van start te gaan

met de bespreking van het METANET-model vatten we eerst de belangrijkste voor- en nadelen van het METANET-model samen. De belangrijkste voordelen van het METANET-model zijn:

- het METANET-model heeft een relatief lage rekencomplexiteit wat het geschikt maakt voor de meeste online toepassingen,
- het METANET-model is geschikt voor de simulatie van alle verkeerstoestanden (file, grote dichtheid, en lage dichtheid),
- de rekencomplexiteit van het METANET-model is onafhankelijk van de verkeerssituatie die gesimuleerd wordt,
- het METANET-model voorziet in een term die vertragingen ten gevolge van wevend verkeer nabij het verdwijnen van een rijstrook in rekening brengt,
- het METANET-model voorziet in een term die vertragingen ten gevolge van invoegend verkeer aan een toerit in rekening brengt,
- in bestemmingsgerichte mode kan het METANET-model de routering van substromen voertuigen tussen herkomsten en bestemmingen expliciet in rekening brengen.

Naast voordelen heeft het METANET-model ook een aantal nadelen:

- het voorziet slechts in de simulatie van één voertuigklasse en bijgevolg worden alle voertuigen op dezelfde manier behandeld,
- aangezien het METANET-model een deterministisch model is, kunnen er geen stochastische effecten van de voertuigstromen in rekening gebracht worden,
- het METANET-model maakt gebruik van geaggregeerde variabelen, en
- er worden geen effecten gemodelleerd die zich afspelen tussen meerdere rijstroken, uitgezonderd de effecten van invoegend verkeer nabij een toerit en de effecten van wevend verkeer nabij het verdwijnen van een rijstrook.

Op basis van een afweging van de voor- en de nadelen van het METANET-model werd beslist dit model te gebruiken voor de simulaties in Hoofdstukken 5 en 6. Voornamelijk de relatief lage rekencomplexiteit en de geschiktheid van het METANET-model voor de simulatie van alle verkeerstoestanden waren belangrijke argumenten om het METANET-model te gebruiken. In de context van de anticipatieve MVR-gebaseerde regelaar die in Hoofdstuk 6 wordt

voorgesteld, speelt de routering van verkeer een belangrijke rol en is de beschikbaarheid van een bestemmingsgerichte versie van het METANET-model een belangrijke troef.

Het METANET-model is geschikt voor de modellering en de simulatie van verkeersnetwerken. Een verkeersnetwerk kan worden voorgesteld als een grafe bestaande uit links en knopen. Het METANET-model beschrijft het gedrag van het verkeer in de links en in de knopen met behulp van linkvergelijkingen en knooppuntsvergelijkingen.

Linkvergelijkingen

Het METANET-model is een discreet model waarin de autosnelweg wordt onderverdeeld in segmenten met een typische lengte van 500 m. De tijd is eveneens een discrete variabele in het METANET-model en een typische waarde voor de tijdsstap is 15 s.

In wat volgt geven we een kwalitatieve uiteenzetting van de linkvergelijkingen van het METANET-model. Voor een gedetailleerde beschrijving van de linkvergelijkingen verwijzen we naar Hoofdstuk 3 in het proefschrift. De toestand van een autosnelwegsegment wordt in het METANET-model door volgende toestandsvariabelen beschreven:

- de verkeersdichtheid (voertuigen/kilometer/rijstrook),
- de gemiddelde snelheid in het segment (kilometer/uur), en
- het verkeersdebiet (voertuigen/uur).

Slechts twee van de drie toestandsvariabelen zijn onafhankelijk aangezien er een fundamentele relatie bestaat tussen de drie variabelen die stelt dat het verkeersdebiet in een segment gelijk is aan de gemiddelde snelheid in het segment vermenigvuldigd met de verkeersdichtheid in het segment vermenigvuldigd met het aantal rijstroken.

Naast de fundamentele relatie tussen de drie toestandsvariabelen bevat het METANET-model ook een vergelijking die het behoud van voertuigen in een segment beschrijft. Deze vergelijking stelt dat de verandering in de verkeersdichtheid in een segment het gevolg is van de nettoverkeersstroom die het segment instroomt (of uitstroomt).

Tenslotte voorziet het METANET-model ook in een vergelijking voor de gemiddelde snelheid in een segment. Drie fenomenen, elk gemodelleerd met behulp van een aparte term, beïnvloeden de gemiddelde snelheid in een segment: relaxatie, convectie en anticipatie. De relaxatieterm beschrijft dat het verkeer in

een segment zich wenst te gedragen overeenkomstig met het fundamenteel diagram voor de gemiddelde snelheid en de verkeerdichtheid in dat segment. Indien de gemiddelde snelheid in het segment afwijkt van de gemiddelde snelheid aangegeven door het fundamentele diagram, zal de relaxatieterm een corrigerende invloed op de gemiddelde snelheid uitoefenen. De convectieterm beschrijft het fenomeen dat een voertuig dat een segment verlaat en een nieuw segment binnenrijdt, zijn snelheid niet ogenblikkelijk zal aanpassen aan de heersende verkeerssituatie in het nieuwe segment. Hierdoor beïnvloedt het voertuig de gemiddelde snelheid in het nieuwe segment. De convectieterm brengt dit fenomeen in rekening. De anticipatieterm beschrijft de anticipatie van bestuurders op wat er voor hen gebeurt. Indien het segment stroomafwaarts van het huidige segment waarin de bestuurder zich bevindt, gecongesteerd is, dan zal de bestuurder dit zien nog voor hij in het nieuwe segment aankomt. Zo gauw hij die file ziet, zal hij vertragen wat een invloed heeft op de gemiddelde snelheid in het huidige segment. Behalve de termen die relaxatie, convectie en anticipatie beschrijven, voorziet het METANET-model nog twee extra termen die aan de uitdrukking voor de gemiddelde snelheid in een segment kunnen worden toegevoegd: de weaving-term en de merging-term. Een weaving-term modelleert de vertraging die ontstaat in een segment indien er in het stroomafwaartse segment minder rijstroken zijn en het verkeer in de verdwijnende rijstroken zich in de andere verkeersstroom moet weven. De merging-term wordt gebruikt om de vertraging in een segment ten gevolge van invoegend verkeer afkomstig van een toerit te modelleren.

Knooppuntsvergelijkingen

De knooppuntsvergelijkingen beschrijven hoe het verkeer zich gedraagt in de knooppunten in het verkeersnetwerk. De belangrijkste functie van de knooppuntsvergelijkingen is de routering van het verkeer doorheen het netwerk. Voor de andere functies van de knooppuntsvergelijkingen verwijzen we naar Hoofdstuk 3 in het proefschrift.

Een knooppunt in een netwerk heeft geen opslagcapaciteit met als gevolg dat alle verkeer dat langs de inkomende links het knooppunt instroomt het knooppunt ogenblikkelijk terug verlaat langs een van de uitgaande links. De wijze waarop de verkeersstromen in de knooppunten aan de verschillende routes in het netwerk worden toegekend, verschilt naargelang we de niet-bestemmingsgerichte of de bestemmingsgerichte versie van het METANET-model gebruiken.

Het niet-bestemmingsgerichte METANET-model maakt bij de verdeling van het verkeersdebiet dat het knooppunt instroomt over de verschillende uitgaande links, geen onderscheid tussen de substromen met verschillende bestemming. De toekenning aan de uitgaande links in een knooppunt van een substroom met een bepaalde bestemming is dus onafhankelijk van de bestemming en voor alle bestemmingen identiek. Voor elk van de uitgaande links in een knooppunt

wordt in het niet-bestemmingsgerichte METANET-model een afslagfractie gedefinieerd. De afslagfractie van een uitgaande link geeft de fractie van de totale instroom in het knooppunt aan die het knooppunt langs deze link verlaat. Het niet-bestemmingsgerichte METANET-model is geschikt voor simulaties waarbij de routering van de voertuigstromen in het netwerk niet in detail wordt beschouwd.

In het bestemmingsgerichte METANET-model wordt wel expliciet met de routering van het verkeer doorheen het netwerk rekening gehouden. Dit wordt gerealiseerd door elke substroom met een bepaalde bestemming een bestemmingsafhankelijke toedeling aan de verschillende uitgaande links te geven. De toedeling van een substroom aan de verschillende uitgaande links wordt gedefinieerd met behulp van splitsingsfracties. Een splitsingsfractie voor een bepaalde link en een bepaalde bestemming geeft de fractie van de substroom met die bepaalde bestemming aan, die het knooppunt langs die link verlaat. De splitsingsfracties kunnen bijgevolg verschillen naargelang de verschillende substromen. Het bestemmingsgerichte METANET-model wordt gebruikt indien de routering van de voertuigstromen expliciet wordt beschouwd zoals bijvoorbeeld het geval is in Hoofdstuk 6.

Tot slot van dit hoofdtuk vermelden we dat de parameters van het METANET-model bepaald dienen te worden op basis van verkeersmetingen. Voor elke link in het METANET-model dienen minstens 6 parameters bepaald te worden. De parameters kunnen bepaald worden door de minimalisatie van een kostcriterium dat het verschil tussen de verkeersmetingen en de simulatieresultaten uitdrukt. De identificatie van het METANET-model wordt in Sectie 3.4 in het proefschrift gedetailleerder besproken.

Hoofdstuk 4 Ontwerp van een toeritdoseringsregelaar

In dit hoofdstuk gaan we van start met de uiteenzetting van het toeritdoseringsconcept. Vervolgens wordt het modelgebaseerd voorspellend regelen (MVR) concept toegelicht en toegepast op de regeling van toeritdoseringsopstellingen.

Toeritdosering

Toeritdosering bestaat uit het gecontroleerd toelaten van voertuigen tot een autosnelweg. Dit wordt in de praktijk gerealiseerd door het plaatsen van een verkeerslicht aan de oprit van de autosnelweg waarbij er doorgaans slechts één voertuig tot de snelweg wordt toegelaten per groene fase. Door de lengte van de rode fase aan te passen kan de stroom voertuigen die tot de snelweg wordt toe-

gelaten (de doseerintensiteit), gevarieerd worden. In dit proefschrift gebruiken we de doseersnelheid als regelparameter voor de toeritdoseringsopstellingen. De doseersnelheid wordt gedefinieerd als de verhouding van de doseerintensiteit en de maximale capaciteit van de toerit.

Het idee waarop toeritdosering gebaseerd is, volgt uit de vaststelling dat verkeer op een snelweg zich gedraagt volgens de fundamentele diagrammen. Uit het fundamenteel diagram voorgesteld in Figuur 3 kunnen we afleiden dat er een dichtheid bestaat, de kritische dichtheid ρ_{crit}, waarvoor het voertuigdebiet op de snelweg maximaal is. Indien de verkeersdichtheid groter wordt dan de kritische dichtheid, daalt het verkeersdebiet en ontstaat er congestie. Door toeritdosering toe te passen trachten we te voorkomen dat de verkeersdichtheid op de snelweg groter wordt dan de kritische dichtheid ρ_{crit} om een optimaal voertuigdebiet op de snelweg te realiseren (zie Figuur 3). Indien de verkeersvraag aan de toerit groter is dan de doseerintensiteit, ontstaat er een wachtrij aan het verkeerslicht op de toerit.

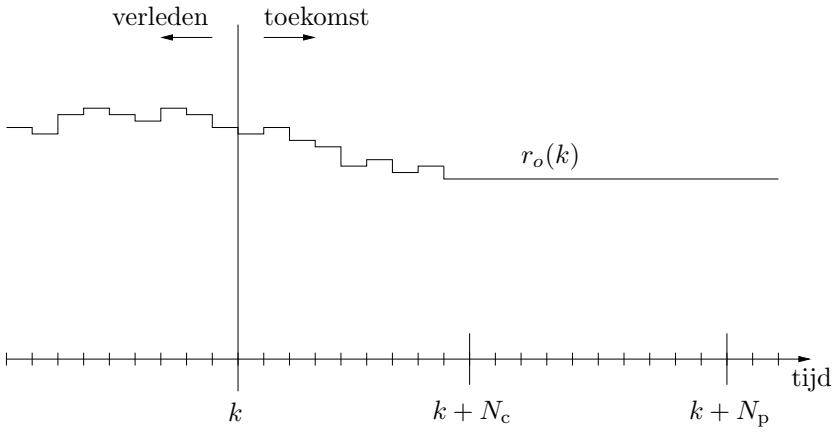
Bij het bepalen van de doseersnelheid voor een toerit dient een afweging gemaakt te worden tussen meerdere conflicterende objectieven: enerzijds wensen we de doorstroming op de autosnelweg zo optimaal mogelijk te houden door de doseersnelheid tijdens congestie op de autosnelweg zo laag mogelijk te kiezen, maar anderzijds mag de doseersnelheid niet te laag zijn omdat dan de wachtrijlengte en ook de wachttijd voor bestuurders in de wachtrij te groot wordt. Andere objectieven die kunnen beschouwd worden bij het bepalen van de optimale doseersnelheid voor toeritdosering, zijn bijvoorbeeld de maximalisatie van de gemiddelde snelheid van de voertuigen, het reduceren van schokgolven op de snelweg uit veiligheidsoogpunt, en het minimaliseren van de totaal gespendeerde tijd die door alle voertuigen in het netwerk (op de autosnelweg en in de wachtrij) wordt gespendeerd.

Er bestaan verschillende algoritmen om de doseersnelheid voor een toeritdoseringsinstallatie te bepalen. Als voorbeeld van een veel gebruikt toeritdoseringsalgoritme vermelden we het ALINEA-algoritme [112] dat gebaseerd is op PID-regeling. Wij zullen het ALINEA-algoritme in Hoofdstuk 5 gebruiken als referentie om de performantie van een toeritdoseringsregelaar op basis van modelgebaseerd voorspellend regelen mee te vergelijken.

Modelgebaseerd voorspellend regelen

In de rest van dit hoofdstuk beschrijven we een nieuwe methode om de doseersnelheid voor toeritdoseringsopstellingen te bepalen. Deze nieuwe methode is gebaseerd op modelgebaseerd voorspellend regelen (MVR).

De implementatie van een MVR-gebaseerde regelaar vereist de definitie van een predictiehorizon N_{p} zoals wordt geïllustreerd in Figuur 4. Tijdens regel-



Figuur 4: Schematische voorstelling van modelgebaseerd voorspellend regelen (MVR). Op regelstap k worden de doseersnelheden voor de hele predictiehorizon N_p berekend. De doseersnelheden variëren alleen gedurende de regelhorizon N_c waarna de doseersnelheid constant gehouden wordt. Alleen de eerste van de berekende doseersnelheden $r_o(k)$ wordt aan de toeritdoseringsopstelling aangelegd. Daarna worden de predictie- en de regelhorizon één regelstap doorgeschoven in de toekomst en start het gehele proces opnieuw.

stap k worden door middel van de optimalisatie van een doelfunctie over de predictiehorizon de optimale doseersnelheden voor de predictiehorizon bepaald. Tijdens de optimalisatie wordt de waarde van de doelfunctie geëvalueerd op basis van een voorspelling van de toekomstige verkeerstoestanden gedurende de predictiehorizon. Deze voorspelling van de toekomstige verkeerstoestanden in het verkeerssysteem wordt bepaald op basis van een verkeerssimulatiemodel. De begintoestand van het verkeerssimulatiemodel wordt bepaald door metingen van de verkeerstoestand in het verkeersnetwerk bij de start van de predictiehorizon. De vereiste invoer voor het verkeerssimulatiemodel gedurende de predictiehorizon wordt geschat. De keuze van het verkeersmodel dat gebruikt wordt om de evolutie van de toestanden over de predictiehorizon N_p te schatten, wordt bepaald door factoren zoals bijvoorbeeld de gedetailleerdheid die de doelfunctie van het verkeersmodel vereist, de rekencomplexiteit van het verkeersmodel, ... maar wordt niet opgelegd door het MVR-concept. Aangezien MVR-gebaseerde regeling werkt met een rollende horizon, wordt nadat de optimale doseersnelheden voor de predictiehorizon gevonden zijn, enkel de eerste doseersnelheid aan de toeritdoseringsinstallatie aangelegd en wordt de predictiehorizon één regelstap verder in de toekomst geschoven, waarna het hele proces opnieuw start.

De doelfunctie die gebruikt wordt bij de optimalisatie van de doseersnelheden van de toerit, beschrijft het verkeersbeleid dat de MVR-gebaseerde regelaar implementeert. De doelfunctie kan uit verschillende termen bestaan die be-

trekking hebben op sociale, economische, milieugerichte, ... aspecten. De verschillende termen in de doelfunctie kunnen gewogen worden overeenkomstig hun belang in het beleid dat de MVR-gebaseerde toeritdoseringsregelaar moet implementeren. Het beleid dat de MVR-gebaseerde regelaar implementeert, kan dus eenvoudig worden aangepast door de doelfunctie te wijzigen. De MVR-gebaseerde aanpak legt immers geen beperkingen op aan de te gebruiken doelfunctie. Indien we gebruik maken van een optimalisatie-algoritme dat in staat is beperkingen in rekening te brengen, dan kunnen we de beschrijving van het verkeersbeleid in de doelfunctie aanvullen door strikte beperkingen op te leggen aan bepaalde toestanden van het verkeerssysteem. Zo kan bijvoorbeeld een maximale wachtrijlengte aan de toeritten opgelegd worden als een strikte beperking om te voorkomen dat een wachtrij te lang wordt en het onderliggend wegennet blokkeert.

Een doelfunctie die vaak wordt gebruikt in de literatuur, is de totaal gespendeerde tijd (TTS) door alle voertuigen in het verkeersnetwerk (zowel op de snelweg als in de wachtrij aan de toerit). In de simulatievoorbeelden in Hoofdstukken 5 en 6 zullen we een lichtjes gewijzigde vorm van de TTS gebruiken als doelfunctie. Immers, aangezien we een regelsignaal wensen dat een zo zacht mogelijk verloop kent, voegen we aan de doelfunctie in deze hoofdstukken nog een term toe die een kost toekent aan variaties van het regelsignaal (de doseersnelheid). De wegingsfactor voor deze term wordt enerzijds groot genoeg gekozen zodat de ergste oscillaties van de doseersnelheid onderdrukt worden, maar anderzijds klein genoeg opdat de TTS in de doelfunctie de dominante term blijft.

Er bestaat een sterk verband tussen de rekencomplexiteit van een optimalisatieprobleem en het aantal dimensies van de parameter ruimte waarin het optimum gezocht wordt. Om de rekencomplexiteit van de optimalisatie in de MVR-gebaseerde regelaar te reduceren wordt daarom vaak een regelhorizon N_c ($N_c \leq N_p$) gedefinieerd. De doseringssnelheden die geoptimaliseerd worden over de predictiehorizon, mogen alleen variëren tijdens de regelhorizon N_c en na de regelhorizon blijft de doseringssnelheid voor de rest van de predictiehorizon constant (zie Figuur 4). Door de definitie van de regelhorizon kan het aantal te optimaliseren parameters beperkt worden, terwijl de predictiehorizon onveranderd blijft. Dit is belangrijk aangezien de lengte van de predictiehorizon bepaalt hoe ver de regelaar in de toekomst kan 'kijken'. Als vuistregel kan gesteld worden dat de lengte van de predictiehorizon best ongeveer gelijk gekozen wordt aan de reistijd voor een voertuig doorheen het gesimuleerde netwerk.

Een tweede manier waarop de rekencomplexiteit van de MVR-gebaseerde regelaar beperkt kan worden, is door een onderscheid te maken tussen de simulatiestap ΔT_{sim} van het verkeerssimulatiemodel (METANET in dit proefschrift) en de regelstap ΔT_{ctrl} van de regelaar. De grootte van de simulatiestap ΔT_{sim} wordt in het geval van het METANET-model langs boven begrensd door de discretisatie van de snelweg in de ruimte (zie Hoofdstuk 3). Een typische waar-

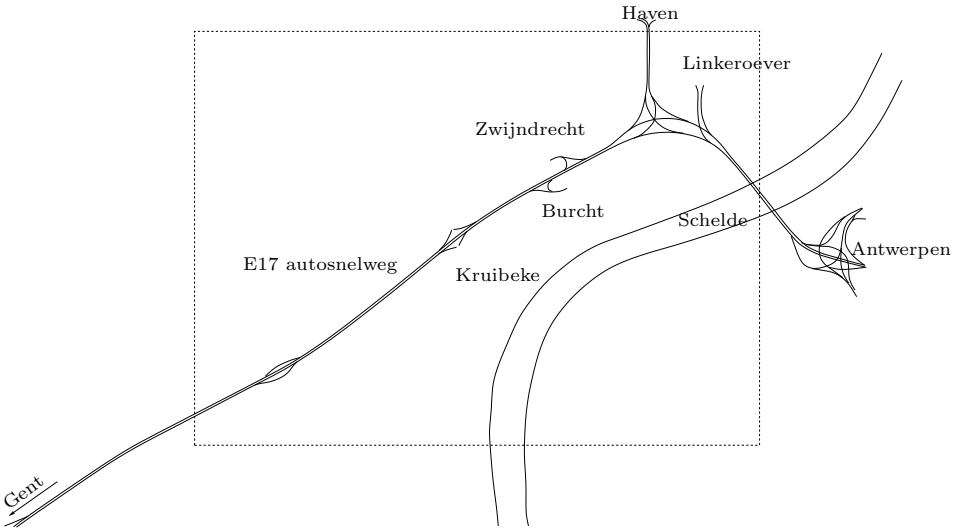
de voor ΔT_{sim} is 10 s. De regelstap ΔT_{ctrl} wordt bepaald door de dynamiek van het verkeerssysteem dat geregeld wordt. Uit de simulatievoorbeelden in Hoofdstukken 5 en 6 zal blijken dat een regelstap $\Delta T_{\text{ctrl}} = 1\,\text{min}$ volstaat. Doordat door de regelstap ΔT_{ctrl} groter is dan de simulatiestap ΔT_{sim} dienen de regelsignalen (de doseersnelheden) minder vaak aangepast te worden dan het geval zou zijn indien we geen onderscheid maken tussen simulatiestap en regelstap.

Tot slot vermelden we de belangrijkste voor- en nadelen van de MVR-gebaseerde regelaar voor toeritdosering. De belangrijkste nadelen van de voorgestelde MVR-gebaseerde aanpak van toeritdosering kunnen we als volgt samenvatten:

- aangezien voor elke regelstap een optimalisatieprobleem, vaak met beperkingen, moet opgelost worden is de MVR-gebaseerde aanpak van toeritdosering vrij rekenintensief,
- de vereiste rekentijd om de optimale doseersnelheden te berekenen varieert,
- de MVR-gebaseerde regelaar vereist een model van het te regelen netwerk, en
- de invoer (bv. verkeersvraag, splitsingsfractie, ...) voor het verkeerssimulatiemodel moet geschat worden voor de hele predictiehorizon.

Tegenover de nadelen van de MVR-aanpak van de regeling van toeritdosering staan echter de volgende voordelen:

- de regelaar kan zich aanpassen aan veranderingen in het verkeerssysteem en in de omgeving van het verkeerssysteem,
- de regelaar kan rekening houden met strikte beperkingen (zoals bijvoorbeeld een strikte beperking op de maximale wachtrijlengte) door gebruik te maken van een optimalisatie-algoritme dat rekening houdt met beperkingen,
- de regelaar is robuust met betrekking tot meetfouten en verstoringen,
- het gewenste gedrag van de regelaar kan makkelijk worden gedefinieerd met behulp van een doelfunctie en een wijziging van het verkeersbeleid kan worden doorgevoerd door de doelfunctie van de regelaar aan te passen, en
- de MVR-gebaseerde regelaar voor toeritdosering kan makkelijk worden uitgebreid tot een regelaar voor de gecoördineerde regeling van meerdere toeritten (zie Hoofdstuk 5) en aangezien de MVR-gebaseerde regeling ook kan gebruikt worden voor andere verkeersmaatregelen, kan de regelaar eveneens worden uitgebreid tot een regelaar voor de geïntegreerde regeling van verschillende verkeersmaatregelen.



Figuur 5: Lay-out van de gevalsstudie: een stuk van 8 kilometer (zie kader) van de E17-autosnelweg Gent–Antwerpen in de rijrichting van Antwerpen.

In het volgende hoofdstuk zullen we de in dit hoofdstuk ontwikkelde MVR-gebaseerde aanpak van de regeling van toeritdosering testen aan de hand van een simulatiemodel van de E17-autosnelweg Gent–Antwerpen.

Hoofdstuk 5 Gevalsstudie

In dit hoofdstuk definiëren we een stuk van de E17-autosnelweg Gent–Antwerpen als een gevalsstudie die we modelleren gebruik makend van het METANET-verkeersstroommodel dat werd voorgesteld in Hoofdstuk 3. We gebruiken het simulatiemodel van de gevalsstudie om de performantie van de MVR-gebaseerde toeritdoseringsregelaar te vergelijken met de performantie van ALINEA-gebaseerde toeritdoseringsregeling. We gaan van start met een beschrijving van de eigenschappen van het beschouwde stuk autosnelweg om vervolgens de volgende situaties met elkaar te vergelijken: de simulatie van de gevalsstudie met ALINEA-gebaseerde toeritdosering en de simulatie van de gevalsstudie met MVR-gebaseerde toeritdosering.

Beschrijving van de gevalsstudie

In dit hoofdstuk definiëren we een stuk van 8 kilometer van de E17-autosnelweg Gent–Antwerpen als een gevalsstudie. In deze gevalsstudie beschouwen we al-

leen het verkeer dat in de richting van Antwerpen rijdt. Een schematische voorstelling van de gevalsstudie is voorgesteld in Figuur 5. We kiezen de E17-autosnelweg als gevalsstudie wegens de aanwezigheid van structurele congestie tijdens de ochtendspits. Bovendien zijn voor de beschouwde gevalsstudie ook voldoende meetgegevens beschikbaar. De beschouwde gevalsstudie bevat 5 toeritten en 4 afritten (zie Figuur 5) waarvan de twee laatste toeritten linkse toeritten zijn, die eerder uitzondelijk zijn in België. Aan het eind van de gevalsstudie bevindt zich de Kennedytunnel onder de Schelde.

Aangezien er in de beschouwde gevalsstudie geen alternatieve routes tussen de herkomsten en de bestemmingen aanwezig zijn, en we bijgevolg geen routering tussen herkomsten en bestemmingen beschouwen, modelleren we de autosnelweg met behulp van het niet-bestemmingsgerichte METANET-model uit Hoofdstuk 3. De parameters van het METANET-model voor de autosnelweg in de gevalsstudie worden bepaald op basis van verkeersmetingen (zie Sectie 5.2 in het proefschrift). Vervolgens wordt in Sectie 5.3 van het proefschrift door de combinatie van verkeersmetigen met de lay-out informatie van de autosnelweg een schatting gemaakt van de verkeersvraag aan de verschillende herkomsten voor een gemiddelde weekdag en van de afslagfracties voor een gemiddelde weekdag. De gegevens voor een gemiddelde weekdag worden bekomen door de verkeersmetingen van dinsdag 22 februari 2000 tot en met deze van vrijdag 25 februari 2000 te combineren.

Aangezien de congestie tijdens de ochtendspits zich localiseert nabij de vierde en de vijfde toerit van de gevalsstudie, zullen we in onze experimenten toeritdosering simuleren voor deze toeritten. We simuleren het belangrijkste deel van de ochtendspits, namelijk de periode van 5 uur tot 10 uur en vergelijken de volgende scenario's:

- simulatie van de gevalsstudie met ALINEA-gebaseerde toeritdosering, en
- simulatie van de gevalsstudie met MVR-gebaseerde toeritdosering.

Om de performantie van de bovenstaande scenario's met elkaar te kunnen vergelijken definiëren we een performantiemaat P_{TTS} die bestaat uit de totale tijd gespendeerd door alle voertuigen in de gevalsstudie gedurende de vijf uur durende simulatie. Hoe kleiner de P_{TTS}, hoe efficiënter de verkeersafwikkeling op de snelweg. Tenslotte leggen we ook een beperking op aan de maximale wachtrijlengte aan de toeritten. Aangezien we moeten voorkomen dat de wachtrij te lang wordt en het onderliggend wegennet blokkeert, leggen we een bovengrens van 100 voertuigen op aan de wachtrijlengte.

Simulatie van de gevalsstudie met ALINEA-gebaseerde toeritdosering

Een ALINEA-gebaseerde regelaar voor toeritdosering is niet in staat om een beperking op de maximale wachtrijlengte in rekening te brengen in zijn regelwet. De enige mogelijkheid om te voorkomen dat de wachtrij aan de toerit ten gevolge van ALINEA-gebaseerde toeritdosering te lang wordt, is om de ALINEA-gebaseerde toeritdosering uit te schakelen zo gauw de wachtrijlengte een bepaalde drempelwaarde overschrijdt. Zo gauw de wachtrijlengte terug kleiner wordt dan de drempelwaarde, wordt de ALINEA-gebaseerde regelaar terug ingeschakeld. Op basis van een aantal experimenten die in Hoofdstuk 5 beschreven worden, bepalen we de optimale waarden voor de parameters van de ALINEA-gebaseerde toeritdoseringsregelaar. Nadat de parameters van de regelaar bepaald zijn, voeren we volgende experimenten uit waarbij we gebruik maken van het METANET-simulatiemodel van de gevalsstudie: toeritdosering op de vierde toerit, toeritdosering op de vijfde toerit, en gelijktijdige toeritdosering op de vierde en op de vijfde toerit. Uit de experimenten blijkt dat ALINEA-gebaseerde toeritdosering op de vierde toerit beter presteert dan op de vijfde toerit. De gelijktijdige toepassing van ALINEA-gebaseerde toeritdosering op de vierde en de vijfde toerit resulteert voor de gesimuleerde ochtendspits in een performantie die lager is (P_{TTS} groter) dan de performantie van ALINEA-gebaseerde toeritdosering op de vierde toerit alleen en in een performantie die hoger is dan de performantie van ALINEA-gebaseerde toeritdosering op de vijfde toerit alleen. Dit is een gevolg van het feit dat ALINEA een lokale regelaar is, waardoor de regelsignalen aan beide toeritten niet op elkaar afgestemd zijn en deze elkaar in het simulatievoorbeeld zelfs tegenwerken.

Ten gevolge van de grote verkeersvraag aan de toeritten tijdens de gesimuleerde ochtendspits groeit de wachtrijlengte aan de toeritten zeer snel, wat uiteindelijk resulteert in een overschrijding van de drempelwaarde en de uitschakeling van de ALINEA-gebaseerde regelaars. We observeren dat de maximale wachtrijlengte overschreden wordt indien we de drempelwaarde waarbij de ALINEA-regelaar wordt uitgeschakeld gelijk kiezen aan de maximale wachtrijlengte. Indien de maximale wachtrijlengte een harde grens is die absoluut niet mag overschreden worden, dan dient de drempelwaarde waarbij de ALINEA-gebaseerde regelaar wordt uitgeschakeld, conservatiever gekozen te worden. We merken op dat, tijdens de periode van grote vraag, oscillaties ontstaan waarbij de ALINEA-gebaseerde regelaar aan- en uitgeschakeld wordt omdat de wachtrijlengte rond de drempelwaarde schommelt. We merken eveneens op dat hoe groter de versterkingsfactor K_R in de ALINEA-regelwet (voor meer informatie over de ALINEA-regelwet verwijzen we naar Sectie 4.3 in het proefschrift), hoe sterker de ALINEA-regelaar reageert op een te hoge verkeersdichtheid op de autosnelweg, maar ook hoe hoger de frequentie van de oscillaties van de verkeersdichtheid en de gemiddelde snelheid op de snelweg. De oscillaties dienen zo veel mogelijk onderdrukt te worden aangezien zij aanleiding kunnen ge-

ven tot gevaarlijke verkeerssituaties. De parameter K_R moet zodanig gekozen worden dat een afweging gemaakt wordt tussen de performantie van de resulterende toeritdoseringsregelaar en de onderdrukking van de oscillaties. Tot slot vermelden we dat uit de simulaties blijkt dat de oscillaties van de verkeersdichteid en de gemiddelde snelheid op de autosnelweg, die veroorzaakt worden door twee onafhankelijke ALINEA-gebaseerde toeritdoseringsregelaars, elkaar kunnen versterken.

Simulatie van de gevalsstudie met MVR-gebaseerde toeritdosering

In deze sectie onderzoeken we met behulp van simulaties de performantie van MVR-gebaseerde toeritdosering voor de gevalsstudie.

Vooraleer we van start kunnen gaan met MVR-gebaseerde toeritdosering, dient beslist te worden welk verkeersbeleid we zullen implementeren. We kiezen een licht gewijzigde vorm van de totaal gespendeerde tijd in het verkeersnetwerk als doelfunctie voor de regelaar. Omdat we oscillaties van de doseersnelheid willen vermijden, voegen we aan de doelfunctie een extra term toe die variaties in het regelsignaal bestraft. Op basis van simulaties werd de weging van deze term zodanig gekozen dat de ergste oscillaties van de doseersnelheid worden onderdrukt, maar zodanig dat de bijdrage van de totaal gespendeerde tijd in de doelfunctie dominant blijft.

Aangezien er een beperking gedefinieerd werd op de wachtrijlengte, gebruikten we in de MVR-gebaseerde regelaar een optimalisatiealgoritme dat deze beperkingen in rekening kan brengen. Om te voorkomen dat de gevonden doseersnelheden overeenkomen met een lokaal minimum van de doelfunctie, werd de optimalisatie voor dit experiment driemaal herhaald met verschillende beginvoorwaarden. Afhankelijk van het gedrag van de doelfunctie kan dit aantal startwaarden verhoogd of verlaagd worden.

De lengte van de predictiehorizon $N_\mathrm{p} = 10\,\mathrm{min}$ werd gekozen op basis van de reistijd doorheen de snelweg onder drukke verkeersomstandigheden. Deze keuze werd met behulp van simulaties van de gevalsstudie geverifieerd. De lengte van de regelhorizon N_c werd enerzijds zo klein mogelijk gekozen om de rekencomplexiteit van de regelaar te beperken, maar anderzijds toch groot genoeg om de performantie van de MVR-gebaseerde regelaar niet te schaden. Een waarde $N_\mathrm{c} = 5\,\mathrm{min}$ bleek aan deze vereisten te voldoen.

Er werden vier scenario's met MVR-gebaseerde toeritdosering tijdens de ochtendspits voor de E17-gevalsstudie onderzocht:

- MVR-gebaseerde toeritdosering op de vierde toerit,
- MVR-gebaseerde toeritdosering op de vijfde toerit,

- niet-gecoördineerde MVR-gebaseerde toeritdosering op de vierde en op de vijfde toerit, en
- gecoördineerde, MVR-gebaseerde toeritdosering op de vierde en de vijfde toerit.

We observeerden dat ook voor MVR-gebaseerde toeritdosering de performantie van toeritdosering op de vierde toerit hoger is (lagere P_{TTS}) dan de performantie van toeritdosering op de vijfde toerit. Dit fenomeen is toe te schrijven aan het verschil in de verkeersvraag voor beide toeritten (zie Figuur 5.6 in het proefschrift) en aan het verschil in de locatie van beide toeritten in de gevalsstudie. De toepassing van MVR-gebaseerde toeritdosering op de vierde toerit leidt tot een relatieve reductie van de P_{TTS} van ongeveer 6 % in vergelijking met het niet geregelde geval.

De performantie van niet-gecoördineerde MVR-gebaseerde toeritdosering op de vierde en op de vijfde toerit was lager dan de performantie van MVR-gebaseerde toeritdosering op de vierde toerit, maar ook lager dan de performantie van MVR-gebaseerde toeritdosering op de vijfde toerit. Door het gebrek aan communicatie tussen beide regelaars beperken beide regelaars de doseersnelheid meer dan nodig is voor het optimaal functioneren van het verkeerssysteem in de gevalsstudie. In de wachtrijen aan de toeritten resulteert dit in een verhoogde wachttijd die niet gecompenseerd wordt door een verhoogd debiet op de autosnelweg.

De gecoördineerde MVR-gebaseerde regeling van toeritdosering op de vierde en de vijfde toerit resulteerde voor de gesimuleerde ochtendspits in de gevalsstudie in dezelfde performantie als MVR-gebaseerde toeritdosering op de vierde toerit alleen. Dit volgt uit het feit dat voor de verkeersvraag tijdens de ochtendspits die we hier simuleren (en die bepaald werd op basis van metingen in het werkelijke systeem) de vierde toerit in staat is om de optimale toestand te realiseren. Bijgevolg dient de toeritdosering aan de vijfde toerit niet in actie te komen tijdens de gecoördineerde toeritdosering.

Gebaseerd op deze simulatieresultaten kunnen we besluiten dat de meest economische oplossing om de congestie tijdens de ochtendspits in de beschouwde gevalsstudie aan te pakken, de implementatie van MVR-gebaseerde toeritdosering op de vierde toerit is. Indien de verkeersvraag in de toekomst groter wordt en/of indien de beperking op de wachtrij strikter wordt, dan dienen bovenstaande simulaties herhaald te worden om na te gaan of de gecoödineerde MVR-gebaseerde regeling van toeritdosering op de vierde en de vijfde toerit niet tot een verbetering van de performantie kan leiden.

Vergelijking van ALINEA-gebaseerde en MVR-gebaseerde toeritdosering

Tot slot van dit hoofdstuk vergelijken we de ALINEA-gebaseerde en de MVR-gebaseerde toeritdoseringsalgoritmen op basis van de resultaten van de simulaties uit dit hoofdstuk.

- Op basis van een vergelijking van de door de regelaars gerealiseerde performanties P_{TTS} voor een ochtendspits in de E17-gevalsstudie, kunnen we besluiten dat de MVR-gebaseerde toeritdoseringsregelaar leidt tot een meer performant (lagere P_{TTS}) verkeerssysteem dan de ALINEA-gebaseerde regelaar. Dit was het geval voor toeritdosering op de vierde toerit, op de vijfde toerit, en op de vierde en de vijfde toerit samen.

- De performantie van niet-gecoördineerde MVR-gebaseerde toeritdosering op de vierde en op de vijfde toerit was lager dan de performantie van MVR-gebaseerde toeritdosering op de vierde of op de vijfde toerit, maar toch hoger dan de performantie van de onderzochte ALINEA-gebaseerde toeritdoseringsconfiguraties.

- We observeerden dat terwijl de ALINEA-gebaseerde regelaar niet in staat is de harde beperkingen op de wachtrijlengte te respecteren (ondanks het uitschakelen van de regelaar indien de wachtrijlengte te lang wordt), de MVR-gebaseerde toeritdoseringsregelaar de beperkingen steeds respecteert. Om te voorkomen dat de lengte van de wachtrij, veroorzaakt door de ALINEA-gebaseerde regelaar de maximaal toegelaten lengte overschrijdt, dient de drempelwaarde van de wachtrijlengte waarbij de regelaar wordt uitgeschakeld conservatief ingesteld te worden, wat een negatieve invloed heeft op de performantie van de regelaar.

- De doseersnelheden, zoals die berekend worden door de MVR-gebaseerde regelaar kennen een veel regelmatiger verloop in de tijd dan de doseersnelheden berekend door de ALINEA-gebaseerde regelaar.

- De rekencomplexiteit van de ALINEA-gebaseerde regelaar is kleiner dan de rekencomplexiteit van de MVR-gebaseerde regelaar. De berekening van de doseersnelheden die de doelfunctie minimaliseren over de predictiehorizon is de meest rekenintensieve stap in het algoritme van de MVR-gebaseerde regelaar. Echter, door een zorgvuldige keuze van de parameters ΔT_{ctrl}, N_{p} en N_{c} kan deze rekencomplexiteit gevoelig worden gereduceerd.

Hoofdstuk 6 Uitbreiding van modelgebaseerd voorspellend regelen van toeritdosering

In dit hoofdstuk presenteren we twee uitbreidingen voor MVR-gebaseerde regelaars voor toeritdosering: een MVR-identificatie aanpak van toeritdosering en anticipatieve MVR-gebaseerde toeritdosering. Hoewel we ons in dit hoofdstuk toeleggen op de regeling van toeritdosering zijn deze twee uitbreidingen eveneens toepasbaar bij de MVR-gebaseerde regeling van andere verkeersmaatregelen.

De MVR-identificatie aanpak van toeridosering

Een MVR-gebaseerde regelaar gebruikt een model om een voorspelling te maken van de verkeerstoestanden over de predictiehorizon. Aangezien allerlei externe factoren, die niet in het model vervat zitten, de verkeerssituatie op de autosnelweg kunnen beïnvloeden, kan er een significante afwijking ontstaan tussen het werkelijke gedrag van het verkeer op de autosnelweg en het gemodelleerde gedrag.

Het is op basis van de voorspellingen van het verkeersmodel dat het optimalisatie-algoritme in de MVR-gebaseerde regelaar de doseersnelheden bepaalt die moeten leiden tot een zo laag mogelijke waarde van de doelfunctie. Afwijkingen tussen het verkeersgedrag op de snelweg en het gemodelleerde gedrag kunnen leiden tot een reductie van de performantie van de regelaar. Dit illustreren we met een simulatievoorbeeld in Hoofdstuk 5 in het proefschrift. In dit simulatievoorbeeld wordt de performantie van een simulatie, waarbij de parameters van het predictiemodel in de MVR-gebaseerde regelaar opzettelijk werden geperturbeerd, vergeleken met een simulatie waarbij de parameters van het predictiemodel van de MVR-gebaseerde regelaar niet werden geperturbeerd. Er werd een duidelijke reductie in de performantie van de regelaar waargenomen ten gevolge van de geperturbeerde parameters in het predictiemodel.

Daarom stellen we voor om het model van de MVR-gebaseerde regelaar op regelmatige tijdstippen te heridentificeren zodat wijzigingen in het gedrag van het verkeer op de autosnelweg, bijvoorbeeld ten gevolge van de weerssituatie (regen, mist, sneeuw, ...), in rekening worden gebracht in het predictiemodel.

De identificatie van de modelparameters van het METANET-model, die beschreven werd in Hoofdstuk 3, is een rekenintensieve taak. We kunnen bij de MVR-identificatie aanpak van toeritdosering echter een aantal maatregelen treffen om de vereiste rekenkracht verlagen. Om te beginnen kunnen we er van uitgaan dat de effecten die niet in het model vervat zitten op een relatief trage tijdsschaal plaatsvinden vergeleken met de tijdsschaal van de regelaar ΔT_{ctrl}. Bijgevolg kunnen we de tijdsstap tussen twee opeenvolgende heridentificaties

groter kiezen dan de controlestap ($\Delta T_{\text{ident}} > \Delta T_{\text{ctrl}}$), wat extra tijd vrijmaakt voor de identificatie. De rekencomplexiteit van de heridentificatie blijft ook beperkt doordat er in het algemeen een goede startwaarde voor de identificatie beschikbaar is. Inderdaad, aangezien we kunnen veronderstellen dat de niet gemodelleerde externe factoren de modelparameters niet te bruusk zullen wijzigen, zijn de oude parameters een goede startwaarde voor de heridentificatie. Tot slot kunnen we de heridentificatie ook implementeren als een adaptief proces, waarbij op regelmatige tijdstippen een voorafbepaald aantal iteraties van het identificatieproces worden uitgevoerd. Dit heeft als bijkomend voordeel dat de rekentijd die aan de heridentificatie gespendeerd wordt nauwkeuriger kan worden ingeschat.

Anticipatieve MVR-gebaseerde regeling voor toeritdosering

Indien we toeritdosering toepassen op een snelwegennetwerk waar alternatieve routes bestaan tussen de herkomsten en de bestemmingen, dan kunnen de bestuurders beslissen om een alternatieve route te kiezen als respons op de regelacties. Hoewel dit herrouteringsgedrag een sterke invloed kan hebben op de uiteindelijke verkeerssituatie in het verkeersnetwerk en op de performantie van het verkeersnetwerk, wordt in de literatuur bij het ontwerpen van verkeersregelaars niet vaak rekening gehouden met herrouteringseffecten. In deze sectie ontwikkelen we een methode om op een efficiënte manier herrouteringseffecten te berekenen en in een MVR-gebaseerde regelaar te integreren. Op deze manier bekomen we een anticipatieve MVR-gebaseerde regeling voor toeritdosering in verkeersnetwerken.

Wardrop [147] stelde reeds in 1952 dat het verkeer in een verkeersnetwerk zich over de alternatieve routes van het netwerk verdeelt, zodanig dat een individuele bestuurder zijn reiskost niet kan verlagen door een andere route te kiezen. Het evenwicht dat aldus ontstaat in de verdeling van de voertuigstromen over de verschillende routes wordt een gebruikers-optimaal evenwicht genoemd, aangezien het ontstaat indien elke individuele bestuurder zijn routekeuze optimaliseert. In een verkeersnetwerk waarin de voertuigstromen zich in een gebruikers-optimaal evenwicht bevinden, vertegenwoordigen alle gebruikte routes tussen een herkomst-bestemmingspaar dezelfde reiskost en hebben de niet gebruikte alternatieve routes een hogere reiskost.

Twee belangrijke factoren die meespelen in de kost die een bestuurder aan een route toekent zijn de reistijd langsheen de route en de lengte van de route. Het belang dat een individuele bestuurder aan elk van deze factoren toekent varieert van bestuurder tot bestuurder. Wij gaan hier uit van een kost toegekend aan de verschillende routes door een gemiddelde bestuurder. We merken op dat de reistijd langsheen een route omgekeerd evenredig is met de heersende snelheid in de snelwegsegmenten die deel uitmaken van deze route. Aangezien de snelheid

in een snelwegsegment afhankelijk is van het verkeersvolume op de snelweg (zie de bespreking van de fundamentele diagrammen in Hoofdstuk 2 van het proefschrift), wordt de kost die een bestuurder aan een route toekent beïnvloed door het verkeersvolume in de segmenten van de route.

Het evenwicht dat beschreven wordt door Wardrop is de situatie die ontstaat als de verkeersvraag lang genoeg constant blijft om alle overgangsverschijnselen te laten uitsterven. Er bestaan verschillende methoden om dit statische evenwicht beschreven door Wardrop te berekenen [126]. In Hoofdstuk 6 gebruiken we hiervoor de methode van de opeenvolgende gemiddelden [119].

In de praktijk varieert de verkeersvraag echter in de tijd. Hierdoor treden er verschuivingen op in het gebruikers-optimaal evenwicht omdat de bestuurders hun routekeuze aanpassen aan de gewijzigde situatie. De aanpassing van de routekeuze van de bestuurders gebeurt echter niet ogenblikkelijk. Dit kunnen we als volgt verklaren: Tijdens hun reis dooheen het verkeersnetwerk verzamelen de bestuurders informatie over de toestand van het verkeersnetwerk. Op basis van de verzamelde informatie over de verkeerstoestand bepalen de bestuurders hun optimale route. De informatie over de verkeerstoestand die de bestuurders op dit ogenblik hebben is dus gebaseerd op informatie uit het nabije verleden. Dit proces kan als volgt gemodelleerd worden: De verkeerstoestand in het nabije verleden volgt uit de verkeersvraag en de doseersnelheden uit het verleden. Het feit dat het verzamelen van informatie van de verkeerstoestand een tijd duurt, modelleren we door de verkeersvraag en de doseersnelheid uit te middelen over een periode τ_{info} in het nabije verleden. Door de verkeersvraag en de doseersnelheid over het tijdsinterval $[t - \tau_{\text{info}}, t]$ uit te middelen, bekomen we informatie over de verkeerstoestand zoals die in het nabije verleden werd ervaren door de bestuurders en op basis waarvan de bestuurders nu hun routekeuze maken. De waarde van de parameter τ_{info} hangt af van allerlei factoren zoals bijvoorbeeld de dimensies van het autosnelwegnetwerk, de topologie van het netwerk, de beschikbaarheid van informatie (bijvoorbeeld verkeersinformatie), ...

De bestuurders gebruiken de verzamelde informatie over de verkeerstoestand bij het maken van hun routekeuze. Gebruik makend van een statische verkeerstoedelingsmethode, bijvoorbeeld de methode van de opeenvolgende gemiddelden, kunnen we de everwichtsverkeerstoedeling (EVT) berekenen die overeenkomt met de huidige door de bestuurders ervaren toestand van het autosnelwegennetwerk. Aangezien we voor het bepalen van de ervaren toestand van het netwerk het gemiddelde hebben genomen van de verkeersvraag over de horizon τ_{info} zal een statische toedelingsmethode goede resultaten leveren.

Het duurt een tijd vooraleer het verkeer zich daadwerkelijk zal reorganiseren volgens de berekende EVT. Dit omwille van het feit dat niet alle bestuurders ogenblikkelijk zullen beslissen om de nieuwe route die overeenkomt met de EVT te gebruiken. De overgang van de huidige verkeerstoedeling (HVT)

naar de EVT modelleren we als een exponentieel verloop. De waarde van de tijdsconstante τ_{evol} dient zo bepaald te worden dat de evolutie van de HVT naar de EVT in een realistisch tijdsbestek plaatsvindt.

Bovenstaand model voor de herroutering van de verkeersstromen in een verkeersnetwerk kan gecombineerd worden met MVR-gebaseerde toeritdosering wat leidt tot anticipatieve MVR-gebaseerde toeritdosering. Anticipatieve MVR-gebaseerde toeritdosering is in staat om te anticiperen op de herrouteringseffecten in het netwerk.

De MVR-gebaseerde toeritdoseringsregelaar berekent de optimale doseersnelheid gebruik makend van een predictiemodel. In dit proefschrift gebruiken we hiervoor het METANET-verkeersstroommodel. Aangezien we hier routering expliciet in rekening brengen, kiezen we voor het bestemmingsgerichte METANET-model. De invoer van het METANET-model voor de simulatie van de verkeerstoestand over de predictiehorizon bestaat uit een schatting van de verkeersvraag en van de splitsingsfracties. In de MVR-gebaseerde toeritdoseringsregelaar worden de splitsingsfracties die aan het METANET-predictiemodel worden aangelegd berekend op basis van de voorspelde exponentiële evolutie van de HVT naar de EVT. Op deze manier worden de doseersnelheden geoptimaliseerd rekening houdend met de herrouteringseffecten in het verkeersnetwerk. Aangezien de herroutering van de verkeersstromen in het netwerk typisch plaatsvindt op een veel tragere tijdsschaal dan de effecten die optreden nabij de toeritten, kan de verkeerstoedeling aan een lager tempo worden herberekend dan de doseersnelheid (bijvoorbeeld elke $\Delta T_{\text{anticip}} = 15\,\text{min}$). In Sectie 6.2.4 in het proefschrift presenteren we een simulatievoorbeeld van anticipatieve MVR-gebaseerde toeritdosering waar we toeritdosering implementeren tijdens een ochtendspits en waar we rekening houden met de herroutering van de verkeersstromen veroorzaakt door de wijzigende verkeersvraag en door wegwerkzaamheden.

Tot slot vermelden we de combinatie van beide uitbreidingen in de gecombineerde, anticipatieve MVR-identificatie aanpak van toeritdosering welke in staat is om het performantieverlies ten gevolge van traag variërende, niet gemodelleerde verstoringen te beperken en die ook de herrouteringseffecten van de verkeersstromen in rekening kan brengen bij het bepalen van de optimale doseersnelheden.

Hoofdstuk 7 Besluiten en suggesties voor verder onderzoek

Besluiten

De voornaamste doelstelling van ons onderzoek was na te gaan hoe de toepassing van regeltechniek kan bijdragen tot een efficiëntere verkeersafwikkeling op autosnelwegen. In het kader van dit onderzoek kwamen onderstaande onderwerpen aan bod.

In Hoofdstuk 2 behandelden we de beschikbare verkeersmetingen evenals de meest gangbare sensor technologieën. We ontwikkelden een efficiënt schattingsalgoritme voor ontbrekende of corrupte metingen dat ook geschikt is voor de schatting van metingen over langere periodes. Daarnaast selecteerden we toeritdosering uit de in dit hoofdstuk behandelde verkeersmaatregelen die kunnen toegepast worden op autosnelwegen. De hoofdredenen voor onze keuze om toeritdosering te behandelen in dit proefschrift zijn de afdwingbaarheid, de relatief lage kostprijs, en de snelle implementeerbaarheid van toeritdosering.

Aangezien er uit praktische en uit veiligheidsoverwegingen geen experimenten kunnen gedaan worden op echte autosnelwegen, besloten we in ons onderzoek gebruik te maken van een verkeerssimulatiemodel. In Hoofdstuk 2 werd een kort overzicht gegeven van een aantal classificatiecriteria voor verkeersmodellen. We kozen voor het macroscopisch METANET-verkeersstroommodel dat we in Hoofdstuk 3 gedetailleerd bespraken.

Aangezien we ons onderzoek zo realistisch mogelijk wilden houden werd in Hoofdstuk 5 een stuk van de E17-autosnelweg Gent–Antwerpen als gevalsstudie gedefinieerd. Er werd een verkeerssimulatiemodel voor de gevalsstudie geïmplementeerd op basis van het METANET-verkeersstroommodel. De parameters van het model werden geïdentificeerd op basis van verkeersmetingen op de E17-autosnelweg. De verkeersvraag die als input aan het simulatiemodel wordt aangelegd, werd eveneens op basis van verkeersmetingen geschat.

In Hoofdstuk 4 presenteerden we het toeritdoseringsconcept en bespraken we ALINEA, een veel gebruikt lokaal toeritdoseringsalgoritme. We bespraken het concept van modelgebaseerd voorspellend regelen (MVR) en we pasten dit concept toe op toeritsdosering. Een MVR-gebaseerde regelaar voor toeritdosering is, in tegenstelling tot een ALINEA-gebaseerde regelaar, een niet lokale regelaar die zich bovendien goed leent tot de gecoördineerde regeling van meerdere toeritten. Door een onderscheid te maken tussen de simulatiestap van het verkeerssimulatiemodel en de regelstap van de toeritdoseringsregelaar waren we in staat om de rekencomplexiteit van de regelaar de reduceren.

In Hoofdstuk 5 werd op basis van simulaties van de gevalsstudie de performantie van MVR-gebaseerde toeritdosering vergeleken met de performantie van

ALINEA-gebaseerde toeritdosering. Hierbij onderzochten we zowel geïsoleerde MVR-gebaseerde toeritdosering als gecoördineerde MVR-gebaseerde toeritdosering. We stelden vast dat voor de simulatie van de gevalsstudie MVR-gebaseerde toeritdosering significant beter presteert dan de ALINEA-gebaseerde toeritdosering. MVR-gebaseerde toeritdosering kon de totale gespendeerde tijd in het verkeersnetwerk sterker reduceren dan de ALINEA-gebaseerde regelaar. Bovendien was de MVR-gebaseerde toeritdoseringsregelaar in staat de strikte beperking op de maximale wachtrijlengte steeds te respecteren en kende de evolutie van de doseersnelheden berekend door de MVR-gebaseerde regelaar een zacht verloop in de tijd. De ALINEA-gebaseerde toeritdoseringsregelaar daarentegen leidde tot een wachtrij die de maximale lengte overschreed en de berekende doseringssnelheid kende een grillig verloop in de tijd.

Op basis van deze simulaties kunnen we besluiten dat de toepassing van MVR-gebaseerde toeritdosering op de E17-gevalsstudie leidt tot een significante verbetering van de verkeersafwikkeling.

Tot slot ontwikkelden we in Hoofdstuk 6 twee uitbreidingen van het MVR-concept: de MVR-identificatie aanpak van toeritdosering en de anticipatieve MVR-gebaseerde toeritdosering. De MVR-identificatie aanpak van toeritdosering voorkomt door een heridentificatie op regelmatige tijdstippen, dat ten gevolge van niet gemodelleerde externe invloeden een te grote afwijking ontstaat tussen het gedrag van het verkeer op de autosnelweg en het door het predictiemodel voorspelde gedrag. De anticipatieve MVR-gebaseerde regelaar voor toeritdosering is in staat om herrouteringseffecten in rekening te brengen bij het berekenen van de doseersnelheid. Op deze manier kunnen we voorkomen dat ongewenste effecten optreden ten gevolge van de combinatie van de toeritdoseringsregelaar met de herrouteringseffecten in het verkeersnetwerk.

Verder onderzoek

Tot slot suggereren we enkele onderwerpen voor verder onderzoek en vermelden we enkele open problemen.

In dit proefschrift hebben we ons beperkt tot toeritdosering. De MVR-gebaseerde regelaar is echter ook in staat andere maatregelen aan te sturen. Daar sommige van deze maatregelen een eerder indicatief karakter hebben (bijvoorbeeld route adviezen) dient rekening gehouden te worden met de opvolgingsgraad van de maatregelen. Andere maatregelen, zoals bijvoorbeeld het openen van extra rijstroken, leiden tot geheeltallige optimalisatieproblemen die zeer rekenintensief zijn. Deze maatregelen dienen verder onderzocht te worden in de context van MVR-gebaseerde regeling. Daarnaast is ook de performantie van de geïntegreerde regeling van deze maatregelen een interessant onderwerp voor verder onderzoek.

Het verschil in performantie tussen MVR-gebaseerde toeritdosering en ALINEA-

gebaseerde toeritdosering dient in een werkelijke situatie gevalideerd te worden. Als mogelijke tussenstap kan eventueel gebruik gemaakt worden van de gedetailleerde microscopische simulatiemodellen om de performantie van beide regelaars te vergelijken.

Aangezien in de literatuur het optreden van instabiliteiten wordt gerapporteerd indien verkeerstoedeling en verkeersregeling worden gecombineerd, dient de stabiliteit van de anticipatieve MVR-gebaseerde regelaar nader onderzocht te worden. Doordat wij het verleden van de verkeerstoestand in de anticipatieve MVR-gebaseerde regeling in rekening brengen, evenals de vertraging in de routekeuze van de bestuurders denken wij deze instabiliteiten te kunnen vermijden. Dit vergt verder onderzoek.

Voor zeer grote verkeersnetwerken met veel geregelde toeritten wordt de rekencomplexiteit van MVR-gebaseerde regeling te groot voor de huidige generatie computers. Hiërarchische regeling en multi-agent regeling kunnen als mogelijke alternatieven onderzocht te worden.

Een belangrijk onderwerp voor verder onderzoek is de robuustheid en de betrouwbaarheid van geregelde verkeersnetwerken. Naarmate een regelaar instaat voor grotere netwerken worden de robuustheid en de betrouwbaarheid van de regelaar belangrijker. De robuustheid van de MVR-gebaseerde regelaars tegen onverwachte gebeurtenissen en verstoringen moet nog verder worden onderzocht.

Chapter 1

Introduction

1.1 Problem statement

The ever-increasing need for road transportation will increasingly cause traffic congestion unless some far-reaching measures are taken in the near future. During rush hours every morning and every evening, the saturation point of numerous motorways around the world is reached, resulting in traffic jams. These traffic jams induce costs to the society at several levels.

As a response to the growing internationalization of the economy, the demand for mobility of people and goods increases (see Figure 1.1). Companies tend to locate their offices and production units near accesses to motorways in order to be able to quickly transport their goods and in order to be easily accessible for their employees and their clients. Delays caused by congestion result in a huge loss of productivity, not only due to the unproductive hours of the commuters but also due to late deliveries of supplies or services. Since storage is costly, more and more companies apply 'just in time' transport strategies. This evolution requires a smooth flow of goods and the predictability of travel times. The delays caused by congestion force these companies to keep larger and thus more expensive stocks in order to guarantee the reliability of the production process.

While the unproductive hours spent in congestion cost money to the companies, the government, the hospitals, ..., congestion has also repercussions on the daily lives of many people. Getting caught in traffic causes stress and forces people to change their travel behavior e.g., by leaving for work earlier, by choosing an alternative mode of transportation, ... The stress and the time losses caused by (unexpected) extended commuting times have a negative

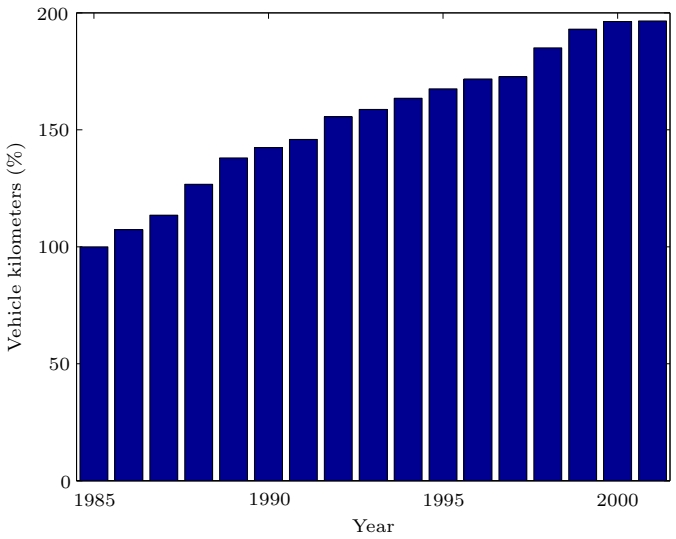


Figure 1.1: An overview of the evolution of the total number of vehicle-kilometers traveled per year in Belgium. The year 1985 is chosen as the reference. Source: Flemish Government, LIN Department.

impact on the quality of life of all individuals involved. Another congestion cost for the society is caused by the occurrence of potentially dangerous situations (e.g., a traffic jam in an environment with reduced visibility). Congested traffic operation results in an increased fuel consumption, which has both an economical and an ecological price. Different strategies can be applied in order to tackle the congestion problem. Some of these techniques include:

- The construction of new roads in order to eliminate the most important bottlenecks or to eliminate 'missing links' in the existing motorway network. As an example of the elimination of missing links we mention the plans of the government to close the ring road around Antwerp in Belgium. Bottlenecks can be reduced or eliminated by adding more lanes. The main disadvantages of this strategy are that the construction of new roads is very costly and that adding more capacity at one location often merely shifts the congestion to another location.

- Stimulate the use of other modes of transportation such as public transportation, car-pooling, transport of goods by train or by ship, ... However, in order to realize a shift from road transportation to other modes of transportation, the other modes of transportation need to offer a viable alternative, which often requires large investments.

- Reduce the demand for road transportation by e.g., public transportation (train, tram, bus, boat, ...), the development of a high-speed communications network (tele-working), additional taxes, ... A decrease in the demand for road transportation can either be established by providing alternatives or by charging taxes on road transportation. The former requires investments of the government while the latter is an unpopular and politically hard decision [29].
- An improved use of the existing infrastructure, which can be implemented by control measures such as road pricing, better transportation planning by the industry, improved traffic control (e.g., dynamic traffic management), ...
- Implementation of Intelligent Vehicle/Highway Systems (IVHS), which apply a combination of control, telecommunication and computer technologies in order to assist the driver in reaching the desired destination while attempting to globally avoid congestion and improve safety. Since IVHS systems require on-board sensors, processing power, ..., which are not readily available in vehicles nowadays, this approach faces resistance of vehicle owners not willing to invest in new technology.

Given the huge costs of congestion to the society and the urgent need for solutions, this text investigates a solution applicable in the short term. More in particular, ramp metering is considered as a traffic control measure in an advanced traffic management system set-up that optimizes the use of the currently available traffic capacity [26,136]. Ramp metering is merely an example of a traffic control measure that can be controlled using the model predictive control framework for advanced traffic management systems that we present in this thesis. We chose to discuss ramp metering as a traffic control measure in this thesis since its main advantages are that it can be realized in the short term and at a minimal cost. Ramp metering can use the existing infrastructure and only relatively limited investments are necessary.

1.2 General overview

In this section, we present a general overview of the topics discussed in this thesis and their relations with respect to each other.

There exist several approaches to deal with congestion on motorways. Given the sub-optimal operation of traffic during congestion and given the dynamic nature of the phenomena that occur during traffic operation, we investigate traffic control on motorways from an automatic control point of view. In this thesis we develop an advanced traffic management system consisting of a model predictive control (MPC)-based ramp metering controller.

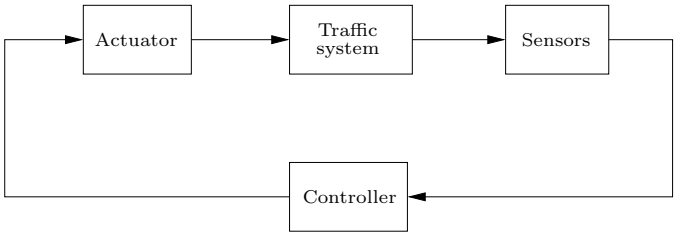


Figure 1.2: Schematic representation of a control loop. The controller determines the optimal control signals to send to the actuators based on the traffic measurements provided by the sensors.

An advanced traffic management system operates according to the control loop concept known from automatic control theory (see Figure 1.2). The advanced traffic management system consists of a controller, which is equipped with traffic measurements. Based on these traffic measurements, the controller dynamically computes the appropriate control signals to steer the traffic system towards a desired state. The controller sends its control signals to actuators or to lower level controllers that interact with the traffic system according to the received control signals.

From the available traffic control strategies that we present in Section 2.3, we chose ramp metering since it can be implemented easily and at a relatively low cost. Moreover, the ramp metering infrastructure can be completely implemented along the roadside and no investments in new in-vehicle technology are required, which improves the acceptance by the public. Currently, ramp metering is already operated in many countries around the world [76, 124]. However, the model predictive control approach to ramp metering presented here is new. As we show in a simulation example in Chapter 5, model predictive control based ramp metering has some advantages over the more classical approaches that are currently used. After the development of the ramp metering controller and some simulation experiments, two extensions are added to improve the robustness of the controller and to take re-routing effects due to the ramp metering actions into account.

The relations between the topics we present in this thesis are presented in Figure 1.3. Let us now discuss each of these topics in more detail.

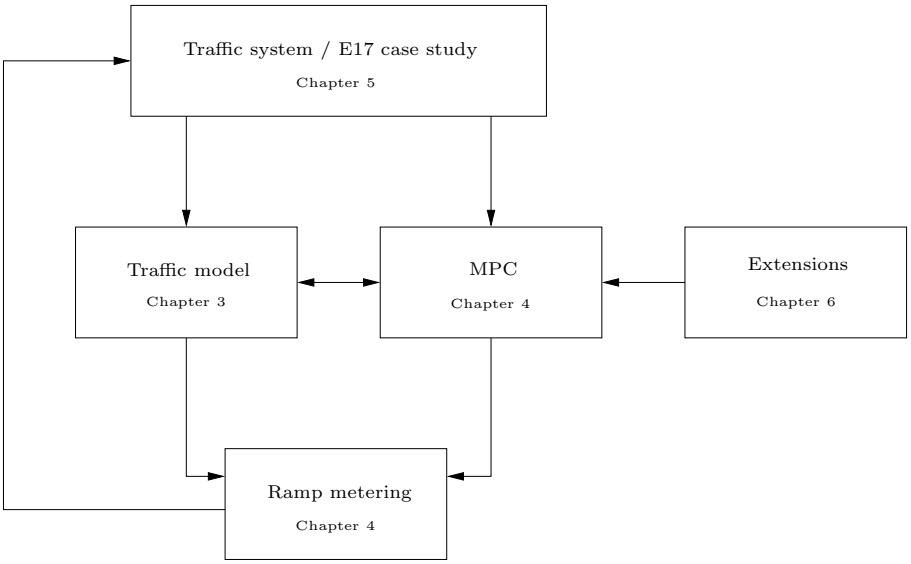


Figure 1.3: General overview of the main topics discussed in this thesis. In this thesis we consider the E17 motorway Ghent–Antwerp as a case study of a traffic system. A traffic model is implemented for the case study and an MPC-based ramp metering controller is developed. In Chapter 6 we provide an extension to the MPC-based ramp metering framework to eliminate external disturbances and an extension to account for traffic re-routing due to control actions.

The E17 motorway traffic system

The traffic network that we consider in this thesis is a case study network consisting of the E17 motorway Ghent–Antwerp in Belgium. We have selected this case study network based on the observation that recurrent congestion occurs on the E17 motorway in the driving direction of Antwerp during the morning rush hour and based on the availability of traffic measurements. The E17 motorway case study is discussed in detail in Chapter 5. Since the traffic measurements on the case study motorway are the main link between the traffic system and the other components discussed in this thesis (see Figure 1.3), we discuss the most popular sensor technologies and the measurements they provide. We illustrate the operation of ramp metering, the traffic control measure of our choice in this thesis, by plotting the fundamental diagrams for the case study using traffic measurements.

Motorway traffic simulation models

We have implemented a traffic model for the case study since due to safety concerns it is not possible to run experiments on the real-life motorway. This

model is used to test the performance of the ramp metering traffic controllers developed in Chapter 4 instead of applying the ramp metering control to the real-life traffic system. Since a variety of traffic models exists and in order to make an informed decision about which model to choose, we describe some ways to classify traffic models based on their properties in Chapter 2. We have selected to model the E17 motorway case study using the deterministic, macroscopic METANET traffic flow model. This model is described in detail in Chapter 3. During the identification phase, we use traffic measurements of the traffic system to determine the appropriate model parameters. The traffic demands, which are needed in order to simulate the model, are also determined using measurements from the traffic system. In the appendix, we discuss an improved first-order macroscopic traffic flow model that trades off accuracy and computational complexity.

Ramp metering for advanced traffic management

In this thesis, we present an advanced traffic management system using ramp metering as a traffic control measure. Note however that the model predictive control (MPC) framework is also applicable to advanced traffic management systems using other traffic control measures or using a combination of traffic control measures. An overview of traffic control measures for motorway traffic is given in Chapter 2. As can be observed in the schematic representation of the control loop in Figure 1.2, the control signals are provided by the controller to the actuator. In a real-life implementation of MPC-based ramp metering, this results in the controller providing metering rates to the ramp metering set-up in the traffic system. However, since we assess the controller performance using a model of the case study motorway instead of the real-life traffic system, the controller provides the metering rates to a ramp metering set-up in the model.

Model predictive control in a traffic framework

In this thesis, we implement the control part of the control loop in Figure 1.2 by a model predictive control based controller, which we describe in Chapter 4. The model predictive control based controller receives traffic measurements from the traffic system and uses a traffic prediction model to determine the metering rates that minimize an objective function. The choice of the prediction model is not imposed by the MPC controller but by the level of detail the objective function requires and by the computational complexity of the model.

Extensions to the MPC framework

In Chapter 6 we add two new extensions to the MPC framework. In order to eliminate the misfit between the traffic prediction model used in the MPC controller and the real-life traffic system, which can reduce the performance of

the controller as we illustrated with a simulation example in Chapter 6, we propose to regularly re-fit the MPC prediction model using traffic measurements. The second extension we have developed is anticipative MPC-based control, which allows the MPC-based controller to take re-routing of the traffic due to control measures into account. This extension is illustrated in Chapter 6 using a simulation example.

1.3 Chapter by chapter overview

In this thesis we study advanced traffic control on motorways. Model predictive control is discussed as a new technique to optimally control a ramp metering set-up. This technique is applied to a real-life case study, which is the motorway E17 Ghent–Antwerp in Belgium. After the discussion of ramp metering control for the case study, some extensions to the MPC framework are presented: model re-fitting and anticipative MPC-based control of ramp metering. Finally, both model re-fitting and anticipative MPC-based control are combined to obtain an anticipative MPC-identification approach to ramp metering.

A schematic representation of the chapters of this thesis and their relations is shown in Figure 1.4. We now present a chapter by chapter overview of the contents of this thesis.

Given the need for a solution that can be implemented in the short term, advanced traffic management systems are presented in Chapter 2 as a way to deal with congestion. Two basic components of an advanced traffic management system are discussed in Chapter 2: data acquisition and traffic control measures. Since the model predictive control approach to traffic control, which we present in this thesis, requires traffic models, we also present a brief overview of different types of traffic models and a traffic model classification in Chapter 2. Based on traffic measurements provided by the traffic sensors, we discuss the fundamental diagram, a well-known concept in traffic flow theory, which plays an important role in understanding the concepts of ramp metering that will be presented in Chapter 4. This chapter is based on the material covered by the papers [7, 8, 40, 41].

In Chapter 3, the METANET traffic flow model is presented in both its non-destination oriented and its destination oriented form. The non-destination oriented METANET model can be used to implement a traffic simulation model for a traffic network in case routing decisions in the network are not considered. If the assignment of traffic over alternative routes from an origin to a destination is considered, the destination oriented version of the METANET model is used.

Chapter 4 discusses the ramp metering controller design. First, the ramp metering concept is presented based on a detailed discussion of the fundamental

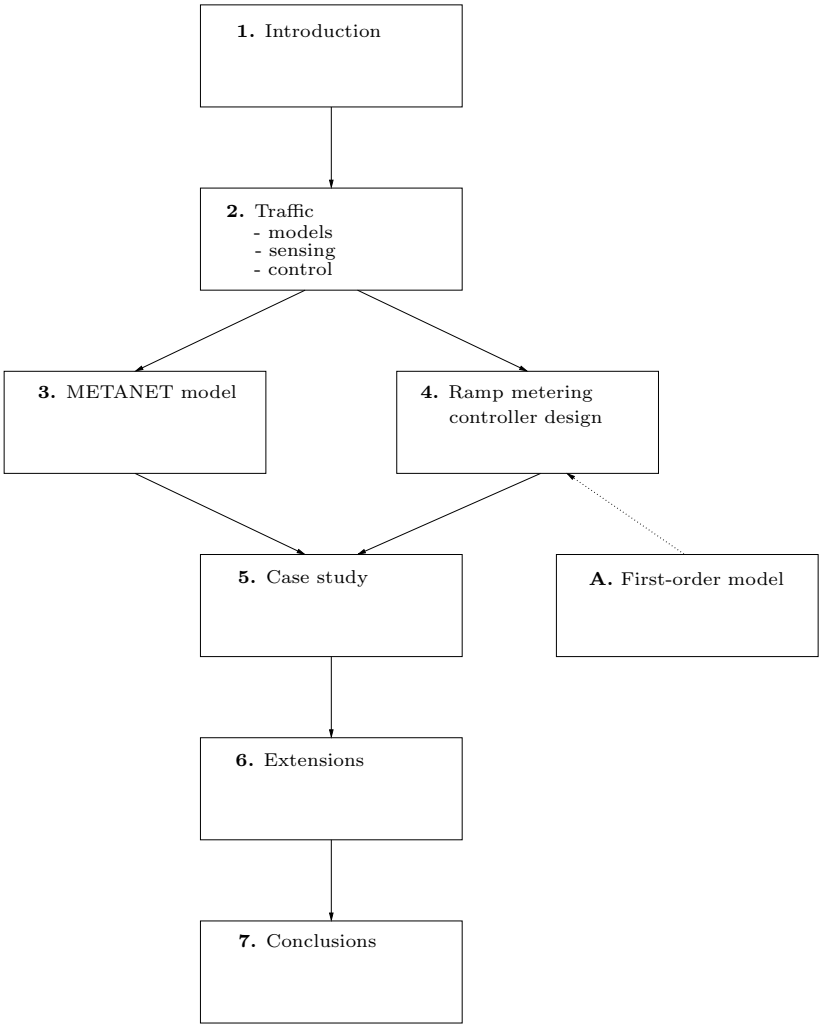


Figure 1.4: Schematic representation of the chapter by chapter organization of this thesis. After the introduction we discuss traffic models, traffic sensors and traffic control measures in Chapter 2. In Chapter 5, the METANET traffic flow model from Chapter 3 and the MPC-based ramp metering controller design from Chapter 4 are applied to the E17 motorway Ghent–Antwerp case study. In Chapter 6, we provide two extensions to the MPC-based ramp metering control framework. We present the conclusions of this thesis and some directions for further research in Chapter 7.

diagram. Next, the low-level implementation of ramp metering is briefly discussed. After a short discussion of some ramp metering algorithms, the widely adopted ALINEA ramp metering algorithm is presented. A more advanced ramp metering algorithm is MPC-based control of a ramp metering set-up. The MPC concept is discussed and applied to the control of a ramp metering set-up in Chapter 4. MPC-based ramp metering operates in a receding horizon framework where the metering rates are optimized using an objective function. The receding horizon concept, the objective function and the computation of the control signals (the metering rates) are all discussed in Chapter 4. The research presented in this chapter resulted in the papers [12, 14].

In Chapter 5 the performance of ALINEA-based ramp metering controllers and MPC-based ramp metering controllers is compared based on a simulation study of the E17 Ghent–Antwerp. First, the case study set-up and the features of the studied E17 motorway stretch are discussed. Since due to safety (and budgetary) concerns, no experiments can be run on the real-life motorway, a computer simulation model for the E17 motorway stretch is implemented using the METANET model. The simulation of the motorway requires information about the traffic demands over the simulated period. Using the traffic measurements discussed in Chapter 2, the traffic demands for the case study are estimated. The following scenarios are investigated: no control, ALINEA-based ramp metering control and MPC-based ramp metering control. The no-control case serves as a reference to assess the potential performance improvement due to ramp metering control on the E17 motorway. MPC-based ramp metering is presented as a new and advanced control method and its performance is compared with the performance of ALINEA, a standard ramp metering control algorithm. Different ramp metering configurations are investigated for the E17 motorway: an isolated ramp metering set-up, non-coordinated ramp metering set-ups, and coordinated ramp metering set-ups. Based on the performance of the controllers, some conclusions are drawn concerning where to install ramp metering on the E17 motorway and which algorithm to use. The research presented in Chapter 5 resulted in the papers [14, 15].

Some extensions to the MPC-based ramp metering controller are presented in Chapter 6. An MPC-based controller uses a prediction model to predict the future traffic states on the motorway. An offset between the real situation and the model can result in a reduced performance of the controller. Hence, a combined MPC-identification approach to ramp metering is developed that re-fits the prediction model using newly available traffic measurements. This topic is also addressed in the papers [10, 11]. Chapter 6 also discusses anticipative ramp metering control. Traffic control measures such as ramp metering influence the route choice behavior of the drivers in the network. In chapter 6 a new approach to anticipative MPC-based ramp metering control is presented. Anticipative MPC-based ramp metering is able to assess the re-routing behavior of the drivers and adapts the control signals (metering rates) such

that the desired traffic state (described by the objective function) is obtained despite re-routing. Anticipative MPC-based ramp metering is illustrated using a simulation example. This method was also presented in [13]. Finally, both extensions are combined in one anticipative MPC-identification approach to ramp metering.

Chapter 7 presents the conclusions of this thesis and presents some directions for future research.

In the appendix, an improved version of a first-order traffic flow model based on the model of Lighthill, Whitham and Richards is presented. Although a first-order model is less detailed than higher-order models such as e.g. the METANET model, which is presented in Chapter 3, first-order traffic flow models are still used in practice especially in on-line control applications for large networks where the computational complexity of second-order models is too high. The improved model of Lighthill, Whitham and Richards was also presented in the paper [9].

1.4 Contributions of this thesis to the state of the art

After the chapter by chapter overview of the previous section, we highlight the contributions of this thesis to the state of the art.

- In Chapter 2, we present an efficient interpolation scheme for the estimation of missing traffic measurement data using a reference day. The presented interpolation method is numerically efficient such that it can be applied in on-line systems. By recomputing the reference day, which can be done off-line, slow changes in the traffic conditions, e.g., due to changing seasons, can be tracked.

- In Chapters 3 and 4, we make a distinction between the simulation time step of the simulation model and the control time step of the controller. This distinction, which is not made in the traffic control literature, results in a lower computational complexity of the traffic controller since the control signals need to be updated less frequently due to the controller time step being larger than the simulation time step.

- In Chapter 4, we apply the model predictive control framework to ramp metering control. The MPC-based ramp metering controller operates in a rolling horizon framework and is able to take hard constraints into account.

- In Chapter 5, we implement a METANET simulation model for a case study consisting of the E17 motorway Ghent–Antwerp using real traffic data and we apply MPC-based ramp metering control to this simulation model. Both MPC-based control of a single on-ramp and MPC-based coordinated control of two on-ramps are investigated using simulations. These simulations illustrate that the MPC-based controller is able to improve the performance of the case study motorway compared to the no control case and compared to the ALINEA-based control case. The hard constraints imposed on the queue length were not violated by the MPC-based ramp metering controller.

- In Chapter 6, we develop the combined MPC-identification approach to ramp metering, which reduces the misfit between the MPC prediction model and the real-life motorway properties.

- Also in Chapter 6, we develop anticipative MPC-based ramp metering control. In the scope of this anticipative MPC-based ramp metering control framework, we develop an anticipative traffic assignment strategy, which models the re-routing behavior of the drivers based on their perception of the traffic state of the motorway network and the traffic demands. We also suggest a combined anticipative MPC-identification approach to ramp metering as a topic for further research.

Chapter 2

Traffic modeling, sensing and control - General overview

Given the huge costs of congestion for the society and the urgent need for solutions, this thesis investigates a solution that is applicable on a short term basis: advanced traffic management systems (ATMS). An ATMS is a dynamic traffic control system that optimizes the use of the currently available traffic capacity using automatic control concepts. The main advantages of this approach are that it can be realized in the short term and at a minimal cost. Indeed, the existing infrastructure can be used and only limited investments are necessary, especially if the costs to implement and to operate ATMS are compared to the cost of the construction of new roads.

In order to set up an ATMS, several key components should be present as shown in Figure 2.1. There is a need to determine the optimal control policy and to examine the effects of certain traffic control measures on the actual traffic situation and on congestion. The optimal control policy is determined by the controller and is based on measurements of the state of the traffic system that are provided by sensors as shown in Figure 2.1. The controller provides the control policy to the actuators in the system, which apply the control policy to the traffic system. ATMS often use a model-based approach to the traffic control problem. In a model-based approach, the controller uses mathematical models and simulation tools to estimate the impact of various traffic control measures on the traffic state and on congestion in the controlled traffic system. In this chapter, we discuss the three key components of an ATMS using model-based control: traffic sensors, actuators and traffic models. Model predictive

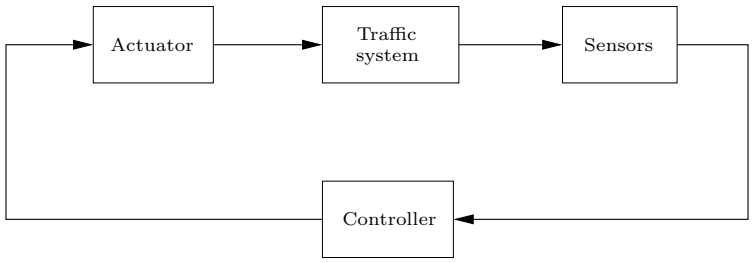


Figure 2.1: Schematic representation of a control loop. The controller determines the optimal control signals to send to the actuators based on the traffic measurements provided by the sensors.

control, a special case of model-based control, will be discussed in Chapter 4. For a more general overview of traffic control, the interested reader is referred to [39, 73, 79, 97, 106] and the references therein.

2.1 Traffic model classification

A wide variety of traffic models exists. Since we will use traffic models later on in this thesis, we provide a brief summary of the classification of the traffic models based on their properties. This section presents four ways to classify traffic models from the point of view of a (traffic) control engineer:

- Physical interpretation;
- Level of detail;
- Deterministic versus stochastic;
- Discrete versus continuous.

While here we will only present a brief summary, the interested reader is referred to [68] for a more detailed overview of the existing traffic models and for a more detailed classification of these models.

2.1.1 Physical interpretation

According to system theory [89], there exist three major approaches towards modeling: white box, black box, and grey box modeling. Each of these approaches is now presented in more detail and their application to traffic models is discussed.

- White box modeling or the deductive approach:
  In the deductive approach physical equations describe the relations between different states of the traffic system. This approach is used for instance to describe a mechanical system by Newton's laws and is called white box modeling (first principles modeling). The relations between the states of the system are described by properties that can be measured (e.g., in Newton's second law of motion the mass of an object relates the force applied to the object with its acceleration).

- Black box or inductive modeling:
  In a black box modeling approach, input and output data of the system are recorded and a generic parameterized model is fitted to the data. For a traffic system, the input-output data typically consist of the evolution of the measured traffic flows, the measured traffic densities and the measured speeds over time. With the black box method there is in general no direct physical relation between the modeled traffic situation and the model structure. An example of black box traffic modeling is the application of neural networks trained to mimic the behavior of the traffic system by e.g., modeling the future traffic density as a generic function of the current density the current speed and the on-ramp and off-ramp flows [65].

- Grey box modeling:
  Finally, an intermediate method between black box and white box modeling can be distinguished. This method is known as grey box modeling, and is a combination of the inductive and the deductive approach. In the deductive phase, parameterized equations between the states of the motorway system are written down. During the inductive phase, the model parameters are tuned by fitting the input-output relation of the traffic model to input-output measurements of the traffic system. Examples of this intermediate approach are the first-order traffic flow model of Lighthill, Whitham and Richards [87,123] and the second-order traffic flow models proposed by Payne [117] and Papageorgiou [96, 133]. The equations in these models have a physical interpretation but contain parameters that need to be fitted using input-output data from the traffic system.

2.1.2 Level of detail

Traffic models can also be classified based on their level of detail. Three classes of traffic models are discussed here: microscopic, mesoscopic and macroscopic traffic models.

- Microscopic traffic models:
  Microscopic traffic models are models in which all vehicles are described individually. Every single vehicle is modeled using models describing e.g., the inter-vehicle interactions and the interactions between the vehicles and the motorway. The two most important microscopic vehicle models are the car following model and the lane changing model [155]. The car following model describes how a vehicle follows its predecessors. The following behavior can be modeled based on the headway a driver preserves between himself and the preceding vehicle, based on how the driver reacts on acceleration or deceleration of the vehicle in front of him, ... and based on combinations of the previous. The lane changing model describes how a driver decides whether or not to overtake its predecessor. Important vehicle and driver properties in the lane changing model are e.g., the speed desired by the driver and the acceleration abilities of the vehicle.
  The behavior of the drivers is not only influenced by inter-vehicle interactions but also by interactions between the motorway and the vehicles, which can be modeled using parameters such as e.g., speed limits on the motorway, the slope or the curvature of the motorway stretch, ...
  Based on similarities between the vehicle parameters, the vehicles can be classified in different vehicle classes such as e.g., cars and lorries. Both the car following and the lane changing models use vehicle and motorway specific parameters.
  Microscopic models can also be implemented as cellular automaton models [101, 102, 154]. A cellular automaton model describes the motorway as a network of interconnected cells. The behavior of the vehicles in the network is described by the way the vehicles are hopping from one cell to the next. The microscopic cellular automaton model consists of a set of rules describing how and when vehicles hop from one cell to another.
  During a microscopic simulation, all the individual vehicles and their interactions are simulated using the microscopic vehicle models. The combination of all simulated vehicle behaviors results in a global image of the traffic situation. Microscopic traffic models are often bundled in commercial traffic simulation software packages such as Paramics, AIMSUN2, FLEXSYT-II, HUTSIM, ... An extensive comparison of commercial microscopic traffic simulators is presented in the SMARTEST study [17]. Since microscopic models describe each vehicle individually, microscopic traffic simulations are often computationally intensive.

- Mesoscopic traffic models:
  Mesoscopic traffic models describe traffic in less detail than microscopic traffic models. Rather than describing each modeled vehicle individually, a mesoscopic traffic model describes vehicles on a more aggregate level. There exist different types of mesoscopic models such as e.g., headway distribution models, cluster models and gas-kinetic models. Headway distribution models [21, 24] describe the distribution of the headways of the individual vehicles. The time headway of a vehicle is defined as the time between the vehicle and its predecessor. Cluster models [19] describe groups or clusters of vehicles with similar properties (e.g., driver aggressiveness, desired average speed, ...). Gas-kinetic models [116, 120] describe the dynamics of the velocity distribution functions of vehicles in the traffic flow. Some gas-kinetic models were designed to model multi-class traffic [66] while other gas-kinetic models were designed to model multilane traffic [61]. A generic gas-kinetic model is presented in [67]. Since mesoscopic traffic models do not describe every vehicle individually, mesoscopic simulations are in general less computationally intensive than microscopic simulations.

- Macroscopic traffic models:
  Macroscopic traffic models are models that use aggregate variables in order to describe the traffic situation. Typically, a macroscopic model defines a relation between the traffic density, the average velocity and the traffic flow. In some models, the traffic density is replaced in the model by the occupancy (see Section 2.2.1 and Remark 2.1).
  Within the class of macroscopic models, a further classification based on the order of the models can be made. The first-order traffic flow model proposed by Lighthill and Whitham in 1955 [86] and independently by Richards in 1956 [123] is based on the analogy of traffic flows with flows of fluids [86]. The only state variable in this traffic model is the traffic density. The model of Lighthill, Whitham and Richards was extended later on in order to be able to better model shock waves and stop-and-go traffic in congested traffic situations [103].
  The model described by Payne in 1971 [117] is a second-order flow model since it uses two traffic state variables: the traffic density and average velocity. This model was later on extended by Papageorgiou to result in the METANET model, which will be discussed in Chapter 3. An overview of recent extensions of the Payne model is given in [88]. Helbing [60] proposed a third-order macroscopic traffic model with the traffic density, the average velocity and the variance on the velocity as state variables.
  Since macroscopic traffic models only work with aggregate variables and do not describe the traffic situation at the level of independent vehicles or at the level of packages of vehicles, they are in general less computationally intensive than microscopic and mesoscopic models.

For a more detailed discussion of the traffic model classification based on the level of detail the interested reader is referred to [68].

2.1.3 Deterministic versus stochastic

In a deterministic traffic model there is a deterministic relation between the input, the traffic states and the output of the model. So, if a traffic situation is simulated twice using a deterministic model, starting from the same initial conditions and with the same inputs and boundary conditions, the outputs of both simulations are the same. The METANET model, discussed in Chapter 3, is an example of a deterministic traffic model.

A stochastic traffic model contains at least one stochastic variable. This implies that two simulations of the same model starting from the same initial conditions, the same boundary conditions and the same inputs may return different results, depending on the value of the stochastic variable during each simulation. A stochastic variable can be characterized by its distribution function or by its histogram. For example, in the microscopic simulation package Paramics, a distribution of the level of patience over different drivers needs to be defined. During simulation, this distribution is sampled to determine an individual level of patience for every driver simulated in the network. A second simulation, and thus a new sampling of the distribution, will result in drivers with other levels of patience. Therefore, stochastic models need to be simulated repeatedly and the results need to be averaged in order to be able to draw conclusions. Stochastic variables are used to model stochastic processes present in real-life traffic situations.

2.1.4 Discrete versus continuous

A motorway traffic model describes the evolution of the state of the traffic network over time. There are two independent variables in a motorway traffic model, namely space and time. Depending on the type of model, these independent variables can be continuous or discrete.

Since the continuous traffic models for real-life situations are in general too complicated to solve analytically, especially if the size of the considered traffic network is large, they are in practice discretized in time and in space in order to simulate the traffic behavior using a computer [32, 83].

2.2 Traffic state measurements

2.2.1 Traffic variables

If a traffic situation is observed or modeled based on traffic measurements, different traffic variables can be measured to describe the traffic state on the motorway [95]:

- the traffic flow or the traffic intensity on the motorway,
- the speed of the vehicles,
- the traffic density,
- the occupancy level of the highway,
- the time headways,
- the distance headways,
- ...

One of the easiest variables to measure is the **traffic flow** (veh/h) or the traffic intensity on the motorway. The traffic flow is measured as the number of vehicles that pass a fixed point on the motorway during a certain time period.

The **mean speed** (km/h) of the vehicles is another important parameter to assess the traffic situation on the motorway. In practice, a value for the mean speed is obtained by averaging the speeds of the vehicles passing the sensor over a certain time period. The mean speed can be measured for each lane separately or it can be measured for all the lanes together.

The **traffic density** (veh/km/lane) is the number of vehicles on the motorway per unit of length at a certain time instant. The traffic density can vary from lane to lane. The total traffic density on the motorway is the sum of the lane traffic densities. Note that if the traffic is homogeneously distributed over the lanes of the motorway the motorway traffic density is given by the lane traffic density times the number of lanes[1].

The **occupancy** of a motorway or a lane is measured as the fraction of the time the detector detects a vehicle.

The **time headway** (h) between two vehicles is the time that passes between the arrival of the front of a vehicle and the arrival of the front of the next vehicle at a reference point.

[1] In the remainder of this text, the traffic is assumed to be homogeneously distributed over the different lanes.

The **distance headway** (km) is the distance between the front of a first vehicle and the front of the second vehicle. The distance headway consists of the length of the first vehicle plus the length of the gap between both vehicles.

Depending on the setup and the features of the measurement system, several levels of detail can be distinguished. Average values or instantaneous values for each vehicle can be considered. The averages can be averages over time (e.g., the time mean speed is the average speed at a certain location averaged over a time period) or the averages can be averages over space (e.g., the space mean speed is the average speed in a motorway stretch at a certain time instant)[2]. Depending on the application, the period over which the measurements are averaged can range from seconds to minutes to hours and even to longer periods. The traffic parameters can be measured for every motorway lane separately or a value for the motorway as a whole can be obtained. Furthermore, sometimes vehicles are classified in several vehicle types or classes such as cars, lorries, ... For every vehicle class certain measurements can be collected e.g., the average speed of the lorries traveling on the motorway. As a rule of thumb, the higher the level of detail required and the more parameters need to be measured, the higher the cost of the measurement system involved.

The most commonly measured traffic parameters are the traffic flow, the speed and the occupancy of the motorway traffic [48]. The choice of these parameters is influenced by their importance in traffic theory as well as by the ease they can be measured with the most common traffic detector technologies.

In the next section, the most widely used traffic measurement systems or sensors are presented as well as the problem of collecting all the distributed measurements in a central database.

Remark 2.1 Many macroscopic traffic flow models use the traffic density on the motorway as a state variable. The traffic density is often not available as a traffic measurement since, by its definition, it needs to be measured along a distance [50] and most popular sensors provide point measurements [48]. However, there is a relation between the traffic density and the occupancy [29]:

$$\text{Density (veh/km)} = \frac{\text{Occupancy}}{\text{Average vehicle length (km/veh)}} \ .$$

□

[2]Note that the time mean speed and the space mean speed need not be equal. However, the practical significance of the difference between time mean speed and space mean speed is found to be minimal [48, 50].

2.2.2 Sensor systems

There exist a wide variety of technologies [72, 75] to measure the traffic parameters described in the previous section e.g., pneumatic detection systems, inductive loops, cameras, ultrasonic sensors, microwave sensors, active and passive infrared sensors, passive acoustic arrays, magnetometers, ... In this section the three most widely used systems are discussed:

- inductive loops,
- traffic cameras and
- pneumatic tube detection systems.

The advantages and restrictions of each of these technologies are briefly summarized. For a more elaborate discussion on traffic sensor technologies and a comparison of their performance, the interested reader is referred to [1, 2, 75].

Inductive loops were introduced as a traffic detection system around 1960 and have become the most widespread traffic detection system to date [100]. An inductive loop consists of a conducting loop that is put in or on the surface of the motorway (see Figure 2.2). This loop is connected to an electronic circuit that produces an oscillating current through the loop. If a vehicle enters the loop, the metal parts of the vehicle cause the oscillation frequency of the electronic circuit to shift. This frequency shift is registered and the vehicle is detected.

There are two main configurations of inductive loop systems on motorways: the single-loop and the double-loop configuration. In the single-loop configuration, the traffic flow and the occupancy level of the motorway can be measured. Depending on the implementation, these measurements are available per lane or for all the lanes together. By defining an average length of a vehicle, an estimation of the speed can be computed based on the the time a vehicle spends in the loop. However, using a single inductive loop setup and an average vehicle length, it is hard to distinguish a fast driving lorry from a slow driving car, since the time the vehicle is in sight of the detector (in the loop) can be the same [50, 75]. In order to solve this problem and in order to improve the accuracy of the measurements, inductive loops are also implemented in double-loop configurations. In a double-loop configuration, two loops are installed with a distance Δx in between both loops along the driving direction (Figure 2.2). The registration of the occupancy level and the traffic flow is similar to the single-loop case, except that it can be made more accurate since the information of both loops can be combined. In a double-loop configuration the speed of the vehicles is calculated based on the time lag Δt between the activation of the first and the second loop. Once the speed and the occupancy time of a vehicle are known, the length of the vehicle can be computed allowing for a classification as e.g., a car or a lorry. The main advantages of the double inductive loop setup

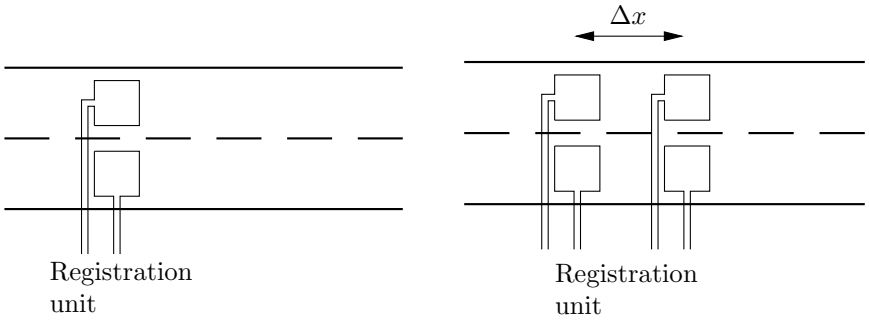


Figure 2.2: Schematic representation of a two lane motorway with a single-loop configuration (left) and a double-loop configuration (right). In the double-loop configuration the speed of the vehicles can be determined based on the delay Δt between the excitation of the loops and the distance Δx between the loops.

over the single-loop configuration are the improved accuracy and the additional traffic parameters that can be recorded. The main disadvantage is the higher cost with respect to the single-loop configuration. The main disadvantages of the inductive loop technology are the sensitivity of the loops during road maintenance works and the special installation and maintenance requirements resulting from the presence of the loops in (on) the road surface.

The presence of the inductive loops in the road surface causes very specific problems regarding e.g., maintenance, repairs, ... Traffic detection using cameras emerged around 1980 and is a non-intrusive technology that is becoming more and more popular [100]. A fixed camera is mounted above the motorway and its images are sent to a video processing unit that extracts the desired parameters using image recognition algorithms. An example of an image a traffic camera provides is shown in Figure 2.3. If the video images are sent to the traffic center they can also serve as a visual inspection tool. The main advantages of the camera traffic detection technology are the higher average accuracy of the measurements and the visual inspection possibilities. However, the main disadvantages are the high cost of the installation and the reduced accuracy during whether conditions with low visibility [1, 2, 77].

Pneumatic traffic detection systems [75] consist of a tube that is led over the road as shown in Figure 2.4. A vehicle driving over the tube produces pneumatic shock waves[3] in the tube which are detected and processed. Based on the detected shock waves in the tube and the delays between consecutive shocks, an estimation of the traffic flow, the speed and even a classification of the vehicles can be made. The main advantage of this technology is the mobility of the installation due to its ease of installation. It allows for a quick set up of a short term measuring point along a motorway. Since the tubes are led on the road

[3]Every axle of the vehicle produces a pneumatic shock wave in the tube.

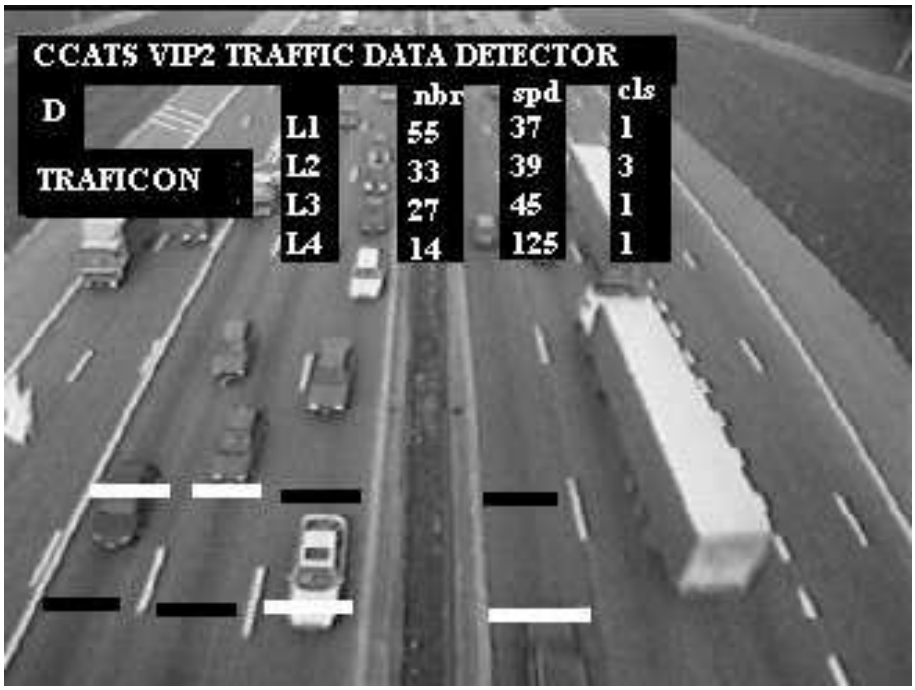


Figure 2.3: A view of a motorway as seen by a traffic detector camera. The video images are sent to a video processing unit that extracts the desired traffic parameters for the different motorway lanes. (Photo courtesy of Traficon NV)

Figure 2.4: The installation of a mobile pneumatic tube traffic detector. The tube is fixed on the road surface and connected to a registration unit at the roadside. (Photo courtesy of PET Corp)

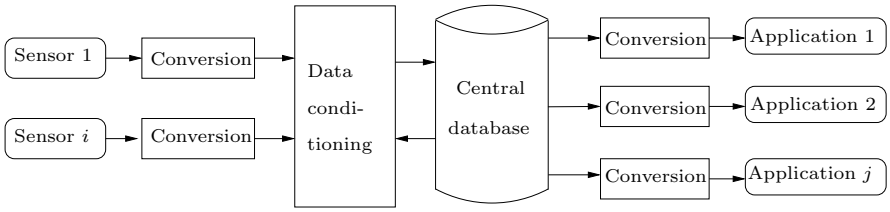


Figure 2.5: Schematic view of data flows from sensors towards the database and from the database towards the applications. The conversion steps provide modularity and flexibility to the system. In the data conditioning step missing or corrupt data can be detected and/or estimated.

surface they are susceptible for wear and tear caused by the traffic, hence the pneumatic sensors are not suitable for long term operation. A disadvantage of pneumatic tube detectors is their limited accuracy.

Interfacing

In order to obtain an overview of the state of a motorway network, traffic sensors collect traffic state measurements at different locations throughout the motorway network. If a motorway network is studied as a whole or if an ATMS is implemented, the distributed measurements need to be centralized. For temporary measurement setups (e.g., pneumatic tube detectors) the data can be collected manually. However, for ATMS systems the collection of the measurement data is automated and the measurements are stored in a central database. The central traffic measurement database is kept on a computer that is networked to the sensors at the different measurement sites. At regular time intervals (e.g., every minute in Belgium), the computer polls the sensors for the newly measured parameters and stores the values in the database.

Since different traffic detector technologies and detectors from different vendors often provide their data in (slightly) different formats, the measurements often need to be converted to the proper standard data format before they are stored in the centralized database. The main advantage is the modularity of this approach as is shown in Figure 2.5. If a sensor is replaced with a sensor of another brand or technology, only the conversion step needs to be adjusted and the database remains unchanged. With camera technology becoming less expensive, defective loops could be replaced by traffic cameras. This would only require an adaptation of the conversion step in order to be able to operate the camera with the existing network and the existing central database. E.g., on the E17 motorway Ghent–Antwerp traffic cameras were installed in the scope of a European research project. A conversion step converts the output of the measurement cameras to the output format of the inductive loops such that the cameras can be hooked up to the existing network. The traffic data used

in the case study in Chapter 5 were captured by these traffic cameras and were downloaded from the central database computer although the system was originally designed to deal with inductive loop data.

The modularity of the system shown in Figure 2.5 is also useful for the applications that use the data in the central database. The data collected in the central database can be used for many different applications [49] such as e.g., traffic control, incident management, archiving, providing travel information, ... The data that we use in this thesis were obtained using a custom application to extract traffic data from the database. An upgrade of one of the applications using the database or adding a new application implemented by a different vendor might require a data format that is not compatible with the data format of the database. This compatibility problem is solved by the modular approach.

2.2.3 Data conditioning

Since failures and disturbances can occur at the sensors or during the transmission of the measurements, it can be useful to perform a data conditioning step before the data is used. First, the detection of corrupt and missing data is considered followed by a discussing of the estimation of missing measurements.

Corrupt and missing data

After the sensor data is transferred over the network to the computer containing the database, the data is typically checked for corrupt or missing data values [49].

Every measured parameter is restricted to an interval of possible values that is based on physical laws and common sense assumptions e.g., a vehicle count cannot be negative, the speed of a vehicle on a motorway cannot be negative, the speed of a vehicle is limited, ... If a measured value does not satisfy the constraints, it is marked as unavailable. In this way the corrupt data values are transformed into missing data values.

Rejected measurement values due to corrupt data is only one cause of missing measurement values. Sometimes the central computer gets no answer to its polls because a sensor is out of order due to works on the road, power failure, ... or because the network is not working properly. Since some applications require data records which are complete and contain no gaps, some methods to estimate the missing values are presented next.

Estimation of missing data

Different estimation techniques can be used if a measurement is missing. In general, the missing measurements are estimated using a black box model that is combined with available measurements. The model to approximate the missing values can be a piecewise-linear function, a polynomial function, splines, ... The missing values can then be determined by interpolation or by extrapolation.

The main advantage of linear interpolation or extrapolation is its computational efficiency. The main disadvantage is its limited accuracy especially if extrapolation is used, i.e., if the time index of the missing value lies outside the time range of the available data points. This becomes important when a measurement is unavailable for a longer time, since in that case estimates are needed for successive values. Interpolation and extrapolation are in fact local estimation methods that use an approximation that is valid in the neighborhood of the available measurements. The use of this approximation over a larger interval can give rise to large deviations of the estimation from the real value.

Better and more general approximators can be used such as polynomial approximations (e.g., seasonal autoregressive integrated moving average (ARIMA) models [20,151]), nonparametric regression [127], neural networks [158], splines, ... The use of more data points in these models results in a higher attainable accuracy, but also significantly increases the computational efforts [128]. In the next section we present a simple estimation scheme that is a good trade-off between complexity and accuracy when estimating missing values.

A numerically efficient estimation scheme

The method presented in this section is an interpolation of the missing value using a reference day. The reference day is constructed off-line using many data points. The reference day is a day of which the traffic situation is similar to the day under study. The missing value at a given time instant of the current day is determined based on the last available measurement and on information about the traffic situation at the corresponding time instant in the reference day. The missing value is looked up in the reference day and scaled to the traffic intensity of the studied day. This scaling is done using the following formula:

$$x_{\text{data}}(p) = \frac{x_{\text{data}}(p-1)}{x_{\text{reference}}(p-1)} x_{\text{reference}}(p)$$

where $x_{\text{data}}(p)$ is the measurement (or the estimate) of the variable at the p-th time sample of the current day, and $x_{\text{reference}}(p)$ is the corresponding value in the reference day. More advanced interpolation methods can be found in the signal processing literature [156].

The reasoning behind this method is that every day shows a traffic pattern that can be predicted, e.g., the build-up of congestion in the morning and the evening rush hours. If a day with similar characteristics as the current day can be found, i.e., the reference day, the information about the traffic pattern in the reference day is used to estimate the missing values in the current day. This is realized by scaling the value from the reference day to the intensity level of the current day using the scaling factor $\frac{x_{\text{data}}(p-1)}{x_{\text{reference}}(p-1)}$. This scaling is necessary since the traffic intensities of the present day and the reference day are not necessarily the same in absolute numbers whereas their traffic patterns (i.e., the evolution of the traffic intensities) are assumed to be similar [128]. The reference day is constructed by taking the average over several days that are expected to have the same traffic pattern as the current day e.g., Mondays, Fridays, weekdays, weekends, ... For each time instant the average is taken by first adding the corresponding available measurements and then dividing by the number of available values in the sum. This way, it becomes easier to construct a reference day, since also days with missing measurements can be used to construct the reference day. This method of determining the reference day averages out incidental fluctuations, which leads to a more dependable reference day and to more robust estimates.

The traffic patterns during the holidays are completely different from the patterns of an ordinary working day, even if both days are the same week days. Also, certain happenings (pop concert, football game, ...) or incidents (accident, works on the road, ...) can influence the traffic intensity. Last but not least, the weather also plays an important role in the way a traffic situation develops. Therefore, it must be avoided to include days that can be expected to exhibit a different traffic pattern into the calculation of the reference day. It is clear that a method that averages over several days will be more robust against occasional deviations.

If a long averaging window is used, which means averaging over many days, deviating values will be suppressed. E.g., the average over all Mondays for a whole year could be computed to construct a reference day to use for Mondays. However, this is not a realistic situation since there will be no distinction between the traffic pattern of a Monday in winter and the traffic pattern of a Monday in summer, even though the seasons influence the traffic patterns. There is a trade-off to be made between the averaging out of stochastic effects and tracking behavior when constructing a reference day. If the averaging window is made smaller, e.g., two months, which corresponds to eight Mondays, some ability to suppress deviating values in the data will be lost but the changes caused by the changing seasons will better be tracked.

This interpolation method using a reference day is computationally as efficient as ordinary interpolation or extrapolation while providing more accurate results. The construction of the reference day required some computational

effort but this task can be performed off-line. The interpolation method using a reference day offers a good trade-off between accuracy and computational efficiency.

2.2.4 The fundamental diagrams

If we plot the measurements of the traffic density, the average speed and the traffic flow near a bottleneck on a motorway, some interesting phenomena can be observed. In this section, we will plot and discuss the following diagrams based on measurements on a real-life motorway: the speed-density diagram, the flow-density diagram, and the speed-flow diagram. These diagrams are known in traffic theory as the fundamental diagrams [95].

The traffic data presented in these diagrams were recorded by a traffic camera on Wednesday, February 23, 2000 during the morning rush hour on the E17 motorway Ghent–Antwerp in Belgium. The camera is located in between the two left-hand sided on-ramps in Figure 5.2, since this is bottleneck where congestion sets in during the morning rush hour. In Chapter 5 we give a more detailed description of the E17 motorway configuration.

In Figure 2.6, the measured average speed on the E17 motorway is plotted as a function of the traffic density. It can be observed in Figure 2.6 that the average speed on the motorway decreases with increasing traffic density. This relation between the average speed and the traffic density on the motorway is in traffic theory often described with an empirical formula [73,95] that needs to be fitted to the data points. Many different empirical formulas for the speed-density relation are proposed in the literature [44, 45, 52, 53, 118, 137]. The solid line in Figure 2.6 represents a least-squares fitted empirical speed-density formula proposed by May [95]:

$$v(t) = v_{\text{free}}\left[1 - \left(\frac{\rho(t)}{\rho_{\text{jam}}}\right)^{\alpha}\right]^{\beta}, \tag{2.1}$$

where α and β are fitting parameters, t is the time, v_{free} is the free flow speed and ρ_{jam} is the jam density. We chose the speed-density formula proposed by May since it is a generic speed-density function[4]. Most other speed-density functions proposed in the literature are special cases of (2.1) [30]. The free flow speed v_{free} is the theoretical maximal average speed on the motorway. The free flow speed is reached as the traffic density on the motorway goes to 0. The traffic density where the average traffic speed is 0 or in other words where the traffic comes to a stand-still is called the jam density ρ_{jam} (see Figure 2.6). The jam density is the theoretical maximal density that can be achieved on

[4]Note that if we choose a speed-density function to be used in the first-order traffic flow model of Lighthill, Whitham and Richards, the chosen speed-density function must satisfy certain criteria in order for the resulting model to be numerically stable as we point out in Remark 2.2.

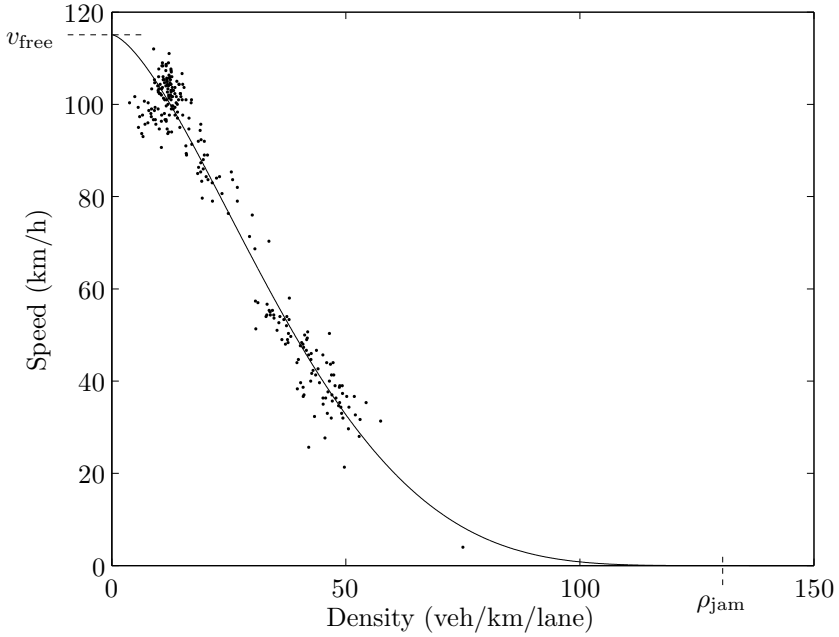


Figure 2.6: A plot of the measurements of the speed and the traffic density on the E17 motorway in Belgium (dots) recorded on Wednesday, February 23, 2000 from 5 a.m. to 11 a.m. The speed on the motorway is observed to decrease with increasing traffic density. An empirical speed–density relation (2.1) is fitted to the measurements (solid line). The parameters v_{free} and ρ_{crit} are also shown on the plot.

the motorway. By least-squares fitting (2.1) to the measurements on the E17, the values of v_{free} and ρ_{jam} were determined [134]. The values $\alpha = 1.5$ and $\beta = 5$ were chosen as proposed by May [95]. The values of v_{free} and ρ_{jam} were found to be 115 km/h and 136 veh/km/lane respectively. A better fit of the parameters v_{free} and ρ_{jam} is computed in Chapter 5 using measurements of multiple days.

There exists a fundamental relation between the traffic density $\rho(t)$, the average speed $v(t)$ and the traffic flow $q(t)$ on a motorway at time t [95]:

$$q(t) = \rho(t)v(t)n \tag{2.2}$$

where n is the number of lanes on the motorway. This relation is always valid and allows to calculate the third quantity based on the knowledge of the two other quantities.

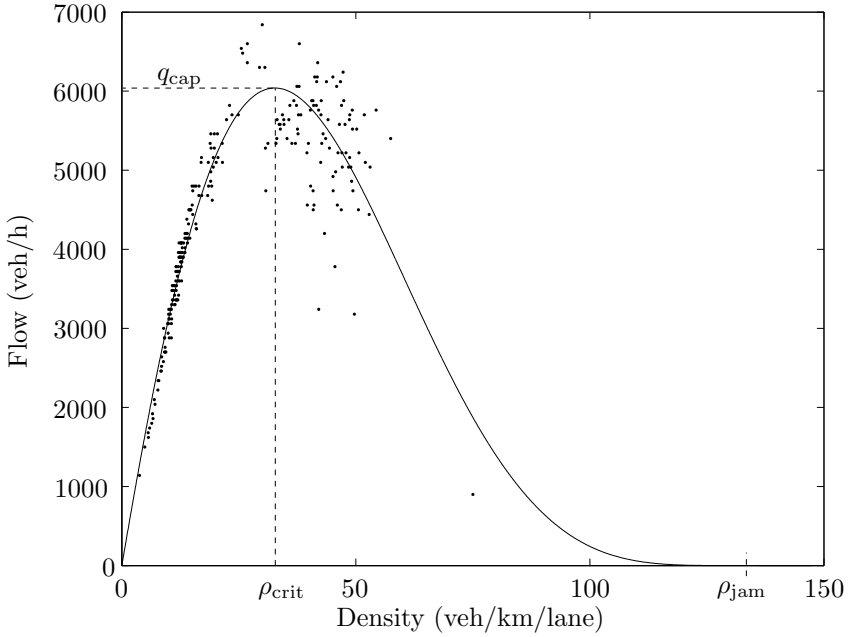


Figure 2.7: A plot of the measurements of the traffic flow and the traffic density on the E17 motorway in Belgium (dots) recorded on Wednesday, February 23, 2000 from 5 a.m. to 11 a.m. The presented traffic flow is the total flow on all three of the lanes of the motorway. The flow–density relation resulting from the choice of the empirical relation defined in (2.1) is plotted in the solid line. The traffic flow initially increases with increasing traffic density until a maximal traffic flow q_{cap} is reached corresponding to the critical density ρ_{crit}. If the traffic density on the motorway continues to increase, the traffic flow starts decreasing with further increasing motorway traffic density.

In Figure 2.7 the measurements of the flow versus the traffic density on the E17 motorway are presented. By substituting the empirical speed-density relation (2.1) in (2.2), an expression for the flow as a function of the traffic density can be found. From Figure 5.2, we know that the number of lanes n at the camera location (CLO3) on the E17 is 3. The flow-density relation resulting from the substitution of the empirical speed-density relation (2.1) in (2.2) is plotted as a solid line in Figure 2.7. Under low traffic conditions, the traffic flow increases in a nearly proportional way with increasing traffic density. If the traffic density keeps increasing, the traffic flow starts saturating until a maximal flow is reached at the critical density ρ_{crit}. This saturation of the traffic flow is due to the decrease of the traffic speed with increasing density,

which was observed in Figure 2.6. For the relation between speed and density given in (2.1), the critical density ρ_{crit} is given by:

$$\rho_{\text{crit}} = \rho_{\text{jam}} \left(\frac{1}{1+\alpha\beta} \right)^{\frac{1}{\alpha}} . \tag{2.3}$$

This equation for the critical density ρ_{crit} is obtained by substituting (2.1) in (2.2) and by calculating the traffic density for which the derivative of the traffic flow to the traffic density becomes 0. By substituting the values of the fitted parameters into (2.3), we find a critical density of 32.5 veh/km/lane for the E17 motorway based on this morning rush hour dataset. The theoretical maximal flow associated with the critical density ρ_{crit} is called the capacity q_{cap} of the motorway. By substituting the expression for ρ_{crit} into (2.2), we find a capacity of 6040 veh/h for the three lane motorway or a capacity of 2013 veh/h/lane under the assumption the traffic is homogeneously distributed over the three lanes. Once the critical density is reached, traffic breakdown occurs and the traffic flow starts decreasing with further increasing traffic density. As soon as breakdown of the traffic flow at ρ_{crit} occurs, congestion sets in and traffic starts operating in congested regime. The congested regime is unstable in the sense that a perturbation that momentarily increases the density in the motorway segment will cause the traffic flow to decrease thus giving rise to an even larger traffic density. If the jam density ρ_{jam} is reached, the speed on the motorway becomes equal to 0 and according to (2.2) the flow becomes also 0. Stable, free flowing traffic operation can only occur at densities below the critical density ρ_{crit}.

The third fundamental diagram is a plot of the speed on the motorway as a function of the traffic flow as presented in Figure 2.8. We see that a measurement of the traffic flow does not uniquely determine the traffic state on the motorway. In general, there exist two traffic regimes, which result in the same traffic flow on the motorway: one regime with a high speed and one regime with a low speed. Indeed, a flow of low density traffic at high speed can result in the same traffic flow on the motorway as a flow of high density traffic at a low speed. It can be verified in Figure 2.6 that the traffic states with high speeds correspond to low traffic densities and vice versa, the traffic states with low speeds correspond to high traffic densities. Only for capacity flow, there is only one corresponding speed as can be seen in Figure 2.8.

Remark 2.2 Equation (2.1) represents a generic form [30] of the empirical speed-density formula that can be used both in the first-order traffic flow model of Lighthill, Whitham and Richards (LWR) [87, 123] and in the second-order macroscopic traffic flow model of Payne [117] and Whitham [150].

In the first-order LWR traffic flow model, the speed on the motorway adapts instantaneously, and according to the empirical speed-density formula that was

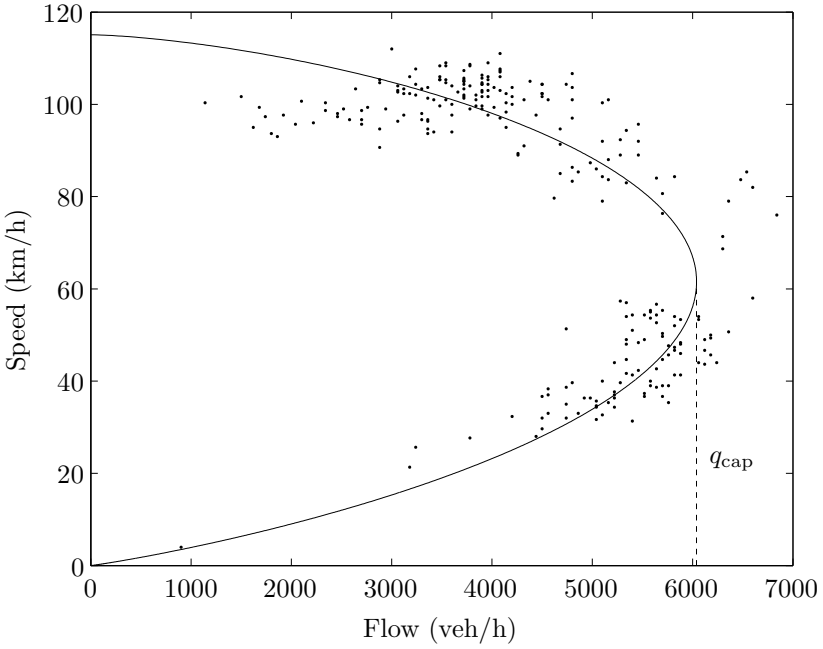


Figure 2.8: A plot of the measurements of the speed and the flow on the E17 motorway in Belgium (dots) recorded on Wednesday, February 23, 2000 from 5 a.m. to 11 a.m. The empirical speed–flow relation resulting from the choice of the empirical relation defined in (2.1) is plotted in the solid line. We observe that for each traffic flow smaller than the maximal flow q_{cap} there exist two traffic regimes: one with a high speed and a low traffic density and one with a low speed and a high traffic density.

chosen for the model, to the traffic density on the motorway [36]. Zhang [161] noted that in the context of the first-order LWR model, a choice of a speed-density formula that leads to a non-concave flow-density diagram results in numerical instabilities in the simulation model. Hence, for the first-order LWR model, the empirical speed-density formula needs to be chosen such that combined with (2.2) it results in a concave speed-density diagram.

However, in the context of the second-order traffic flow model of Payne and Whitham (and hence also in the METANET model that will be discussed in Chapter 3), this is not an issue since extra terms are added to the speed equation to take phenomena such as relaxation, anticipation and convection into account [107, 162]. □

2.3 Traffic control measures

This section provides a brief discussion of some of the alternatives for traffic control on motorways that are available. In the control loop presented in Figure 2.1, the traffic control measures correspond to the actuator that steers the traffic system according to the control signals provided by the controller. Differences between these alternatives that need to be considered include e.g., the costs of the implementation, the fact whether the control is informative or is enforced, the responsiveness of the control to varying traffic conditions and the complexity of the control problem.

- Peak traffic lanes
  Peak traffic lanes are lanes that can be opened or closed depending on the current traffic demand e.g., the opening of shoulder lanes and lane closures [133]. If the demand is high, an additional lane can be opened to the traffic. This control technique requires the presence of an extra lane. The optimal control problem of peak traffic lanes results in a integer optimization problem because the control variable is the number of lanes, which can only be an integer value. Sometimes the additional lane is reserved for high occupancy vehicles as is the case at several locations in the United States of America.

- Tidal flows
  Tidal flows [133] are quite similar to peak traffic lanes but the main difference is that the direction of the traffic flow on the lane can be reversed depending on the traffic demands. The dynamics of this technique are obviously very limited: because of safety considerations it is required that once a lane is opened in one direction it takes much time before it can be opened in the other direction.

- Velocity harmonization
  Velocity harmonization tries to enhance traffic safety by minimizing the speed differences between vehicles. This can be implemented as an automated system that provides the drivers with a desired speed through variable message signs [129, 130]. This desired speed can be informative or it can be a dynamic hard speed limit that is enforced. As an alternative to variable message signs, the police can impose the speed limits in the following way. During days with peak traffic demands during the summer holidays, the Belgian police creates platoons of vehicles on the E40 motorway, which are lead by a police vehicle. These platoons travel with a constant speed and overtaking is prohibited. This technique is known as 'blokrijden'.

- Intelligent speed adaptation systems
  Currently, experiments are being conducted with intelligent speed adaptation systems in Belgium and The Netherlands [23]. An intelligent speed

Figure 2.9: Picture of a gantry of the electronic road pricing (ERP) system in Singapore. Vehicles are detected and tolled according to the time of day and the chosen route.

adaptation system provides every vehicle with a desired maximal speed depending on factors such as the traffic situation, the location of the vehicle, the weather conditions, ... The intervention level of the intelligent speed adaptation system can be advisory (the driver is informed about the speed limit), voluntary (the system automatically limits the maximal speed of the vehicle but the driver can override the system) or mandatory (the driver cannot override the system) [140].

- Electronic road pricing
  Electronic road pricing (ERP) or electronic road tolling is a traffic control strategy where the traffic demand is influenced by pricing strategies. Figure 2.9 shows a picture of a gantry of the ERP system in Singapore. Drivers are detected and tolled based on the roads they use and when they used the roads. By choosing an adequate tolling strategy, the traffic demand can be reduced, spread in time (higher tolling during peak hours, during congestion [148], ...) and spread in space (higher tolling in certain areas).

Figure 2.10: A dynamic route information panel in Singapore providing expected travel times to several landmarks.

- Traffic re-routing
  If alternative routes are available, re-routing can be implemented [142]. Dynamic route information panels provide the drivers with information about alternative routes such as e.g., the expected travel times (Figure 2.10). Re-routing can be either obligatory or information-based.

- Ramp metering
  The ramp metering control strategy consists of preventing the traffic density on a motorway to grow too large by restricting the on-ramp traffic flows. Ramp metering is implemented by installing a traffic signal at the on-ramp of the motorway and is as such easily enforcable e.g., by means of a red light running camera. Ramp metering can be easily implemented in many of the current real-life traffic situations. Actually, ramp metering is already in use at several locations throughout the world e.g., in the United States of America [26], in The Netherlands, and in Belgium [76]. Figure 2.11 shows a picture of the setup in Kessel-Lo, Belgium. In Chapter 4 we give a detailed discussion of ramp metering.

Figure 2.11: Picture of the ramp metering setup on the E314 on-ramp in Kessel-Lo, Belgium.

2.4 Conclusions

In this chapter, we have discussed advanced traffic management systems from an automatic control perspective. We have discussed the three components in the control loop presented in Figure 2.1: the traffic models used in the traffic controller, the sensors, and the actuators.

The controller used in advanced traffic management systems is often a model-based controller that determines the appropriate control actions using a traffic model. Since the model properties such as e.g., the level of detail, the computational complexity, ... play an important role in the choice of the appropriate model to be used in the controller, this chapter started with a brief summary of criteria based on which traffic models can be classified.

In order to dynamically respond to the traffic state in the traffic system, the controller presented in Figure 2.1 depends on traffic measurements that are provided by sensors. First, the traffic parameters that are most often available and used in real-life situations were presented. Next, the most popular traffic measurement technologies were discussed. Failures or errors in the traffic sensors or in the data network during transmission can introduce corrupt or missing data. During a data conditioning step, the corrupt and the missing values are detected and often an estimation for the missing data is made. We have presented a computationally efficient estimation scheme that takes the expected evolution of the traffic state into account in its estimates. This is important if bursts of missing data need to be estimated. Traffic measurements can be investigated by plotting the fundamental diagrams. We have plotted the fundamental diagrams for the E17 motorway Ghent–Antwerp based on traffic measurements obtained by a camera during a morning rush hour. We have fitted a generic empirical speed-density relation to the traffic data.

The final component in the control loop is the actuator. Advanced traffic management systems can use a variety of traffic control measures that act as actuators for the traffic system. The controller provides the desired control policy to the actuators, which apply the control strategy to the traffic system. We have concluded this chapter with an overview of dynamic traffic control measures that can be applied in advanced traffic management systems.

The main contribution of this chapter to the state of the art is the development of an efficient interpolation scheme using a reference day. The presented interpolation method is numerically efficient such that it can be applied in on-line systems. By recomputing the reference day, slow changes in the traffic conditions, e.g., due to changing seasons, can be tracked. The reference day can be computed off-line.

Chapter 3

Metanet model

After the overview of traffic model classification in Section 2.1, this chapter discusses the METANET traffic flow model [96,110] in detail. The METANET model will be used in the simulation examples presented in Chapters 5 and 6.

Before discussing the METANET model, we provide a concise history of the second-order traffic flow models. In this overview of the history of the second-order traffic flow models, we provide the arguments of some of the adversaries and supporters of the second-order traffic flow model.

Next, we present the METANET model in its non-destination oriented form, which is useful in cases where traffic assignment is not explicitly considered. The non-destination oriented METANET model does not keep track of how the traffic traveling to a certain destination is assigned to the alternative routes to that destination. Instead, for every leaving link of every node, a turning rate describes the fraction of the total inflow of the node that uses that particular link. We will use the non-destination oriented METANET model in Chapter 5 for the simulation of a traffic corridor.

Since we will consider traffic assignment in Chapter 6, we present the destination oriented METANET model in Section 3.3. The destination oriented METANET model keeps track of how the traffic flow destined to a particular destination is assigned to the alternative routes to that destination. For every destination reachable from a node, splitting rates describe how the sub-flow in the node destined to that destination distributes over the links leaving the node. This model feature allows the traffic assignment of the traffic destined to a particular destination to be altered by updating the corresponding splitting rates. We conclude Section 3.3 with a summary of the advantages and the disadvantages of the METANET traffic flow model.

In the last section of this chapter, some notes regarding the identification of the model parameters are made.

3.1 A concise history of second-order traffic flow models

Before presenting the METANET traffic flow model in the next two sections, we provide a concise history of the second-order traffic flow model that is the basis of the METANET model.

The METANET model [96, 110, 133] is a second-order, deterministic, macroscopic traffic flow model that consists of the Payne model [117] discretized in both space and time with some model enhancements proposed by Papageorgiou added. The model presented by Payne was also independently described by Whitham in 1974 [150]. The model presented by Payne and Whitham, and thus also the METANET model, is based on the analogy between traffic flows and flows of fluids.

In 1995, Daganzo published a critique on the second-order fluid approximations of traffic flow [36]. Daganzo bases his critique mainly on the flaws in the analogy between traffic flows and flows of fluids such as e.g., contrary to a fluid particle, a vehicle is an anisotropic particle in the flow that mostly responds to the traffic situation ahead. According to Daganzo, other flaws in the analogy between traffic flows and flows of fluids are that the width of a shock wave in a traffic flow only encompasses a few vehicles and that the drivers have a personality, which remains unchanged during their journey. Using an example, Daganzo shows that negative speeds and negative flows on the motorway can result due to the intrinsic assumption of isotropy of the vehicles in the Payne model. In the conclusion of his paper, Daganzo suggests to use the first-order traffic flow model by Lighthill, Whitham [87] and Richards [123] instead of the higher order traffic flow models.

In 1998, Papageorgiou points out in [107] that the second-order models were developed in order to remove some of the deficiencies of the first-order model of Lighthill, Whitham and Richards. Where the first-order traffic flow model assumes an instantaneous adaptation of the drivers to a change in the traffic density on the motorway, the second-order Payne model is derived from a car following model that takes the reaction time of the drivers into account. This results in a dynamic equation for the mean speed containing a relaxation term. The impact of lane drops and of merging behavior at on-ramps can also be taken into account in this dynamic equation for the mean speed. Papageorgiou argues that, given the fact that the traffic flow models are empirical models, the problem of the negative speeds and the negative flows described by Daganzo can

simply be resolved by computing the speed on the motorway as the maximum of zero (or another minimal speed) and the speed provided by the model. In order to further counter the critique of Daganzo, Papageorgiou refers to a validation study where the accuracy of a variant of the model of Payne and Whitham is compared with the accuracy of the first-order traffic flow model by Lighthill, Whitham and Richards for a real-life motorway [32]. The accuracy of the second-order model of Payne and Whitham, which is the basis of the METANET simulation model, was found to be significantly better than the accuracy of the first-order model by Lighthill, Whitham and Richards.

In 1999, a new critique on the second-order traffic flow models was formulated by Heidemann [59], who claimed an internal inconsistency in the model of Payne and Whitham. However, in 2003, Zhang refuted this inconsistency [162]. In the same publication, Zhang stated that the smooth speed-density relation of the first-order traffic flow model is a good approximation of the traffic dynamics on the motorway although start-stop traffic cannot be modeled using the Lighthill, Whitham and Richards

first-order model. According to Zhang, the scattering of the measurements around the empirical speed-density relation (see e.g. Figure 2.6) cannot be explained by measurement noise alone, but results from finer traffic dynamics that are not captured by the first-order traffic flow model, a conclusion that was also drawn by Papageorgiou in [107]. Relaxation and anticipation are two finer traffic dynamics that are captured by the second-order traffic flow model. Relaxation and anticipation will be discussed in more detail in Section 3.2.

In 2000, Aw and Rascle presented a new second-order traffic flow model [5] that is based on the model presented by Payne and Whitham but that includes some adaptations to eliminate the inconsistencies discussed by Daganzo [36]. E.g., the Aw and Rascle model takes the anisotropy of the vehicles in the traffic flow into account and avoids in this way negative values of the speed and the flow. In 1999, Zhang presented a traffic flow study based on the Lighthill, Whitham and Richards model that takes the hysteresis phenomenon into account [160]. Zhang defines three kinds of traffic flows: acceleration, deceleration and equilibrium traffic flows. Due to the hysteresis phenomenon, the speed-density diagrams for each of these classes differ. Zhang states that it should be possible to model start-stop traffic if transitions between the different speed-density diagrams are considered. However, a working simulation model is not available yet.

As becomes clear from the discussion above, there are supporters and adversaries for the use of the second-order traffic flow model presented by Payne and Whitham. Since Papageorgiou claims that, for practical applications, most of the critique of Daganzo can be countered, and since the critique of Heidemann was countered by Zhang, we choose the METANET traffic flow model for the simulations in Chapters 5 and 6.

In the next two sections, we present the METANET model in both its non-destination oriented and in its destination oriented form.

Remark 3.1 It is important to note that the choice of the METANET traffic model presented in this chapter is not imposed by the model predictive control (MPC) framework discussed in Chapter 4 but by a trade-off between model accuracy (level of detail) and computational complexity for the case study that will be presented in Chapter 5. So other traffic models than the one presented here could also be used in an MPC framework, such as, e.g., a gas-kinetic model [62, 63, 69], a first-order [87, 123] or a second-order [117] macroscopic model or a microscopic [135] traffic simulation model such as they are implemented in Paramics, AIMSUN [6], FLEXSYT-II [99], HUTSIM, ... □

3.2 METANET motorway traffic flow model

In this section the non-destination oriented METANET traffic flow model is presented. In Section 3.3 the more general destination oriented METANET model will be presented. The non-destination oriented METANET traffic flow model is used when traffic assignment aspects (routing) are not taken under consideration. When traffic assignment is an issue, the destination oriented METANET traffic flow model is used. In Chapter 5, the non-destination oriented METANET model presented in this section will be used to simulate a case study.

The METANET model [96, 110, 133] is a second-order, deterministic, macroscopic traffic flow model that consists of the Payne model [117] discretized in both space and time with some model enhancements proposed by Papageorgiou added. A motorway network is represented by a directed graph with the links corresponding to the motorway stretches. Each motorway stretch has uniform characteristics, i.e., no on-ramps or off-ramps and no major changes in geometry. Where major changes occur in the characteristics of the motorway stretch or in the road geometry (on/off-ramp), a node is defined. Each link m is divided into several segments as presented in Figure 3.1.

Since the METANET model presented below cannot account for high speed vehicles that travel through a segment (i.e., that travel the length $l_{m,i}$ of the segment) in less time than the simulation time step ΔT_{sim}, the following inequality must hold for all segments i in all links m in order to guarantee numerical stability of the model [109, 133]:

$$l_{m,i} > v_{\text{free},m} \Delta T_{\text{sim}}, \tag{3.1}$$

where $v_{\text{free},m}$ is the free flow speed of link m. The interpretation of the free flow speed will be discussed below. A typical value for the discretization in

Symbol	Unit	Description
m		link index
i		segment index
l		simulation step counter
ΔT_{sim}	s	simulation step
$v_{\text{free},m}$	km/h	free flow speed in link m
$\rho_{m,i}(l)$	veh/km/lane	traffic density in segment i of link m at simulation step l
$v_{m,i}(l)$	km/h	mean speed in segment i of link m at simulation step l
$q_{m,i}(l)$	veh/h	traffic flow (intensity) in segment i of link m at simulation step l
n_m		number of lanes in link m
$l_{m,i}$	m	length of the i-th segment of link m
$q_{\text{in},m,i}(l)$	veh/h	traffic inflow into segment i of link m at simulation step l
$q_{\text{out},m,i}(l)$	veh/h	traffic outflow out of segment i of link m at simulation step l
τ_m	s	time constant of the speed relaxation term
$V[\rho_{m,i}(l)]$	km/h	equilibrium speed in segment i of link m at simulation step l as a function of the density $\rho_{m,i}(l)$
a_m		fitting parameter equilibrium speed function (fundamental diagram)
ν_m	km^2/h	speed anticipation term model parameter
κ_m	veh/lane/km	speed anticipation term model parameter
δ_{m_o}		merging term parameter
$\rho_{\text{crit},m}$	veh/km/lane	critical density of link m
$q_{\text{on},o}(l)$	veh/h	service rate of on-ramp o at simulation step l
N_m		highest segment index in link m
ϕ_m		weaving term parameter
$w_o(l)$	veh	queue length at on-ramp o at simulation step l
$D_o(l)$	veh/h	traffic demand at on-ramp o at simulation step l
$Q_{\text{cap},o}$	veh/h	on-ramp capacity of on-ramp o
$\rho_{\text{jam},m}$	veh/km/lane	jam density of link m
$r_o(k(l))$		metering rate at on-ramp o at control step $k(l)$
$Q_n(l)$	veh/h	total traffic flow entering (leaving) node n at simulation step l

Table 3.1: Table with the symbols used for the parameters and the variables of the METANET model. The second column in the table presents the units of the parameters and the variables.

Symbol	Unit	Description
I_n		set of all links entering node n
$q_{m,0}(l)$	veh/h	traffic flow leaving a node through link m
$\beta_n^m(l)$		turning rate (non-destination oriented model) at node n to link m
O_n		set of links leaving node n
$\rho_{m,i,j}(l)$	veh/km/lane	partial density of the traffic with destination j in segment i of link m at simulation step l
$\gamma_{m,i,j}(l)$		composition rate for the traffic with destination j in segment i of link m at simulation step l
J_m		set of destinations reachable through link m
$q_{m,i,j}(l)$	veh/h	partial flow with destination j in segment i of link j at simulation step l
$w_{o,j}(l)$	veh	partial queue length for the traffic with destination j at on-ramp o at simulation step l
$\gamma_{0,j}(l)$		composition rate of the demand on-ramp o
$Q_{n,j}(l)$	veh/h	total traffic with destination j entering node n at simulation step l
$\beta_{n,j}^m(l)$		splitting rate in node n for the traffic with destination j at simulation step l

Table 3.2: Table with the symbols used for the parameters and the variables of the METANET model (continued). The second column in the table presents the units of the parameters and the variables.

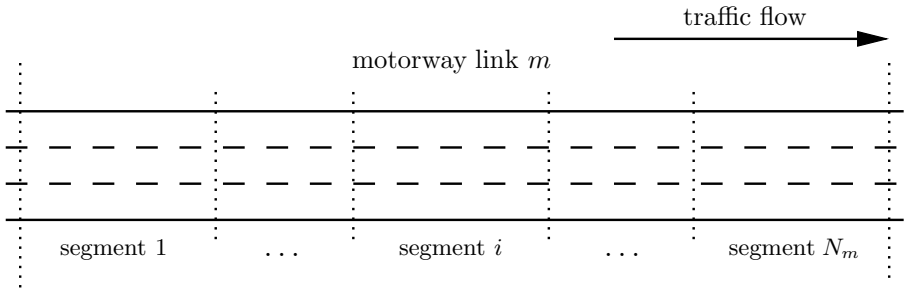


Figure 3.1: In the METANET model, motorway stretches are modeled by links. Every link is divided into segments. A segment is characterized by the couple (m,i) where m is the link index and i is the segment index in the link.

space (i.e., the length of the segments) is $l_{m,i} = 500\,\text{m}$. A typical value for the discretization in time is $\Delta T_{\text{sim}} = 10\,\text{s}$. For these choices of $l_{m,i}$ and ΔT_{sim}, combined with a typical free flow speed of $v_{\text{free},m} = 102\,\text{km/h}$ [95], the inequality (3.1) is satisfied and a sufficient safety margin is included.

The discrete METANET simulation model provides the traffic states of the segments at discrete time instants. The simulation counter l corresponds to the time instant $t = l\Delta T_{\text{sim}}$ and ΔT_{sim} is the time step used for the simulation of the traffic model.

The METANET model consists of link equations and node equations. The link equations describe the traffic behavior in the motorway stretches. The node equations describe the traffic behavior at bifurcation nodes in the network. First, the link equations for the non-destination oriented METANET model are presented followed by a discussion of the node equations.

Link equations

The METANET link equations describe the state of each segment i of a link m. The state of a segment at time $t = l\Delta T_{\text{sim}}$ is characterized by the following state variables:

- the *traffic density* $\rho_{m,i}(l)$ (veh/lane/km),
- the *mean speed* $v_{m,i}(l)$ (km/h), and
- the *traffic intensity* or *traffic flow* $q_{m,i}(l)$ (veh/h),

two of which are independent because of the fundamental relation (2.2). The model state equations that describe the evolution of the flow, the average speed and the density for every segment of the motorway are now discussed in detail.

The first model state equation describes the evolution of the traffic density in a segment and states the conservation of vehicles in every segment:

$$\rho_{m,i}(l+1) = \rho_{m,i}(l) + \frac{\Delta T_{\text{sim}}}{n_m l_{m,i}}[q_{\text{in},m,i}(l) - q_{\text{out},m,i}(l)], \tag{3.2}$$

where $\rho_{m,i}(l)$ is the traffic density in segment i of link m at simulation step l, $q_{\text{in},m,i}(l)$ is the inflow into the segment during the time interval $[l\Delta T_{\text{sim}}, (l+1)\Delta T_{\text{sim}})$, $q_{\text{out},m,i}(l)$ is the outflow of the segment during the same time interval, and n_m and $l_{m,i}$ are respectively the number of lanes in link m and the length of segment i of link m[1]. The outflow $q_{\text{out},m,i}(l)$ of a segment i in a link m is

[1]Note that since in a link the characteristics of the corresponding motorway stretch are constant, the number of lanes is the same in all segments of a link hence the single index in n_m. The length of the segments in a link m results from the discretization of the link in space and can vary over the segments hence the second index in $l_{m,i}$.

the inflow $q_{\text{in},m,i+1}(l)$ of the downstream segment $i+1$ in link m.
Equation (3.2) results from discretizing the conservation equation of the Payne model [117] in both time and space. This conservation law is similar to the conservation law from fluid dynamics [3] and states that the number of vehicles in a segment (the density) can only change due to a net inflow (positive or negative) of vehicles into the segment. The similarity between the conservation equations for vehicles and the conservation equation for fluids was already recognized by Lighthill and Whitham in 1955 [86, 87].

The second link equation describes the average speed in a segment. The average speed in segment i of link m at simulation step $l+1$ is given by:

$$\begin{aligned} v_{m,i}(l+1) = {} & v_{m,i}(l) \\ & + \frac{\Delta T_{\text{sim}}}{\tau_m}\big[V[\rho_{m,i}(l)] - v_{m,i}(l)\big] && \text{Relaxation} \\ & + \frac{\Delta T_{\text{sim}}}{l_{m,i}} v_{m,i}(l)\big[v_{m,i-1}(l) - v_{m,i}(l)\big] && \text{Convection} \\ & - \frac{\nu_m \Delta T_{\text{sim}}[\rho_{m,i+1}(l) - \rho_{m,i}(l)]}{\tau_m l_{m,i}[\rho_{m,i}(l) + \kappa_m]} && \text{Anticipation} \end{aligned} \tag{3.3}$$

where τ_m, ν_m, and κ_m are parameters that need to be fitted to traffic data using conventional identification and/or parameter estimation techniques[2]. In order to obtain a more realistic simulation of the dissolving of congestion, the average speed in a segment is bounded from below by $v_{\min}$. Hence, the average speed in a segment is computed as the minimum of the value given by (3.3) and $v_{\min}$. The physical interpretation of the parameters and terms in (3.3) will be discussed below. Three phenomena contribute to the change of the average speed in a segment: relaxation, convection and anticipation.

- Relaxation:
  The relaxation term states that the average speed in every segment tends to evolve towards a density dependent equilibrium value $V[\rho_{m,i}(l)]$. On the microscopic level all drivers in a segment tend to reach their desired speed. If the actual speed is lower (e.g., because there is/was an obstacle, congestion, ...) the driver will accelerate, if the actual speed is higher (e.g., because the desired speed is dropping rapidly due to onset of congestion, ...) the driver will apply the brakes. When this behavior is averaged over all drivers, we find a density dependent desired equilibrium speed for the segment. If the actual average speed in the segment differs from the equilibrium speed, the average speed will evolve towards the equilibrium speed in the segment. The larger the difference between actual

[2]The identification of the model parameters will be discussed in more detail in Section 3.4.

and equilibrium speed, the larger the drivers' actions and the larger the relaxation term. The time constant τ_m depends on the average drivers' swiftness. The larger τ_m, the slower the drivers will respond to a speed deviating from the equilibrium speed and the smaller the relaxation term. A typical value of τ_m is 18 s [81].

The fundamental diagram or the average speed in a segment in equilibrium is an empirical function of the traffic density and needs to be calibrated. Different formulas are suggested in the literature [81, 95, 110]. The METANET model uses the following expression for the relation between $V[\rho_{m,i}(l)]$ and the traffic density:

$$V[\rho_{m,i}(l)] = v_{\text{free},m} \exp\left(-\frac{1}{a_m} \left(\frac{\rho_{m,i}(l)}{\rho_{\text{crit},m}} \right)^{a_m} \right) . \tag{3.4}$$

The parameter $v_{\text{free},m}$ is the free flow speed or the average speed that a vehicle in a segment obtains when the traffic density in that segment goes towards zero (when the vehicle is the only vehicle in the segment and there are no interactions between vehicles). The parameter a_m is a model parameter. The parameter $\rho_{\text{crit},m}$ is the critical density or the traffic density corresponding to a maximal traffic flow in the fundamental diagram as presented in Chapter 1. In Figure 3.2 the empirical relation between the traffic density in the segment and the equilibrium speed given by (3.4) is plotted. Typical values of these parameters are [81]: $v_{\text{free},m} = 102$ km/h, $\rho_{\text{crit},m} = 33.5$ veh/km/lane and $a_m = 1.867$.

- Convection:
  The convection term in (3.3) takes into account the fact that vehicles entering a segment do not adapt their speed instantaneously. For instance, a driver traveling at high speed through segment $i-1$ will gradually slow down to the speed he feels comfortable with in segment i (equilibrium speed in i). During the braking process, the driver is actually going faster than his desired equilibrium speed. This way, the higher average speed in segment $i-1$ increases the average speed in segment i. The opposite is also true. Cars traveling at low speed in segment $i-1$ need time to accelerate towards the higher equilibrium speed in segment i. During this time, they are traveling at a lower speed, thus pulling the average speed in i down. The importance of the convection phenomenon is proportional to the difference in speed between the two segments since the higher the speed difference, the longer the vehicles will need to accelerate or decelerate and the greater the impact on the average speed in the segment. The effect of convection is also proportional to the average speed in the segment. The higher the speed, the longer it takes to brake or to accelerate due to the increased kinetic energy of the vehicle and thus the greater the effect of convection on the average speed. The impact of convection decreases with the length of the segment. The longer the segment, the longer the distance the vehicles drive at their desired speed

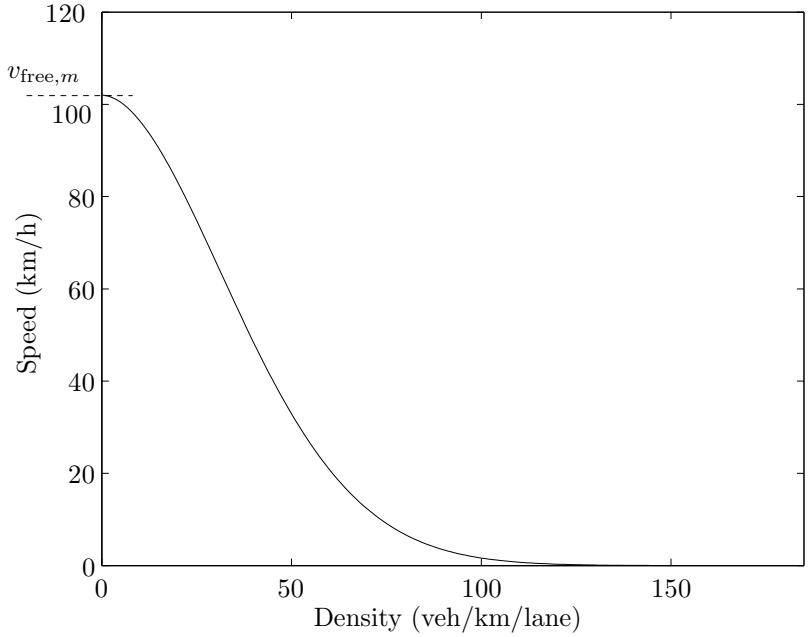


Figure 3.2: Illustration of the evolution of the average equilibrium speed in a segment with varying traffic density according to (3.4). The parameters $v_{\text{free},m}$ = 102 km/h, $\rho_{\text{crit},m}$ = 33.5 veh/km/lane and a_m = 1.867 were taken from [81]. The average equilibrium speed in a segment decreases with increasing traffic density in the segment as was also described in Section 2.2.4.

and the smaller the impact of convection on the average speed. Note that it is assumed here that convection is symmetrical, meaning that drivers will accelerate and decelerate at the same rate (on average).

- Anticipation:
  The anticipation term takes into account the fact that drivers are looking ahead and respond to the density they see in front of them. If a driver sees that the traffic density ahead is higher, he will anticipate and start to slow down (note the sign!) towards the equilibrium speed associated with the increased density ahead. Vice versa, if a driver sees that the traffic density ahead is lower than the current density, he will start to accelerate towards the equilibrium speed associated with this lower density. The impact of the anticipation term on the average speed in segment i is inversely proportional to the length of the segment since the drivers can only start to adapt their speed towards the equilibrium speed of segment

$i+1$ near the end of segment i. The longer segment i, the longer the vehicles will travel without the influence of the density in segment $i+1$. The anticipation term is proportional to the relative density difference between segments i and $i+1$ since according to (3.4) the higher the difference in density, the higher the difference in equilibrium speed between both segments will be. Note that we are working with the downstream segment $i+1$, resulting in an information stream opposite to the driving direction. The time constant τ_m in (3.3) describes the swiftness of the drivers' response as was the case in the relaxation term. The parameters ν_m and κ_m allow for further tuning of the anticipation term. Typical values for the parameters ν_m and κ_m are $60\,\mathrm{km}^2/\mathrm{h}$ and $40\,\mathrm{veh/lane/km}$ respectively [81].

In the special cases where on-ramp traffic merges into a motorway segment or when there is a lane-drop on the motorway, the METANET model provides extra merging and weaving terms in (3.3) to account for the dynamic phenomena that occur in these motorway configurations.

A *merging* term $\Delta v_{\mathrm{merge},m_o,1}(l)$ is added to (3.3) for the segments fed by an on-ramp o (Figure 3.3). This term is necessary since the slow-down of traffic in the vicinity of an on-ramp cannot be explained by the increased traffic density in the segment fed by the on-ramp alone. The merging term in the segment fed by the on-ramp accounts for the additional decrease in average speed due to disturbances caused by the lane changing of the merging on-ramp traffic.

The merging term has the following form:

$$\Delta v_{\mathrm{merge},m_o,1}(l) = -\left(\frac{\delta_{m_o}\Delta T_{\mathrm{sim}}}{l_{m_o,1}n_{m_o}}\right)\frac{q_{\mathrm{on},o}(l)v_{m_o,1}(l)}{\rho_{m_o,1}(l)+\kappa_{m_o}} \tag{3.5}$$

where δ_{m_o} is a tuning parameter for the merging term. A typical value of the parameter δ_{m_o} is 0.0122 [81]. Note that since a node is defined at locations where an on-ramp is present, an on-ramp always feeds into the first segment of the link leaving this node (Figure 3.3). Equation (3.5) represents the reduction in speed in the first segment of link m_o leaving the node due to the merging effects of the traffic flow $q_{\mathrm{on},o}(l)$ on on-ramp o.

The decrease in average speed in the first segment of link m_o due to the merging traffic is proportional to the flow $q_{\mathrm{on},o}(l)$ on the on-ramp. Since merging of on-ramp traffic becomes harder as the speed on the motorway increases, the impact of the (slower) merging cars on the average speed in the motorway segment becomes larger as the speed increases. This can intuitively be understood as follows: in a high-speed, low-traffic-density situation, the impact on the average speed of a fast car that is slowed down due to a lane changing maneuver of a merging car will be larger than in a low-speed, high-traffic-density situation where the reduction of the average speed due to the lane changing maneuver

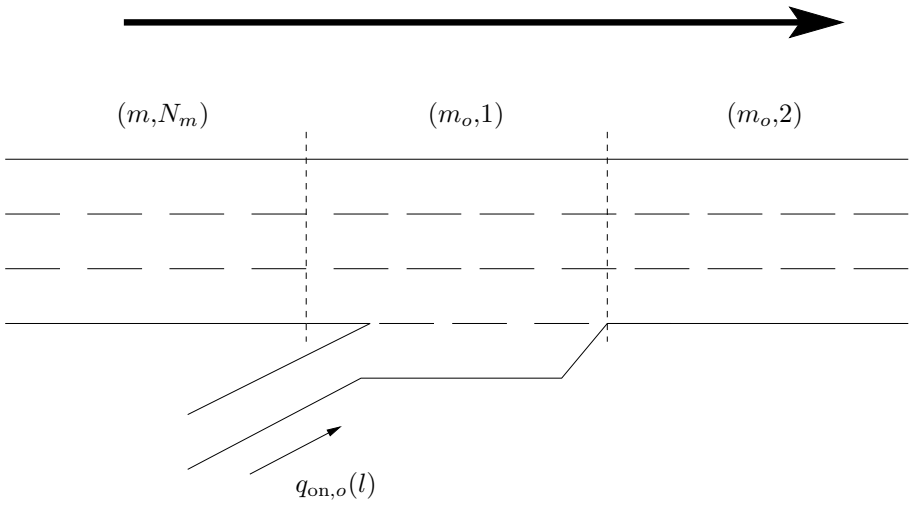


Figure 3.3: On-ramp traffic flow $q_{\text{on},o}(l)$ merging in the motorway traffic flow in the first segment of link m_o. The merging effects occur in segment $(m_o,1)$. The couples (m,i) represent the i-th segment of link m. The arrow denotes the direction of the traffic flow.

will be smaller. The merging term in (3.5) is also inversely proportional to the traffic density $\rho_{m_o,1}(l)$ in the first motorway segment of the link leaving the node. At first sight, this may seem like a contradiction. One might expect that the higher the traffic density is, the harder it becomes to merge into the traffic stream and the higher the impact of lane changing effects. However, since the average speed in the motorway segment at higher densities is already lower (see (3.4)), the net impact of lane changing on the average speed accounted for by (3.5) becomes smaller with increasing traffic density.

A *weaving* term $\Delta v_{\text{weav},m,N_m}(l)$ is added to (3.3) to account for the effects of a reduction in the number of lanes on the motorway. If the number of lanes on a motorway drops, vehicles driving on the disappearing lanes need to weave into the traffic on the other lanes. Since the number of lanes changes, a node is defined. The weaving occurs in the last segment before the lane drop (i.e., the last segment of the link feeding the node) since the vehicles can only leave this segment on the lanes that are also available in (the first segment of) the link leaving the node. A motorway configuration with a lane drop is presented in Figure 3.4. The locally increased traffic density in the last segment of the link feeding the node and the disturbances due to weaving traffic are accounted for by the following term:

$$\Delta v_{\text{weav},m,N_m}(l) = -\frac{\phi_m \Delta T_{\text{sim}}}{l_{m,N_m} n_m}\left(\frac{(n_m - n_{m+1})\rho_{m,N_m}(l)}{\rho_{\text{crit},m}}\right) v^2_{m,N_m}(l) \qquad (3.6)$$

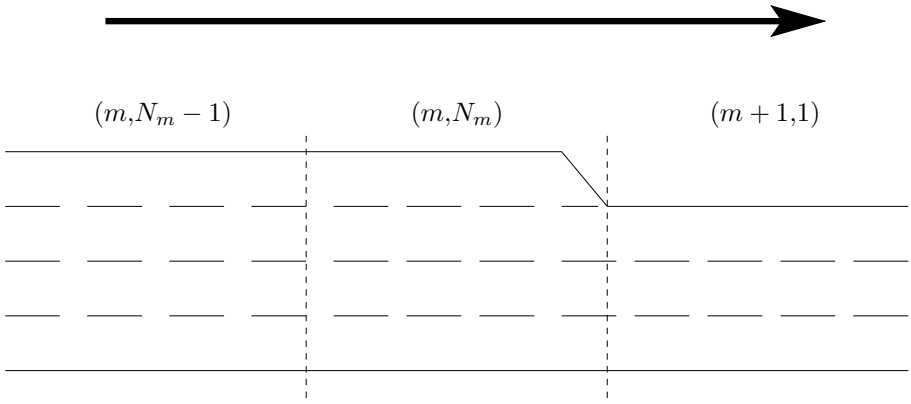


Figure 3.4: Weaving motorway traffic due to a lane drop between links m and $m+1$. The weaving effects occur in segment (m,N_m). The couples (m,i) represent the i-th segment of link m. The arrow denotes the direction of the traffic flow.

where ϕ_m is a tuning parameter. A typical value of the parameter ϕ_m is 2.98 [81]. The segment index N_m denotes the last segment of link m, which is the last segment before the lane drop as depicted in Figure 3.4. The indices m and $m+1$ denote the links before and after the lane drop respectively. The impact of weaving on the average speed is proportional to the number of lanes that is dropped $(n_m - n_{m+1})$. The higher the density $\rho_{m,N_m}(l)$ on the motorway, the more intense and disturbing the weaving process is. The importance of the weaving phenomenon is proportional to the square of the average speed $v_{m,N_m}(l)$ in the segment, expressing that drivers will slow down significantly when weaving occurs in high speed circumstances.

In Chapter 1 a fundamental relation between the traffic density, the average speed and traffic flow was discussed. Using this relation the traffic flow $q_{m,i}(l)$ in segment i of link m can be expressed in terms of the traffic density and the average speed in the segment:

$$q_{m,i}(l) = \rho_{m,i}(l) v_{m,i}(l) n_m \ . \tag{3.7}$$

This fundamental equation provides a relation between the three model states of the METANET model: the traffic density in a segment, the average speed in a segment and the traffic flow in a segment. Equation (3.7) is valid in every segment and at every simulation step. The relation between the states in (3.7) can be substituted in (3.2) to obtain a model with two independent states.

Since it can happen that the traffic demands at the network origins exceed the motorway capacity, queues are modeled at the borders of the simulated

motorway. If the traffic demand $D_o(l)$ at origin link o exceeds the service rate $q_{\text{on},o}(l)$ of the origin link, a queue is formed. The evolution of the queue length $w_o(l)$ at origin link o is given by:

$$w_o(l+1) = w_o(l) + \Delta T_{\text{sim}}(D_o(l) - q_{\text{on},o}(l)). \tag{3.8}$$

The service rate of the origin link is the minimum of the number of vehicles that *want to enter* and the number of vehicles that *can enter* the motorway. This leads to the following expression:

$$q_{\text{on},o}(l) = \min\Bigg[D_o(l) + \frac{w_o(l)}{\Delta T_{\text{sim}}}, \\ Q_{\text{cap},o} \min\left(1, \frac{\rho_{\text{jam},m_o} - \rho_{m_o,1}(l)}{\rho_{\text{jam},m_o} - \rho_{\text{crit},m_o}}\right)\Bigg] \ , \tag{3.9}$$

where $Q_{\text{cap},o}$ is the capacity of the origin link (veh/h) and m_o is the index of the link the origin link feeds into. The number of vehicles that want to enter the motorway through origin link o is determined by the traffic demand $D_o(l)$ and the length of the queue $w_o(l)$. The number of vehicles that can enter the motorway is determined by the capacity of the origin link $Q_{\text{cap},o}$ and by the properties (ρ_{jam,m_o} and ρ_{crit,m_o}) and the state ($\rho_{m_o,1}(l)$) of the motorway segment the origin link feeds into. The number of vehicles that can enter the motorway as a function of the density in the motorway segment is plotted in Figure 3.5.

As long as the traffic density on the motorway remains smaller than the critical density ρ_{crit,m_o}, the capacity $Q_{\text{cap},o}$ of the on-ramp is the limiting factor. If the density on the motorway becomes larger than the critical density, the number of vehicles that is able to enter the motorway decreases proportionally with the density on the motorway. The higher the traffic density on the motorway, the harder it becomes for vehicles to enter the motorway.

A special case of an origin link is an on-ramp with ramp metering. In case a motorway with a metered on-ramp o is modeled, the service rate given by (3.9) is replaced by the following equation:

$$q_{\text{on},o}(l) = \min\Bigg[D_o(l) + \frac{w_o(l)}{\Delta T_{\text{sim}}}, \\ Q_{\text{cap},o} \min\left(r_o(k(l)), \frac{\rho_{\text{jam},m_o} - \rho_{m_o,1}(l)}{\rho_{\text{jam},m_o} - \rho_{\text{crit},m_o}}\right)\Bigg] \ . \tag{3.10}$$

which accounts for the control implemented by the ramp metering setup (see Section 4.1) at on-ramp o. The metering rate $r_o(k(l))$ is discussed in more detail in Section 4.1.3.

Remark 3.2 It is important to note that the control time step ΔT_{ctrl} of the discrete-time ramp metering controller presented in Chapter 4 will in general be different from the simulation time step ΔT_{sim} used in the simulation

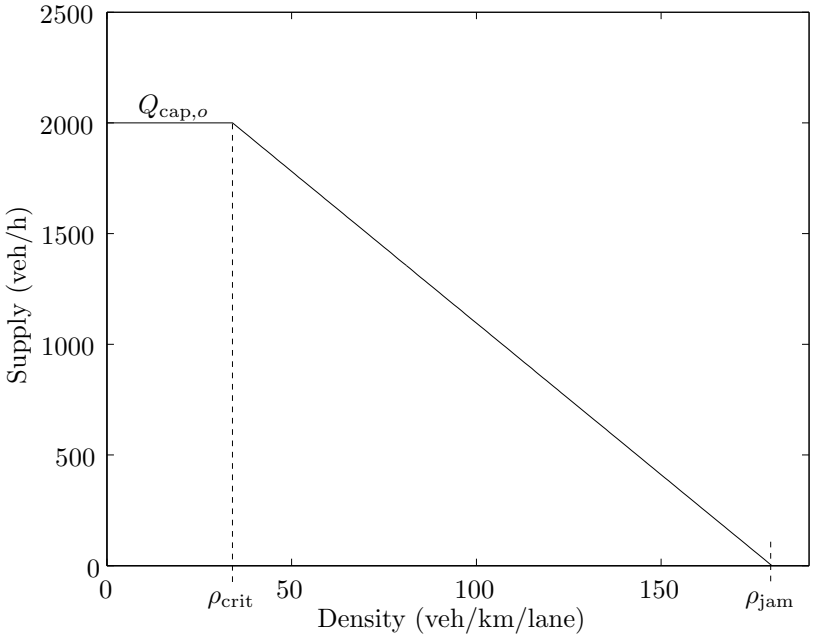


Figure 3.5: The maximum number of vehicles that can enter the motorway through an origin link as a function of the density on the motorway as described by (3.9). If the traffic density in the motorway segment fed by the on-ramp is below the critical density ρ_{crit}, the flow of vehicles that can enter the motorway is equal to the on-ramp capacity $Q_{\text{cap},o}$. For traffic densities in the motorway segment larger than the critical density ρ_{crit}, the flow of vehicles that can enter the motorway decreases proportionally with the traffic density in the motorway segment fed by the on-ramp.

(METANET) model. This difference in control and simulation time step is often neglected in the literature. In order to emphasize this difference, in this thesis the simulation step counter is denoted by l and the control step counter by k. The control step counter value corresponding to simulation step l is defined as $k(l)$, and the (first) simulation step corresponding to control step counter value k is defined as $l(k)$ i.e.,

$$k(l) = \left\lfloor l \frac{\Delta T_{\text{sim}}}{\Delta T_{\text{ctrl}}} \right\rfloor , \quad l(k) = k \frac{\Delta T_{\text{ctrl}}}{\Delta T_{\text{sim}}} \ ,$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to the real number x, and where it is assumed that ΔT_{ctrl} is an integer multiple of ΔT_{sim}. □

Node equations

The node equations of the METANET model [81, 133] describe on an aggregated level how the traffic behaves at the network nodes. The node equations discussed below account for:

- the distribution of the traffic flows in the network,
- the downstream propagation of the speed,
- the upstream influence of the traffic density.

The distribution of the traffic through the network is realized at the network nodes. Several links can enter a network node (Figure 3.6) and several links can leave the node (Figure 3.7). The node equations describing the distribution of the flows in the network provide the flows into the links leaving the nodes. Since the non-destination oriented METANET model does not take the destination of the traffic flows into account, the traffic flows with different destinations entering a node are all distributed in the same way over the links leaving the node.

A node has no storage capacity and thus the total traffic flow entering a node is always equal to the total traffic flow leaving the node. The total traffic flow entering (and leaving) a node n at simulation step l is given by:

$$Q_n(l) = \sum_{\mu \in I_n} q_{\mu,N_\mu}(l) \tag{3.11}$$

where I_n is the set of all links entering node n and where $q_{\mu,N_\mu}(l)$ is the traffic flow leaving the last segment N_μ of link μ at simulation step l.

The traffic flow leaving node n through link m (i.e., the traffic flow entering the first segment of link m) is given by:

$$q_{m,0}(l) = \beta_n^m(l) Q_n(l). \tag{3.12}$$

The turning rate $\beta_n^m(l)$ of node n at simulation step l is the fraction of the total inflow $Q_n(l)$ in node n that leaves the node through link m during simulation step l. The outflow out of node n that feeds into the first segment of link m is modeled as the outflow $q_{m,0}(l)$ of a virtual segment belonging to link m with segment index 0. The traffic flow leaving the network through an off-ramp can also be modeled using a segment leaving a node. The flow on the off-ramp is determined by a turning (leaving) rate and is computed by (3.12). The turning rates of the non-destination oriented METANET model are time dependent since the fraction of the total inflow in a node that chooses a particular link can vary with time (e.g., city inbound traffic during the morning rush hour and city outbound traffic during the evening rush hour).

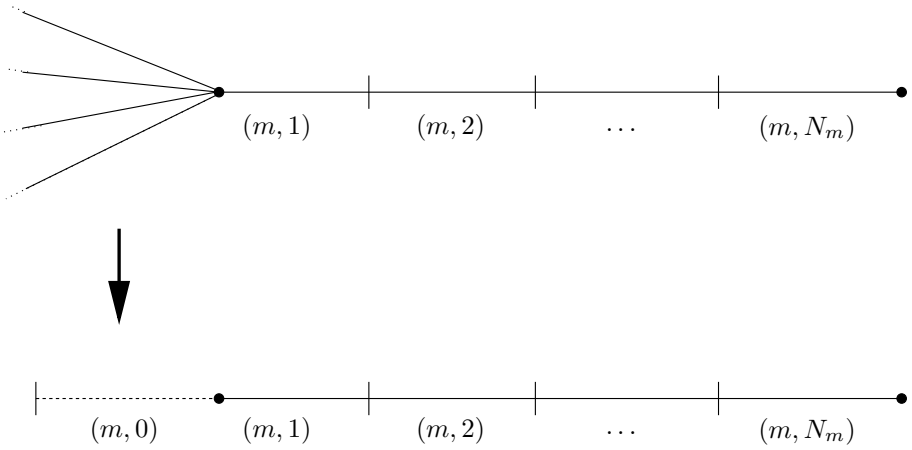


Figure 3.6: A node with one leaving link m and several entering links. The average speeds in the last segments of the entering links are aggregated in a virtual average speed $v_{m,0}(l)$ in a virtual segment according to (3.13). The virtual segment is prepended to the link m and has segment index 0.

The average speed in a segment is influenced by the average speed in the upstream segment as described by the convection term in (3.3). Since several links can enter a node, the speed in the first segment of a link leaving that node is influenced by the speeds in the last segments of the links entering the node as shown in Figure 3.6. The average speeds in the last segments of the links entering a node are aggregated in an average speed $v_{m,0}(l)$ in a virtual segment. The speed in the virtual upstream segment is computed as follows:

$$v_{m,0}(l) = \frac{\sum_{\mu \in I_n} v_{\mu,N_\mu}(l) q_{\mu,N_\mu}(l)}{\sum_{\mu \in I_n} q_{\mu,N_\mu}(l)} \quad . \tag{3.13}$$

The speed in the virtual upstream segment is used in the convection term in (3.3) for every first segment of each link leaving the node. As a special case of downstream propagation of the speed we mention virtual origin links. Virtual origin links are used as boundary conditions for the simulated motorways at the origins. A virtual speed $v_{m,0}(l)$ is assigned to a virtual origin link. This virtual speed propagates downstream to the first segments of the links leaving the origin node through the convection term in (3.3).

The upstream influence of the density at a node can be dealt with in a similar way as with the downstream propagation of the speed. The traffic density in a segment influences the average speed in the upstream segment as described by the anticipation term in (3.3). Since there can be several links leaving a node, the average speed in a segment upstream of a node is influenced by the traffic densities in the first segments of the links leaving the node as shown

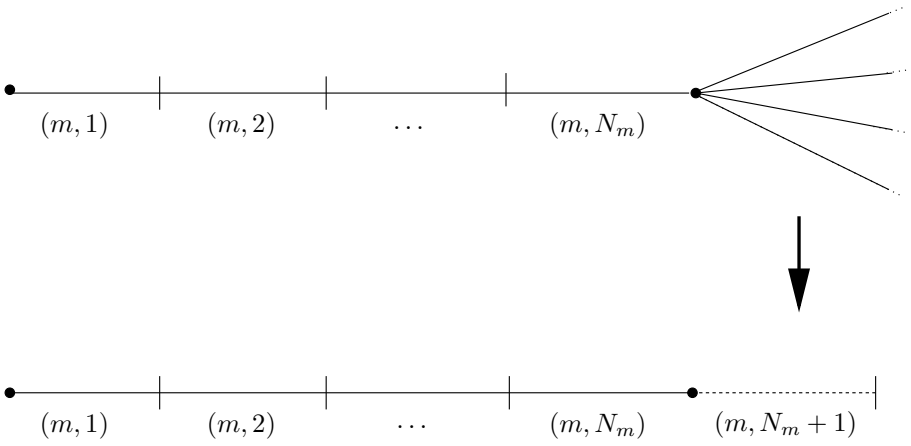


Figure 3.7: A node with one entering link m and several leaving links. The densities in the first segments of the leaving links are aggregated in a virtual density $\rho_{m,N_m+1}(l)$ in a virtual segment according to (3.14). The virtual segment is appended to the link m and has segment index $N_m + 1$.

in Figure 3.7. The densities in the first segments of all links leaving the node are aggregated as a density $\rho_{m,N_m+1}(l)$ in a virtual segment appended to the upstream link m. The traffic density $\rho_{m,N_m+1}(l)$ is computed as follows:

$$\rho_{m,N_m+1}(l) = \frac{\sum_{\mu \in O_n} \rho_{\mu,1}^2(l)}{\sum_{\mu \in O_n} \rho_{\mu,1}(l)}, \tag{3.14}$$

where O_n is the set of all links leaving node n. The density $\rho_{m,N_m+1}(l)$ of the virtual downstream segment is used in (3.3) in the anticipation term of the last segments of the links entering the node. This way, the upstream influence of the density at a node is modeled. Since congestion in the last segment of a link leaving the node has a big impact on the upstream traffic situation, a weighted average of the densities in the first segments of the links leaving the node underestimates the impact of the upstream influence of the traffic density if the other links are uncongested. Hence, we use a sum of squared densities in (3.14) in order to take into account that one congested leaving link can block the upstream segment even if the other leaving links are not congested. As a special case of a link leaving a node, we consider virtual destination links that are used as boundary conditions of the modeled motorways at the destinations. A virtual traffic density is assigned to the virtual destination links. This virtual traffic density $\rho_{m,N_m+1}(l)$ propagates upstream to the last segments of the links entering the destination node through the anticipation term in (3.3).

The combination of the link equations (3.2)-(3.10) and the node equations (3.11)-(3.14) that define the the non-destination oriented METANET model

allows for the simulation of a motorway network. The non-destination oriented form of the METANET model presented above is used if the routing behavior in the network is not explicitly considered. In Chapter 5 we will implement and simulate the non-destination oriented METANET model for a stretch of the E17 motorway Ghent–Antwerp. In the next section we consider the destination oriented METANET model that allows for a simulation of a traffic network where the routing of traffic flows between origins and destinations is explicitly considered. We will use this feature of the destination oriented METANET model in Chapter 6.

3.3 Destination oriented METANET model

A motorway network consists of motorway links that join or bifurcate at network nodes. This section deals with the destination oriented METANET model [96], which takes routing of traffic flows between origins and destinations explicitly into account. The destination oriented METANET model keeps track of how the traffic flow destined to a particular destination is assigned to the alternative routes to that destination. The destination oriented METANET model represents a traffic network as a directed graph with the links corresponding to motorway stretches. The model consists of link equations and node equations. For every destination reachable from a node, splitting rates describe how the sub-flow in the node and destined to that destination distributes over the links leaving the node. First the link equations are presented, followed by the node equations. We conclude this section with a summary of the advantages and the disadvantages of the METANET traffic flow model.

Link equations

Just as in the non-destination oriented model, a link in a motorway network is a motorway stretch connecting two nodes. Within a link the characteristics of the modeled motorway do not vary significantly. In a motorway network with multiple origins and destinations the traffic on a link can be composed of vehicles with different destinations. In order to be able to pin-point the traffic destined to the different destinations, we use partial densities and partial flows in the links. The use of these partial flows and partial densities in the link equations of the destination oriented METANET model is the main difference with the link equations of the non-destination oriented METANET model. The partial density $\rho_{m,i,j}(l)$ is defined as the density induced by the traffic traveling to destination j in the i-th segment of link m at simulation step l. The partial flow $q_{m,i,j}(l)$ is defined in a similar way.

The behavior of each of the segments in the motorway links can be described using the link equations presented below. These link equations are similar to the link equations of the non-destination oriented version of the METANET model

presented in Section 3.2. The difference with the non-destination oriented link equations is the use of the partial densities and the partial flows.

For each segment in the network, we can write down the following equation that describes the conservation of the number of vehicles in a segment:

$$\rho_{m,i,j}(l+1) = \rho_{m,i,j}(l) + \frac{\Delta T_{\text{sim}}}{n_m l_{m,i}} \left[\gamma_{m,i-1,j}(l) q_{m,i-1}(l) - \gamma_{m,i,j}(l) q_{m,i}(l)\right] \tag{3.15}$$

with $\gamma_{m,i-1,j}(l)$ the composition rate of the density (and of the flow) in segment $i-1$ of link m at simulation step l and $q_{m,i-1}(l)$ the flow out of segment $i-1$ of link m at simulation step l. Equation (3.15) can be written for every destination j that is reachable through link m. The composition rate is defined as:

$$\gamma_{m,i,j}(l) = \frac{\rho_{m,i,j}(l)}{\rho_{m,i}(l)} \quad . \tag{3.16}$$

The average speed $v_{m,i}(l+1)$ in segment i at simulation step $l+1$ is the same for all the sub-flows in the segment and can be computed using (3.3) from the non-destination oriented model. The total traffic density in a segment is given by:

$$\rho_{m,i}(l) = \sum_{j \in J_m} \rho_{m,i,j}(l) \tag{3.17}$$

where J_m is the set of destinations reachable through link m. In case of an on-ramp or a reduction in the number of lanes, merging or weaving terms are added to (3.3) as described in Section 3.2.

The partial flow $q_{m,i,j}(l)$ in segment i of link m induced by traffic with destination j at simulation step l is the product of the traffic density $\rho_{m,i,j}(l)$, the average speed $v_{m,i,j}(l)$ and the number of lanes n_m:

$$q_{m,i,j}(l) = \rho_{m,i,j}(l) v_{m,i,j}(l) n_m. \tag{3.18}$$

The total flow in a segment is the sum of all partial flows in that segment. The total flow can alternatively be computed using (3.7) and (3.17).

The dynamics of the queue at on-ramp o can be described in terms of partial queues as follows:

$$w_{o,j}(l+1) = w_{o,j}(l) + \Delta T_{\text{sim}} \gamma_{o,j}(l) \left[D_o(l) - q_{\text{on},o}(l)\right] \tag{3.19}$$

where $w_{o,j}(l+1)$ is the length of the queue at on-ramp o formed by the traffic with destination j at simulation step $l+1$, $D_o(l)$ is the demand at the on-ramp o, $\gamma_{o,j}(l)$ is the composition rate of the demand at the o-th on-ramp, and $q_{\text{on},o}(l)$ is the service rate of the on-ramp given by (3.9) (or by (3.10) in

case of a metered on-ramp). The total queue length is given by the sum of the partial queue lengths. The dynamics of the length of the total on-ramp queue are given by (3.8). Equation (3.19) assumes that the composition of the total queue is homogeneous. Hence, the traffic flow $q_{\text{on},o}(l)$ entering the motorway is assumed to be a homogeneous mixture of vehicles with a composition conform the composition rates $\gamma_{o,j}(l)$ of the demand.

Node equations

Where the link equations describe the state of the traffic in the traffic links, the node equations describe the relations between traffic flows in different links connected to a node. The node equations are important since they describe how traffic is routed through the network.

The major difference with the non-destination oriented METANET node equations is the use of splitting rates at the nodes instead of turning rates. While the turning rates in the non-destination oriented METANET model describe how the total inflow in a node is distributed over the links leaving the node, the splitting rates in the destination oriented METANET model describe how the sub-flows with a given destination that enter a node are distributed over the links leaving the node. In this way, the destination oriented METANET model can independently assign each sub-flow in the node to the links leaving the node. This is an important feature if traffic assignment is considered since, as we will see in Chapter 6, the sub-flows are assigned to the links taking both the destination and the state of the network into account. In a node n, the traffic for a destination j will distribute over the links through which destination j can be reached. The splitting rate $\beta^m_{n,j}(l)$ is the fraction of the total traffic flow $Q_{n,j}(l)$ entering node n and destined for destination j that leaves the node through link m. The following node equations express the relation between the traffic flows entering and leaving the nodes and describe how the traffic flows are routed through the network.

The total traffic volume entering node n and heading for destination j at simulation step l can be computed as:

$$Q_{n,j}(l) = \sum_{\mu \in I_n} q_{\mu,N_\mu}(l)\gamma_{\mu,N_\mu,j}(l), \tag{3.20}$$

where I_n is the set of all links entering node n, where $q_{\mu,N_\mu}(l)$ is the traffic flow leaving the last segment of link μ (i.e., segment N_μ), and $\gamma_{\mu,N_\mu,j}(l)$ is the composition rate of the traffic.

The total traffic flow leaving node n through link m at simulation step l is given by:

$$q_{m,0}(l) = \sum_{j \in J_m} Q_{n,j}(l)\beta^m_{n,j}(l) \tag{3.21}$$

where J_m is the set of reachable destinations through link m, O_n is the set of links leaving node n and where $\beta_{n,j}^m(l)$ is the splitting rate in node n at simulation step l.

The composition rate $\gamma_{m,0,j}(l)$ of the inflow into link m is given by:

$$\gamma_{m,0,j}(l) = \frac{\beta_{n,j}^m(l) Q_{n,j}(l)}{q_{m,0}(l)} \quad . \tag{3.22}$$

The combination of the link equations (3.15), (3.17), (3.3), (3.4), (3.5), (3.6), (3.7), (3.18), (3.19), (3.9), (3.10) and the node equations (3.20), (3.21), (3.22), (3.14) and (3.13) of the destination oriented METANET model allows for a simulation of the motorway network that takes traffic routing explicitly into account.

Advantages and disadvantages of the METANET traffic flow model

After the discussion of both the non-destination oriented and the destination oriented METANET traffic flow model, we conclude this section with a summary of the main advantages and disadvantages of the METANET model.

The main advantages of the METANET traffic flow model are:

- it has a relatively low computational complexity, which allows for real-time use of the model,
- it allows for a simulation of all kinds of traffic conditions (free, dense and congested traffic),
- the computational complexity of the METANET model is independent of the traffic conditions on the motorway,
- it includes terms to account for merging phenomena near on-ramps and weaving phenomena near lane drops, and
- in its destination oriented mode, the METANET model can take the routing of sub-flows between origins and destinations explicitly into account.

However, the METANET traffic flow model has also some disadvantages. The main disadvantages of the METANET model can be summarized as follows:

- it is a single-class model,
- it cannot take stochastic effects of traffic flows into account,
- it uses aggregated traffic variables, and
- it does not take overtaking and other multi-lane effects into account except for the merging and weaving effects.

3.4 Identification and parameter estimation

The traffic flow model presented in Sections 3.2 and 3.3 contains parameters such as $v_{\text{free},m}$, $\rho_{\text{crit},m}$, a_m, τ_m, ν_m, κ_m for the motorway links, and the parameters $Q_{\text{cap},o}$ and ρ_{jam,m_o} corresponding to the on-ramps. These parameters have to be estimated and calibrated on the basis of measurements [78, 110].

The available data to fit the model to are available through sensors as described in Chapter 2. As an illustration, the evolution of the flow and the average speed over a day as measured by a sensor on the E17 motorway Ghent–Antwerp in Belgium are presented in Figure 3.8. The METANET model parameters are collected in a parameter vector θ. The parameter estimation problem can then be formulated as a nonlinear least-squares problem where the set of model parameters θ is sought that minimizes the following cost criterion [89][3]:

$$I(\theta) = \sum_{l \in \mathcal{L}_{\text{meas}}} \Bigg[\sum_{(m,i) \in \mathcal{I}_{\text{m}}} \big(q_{m,i}(l) - \hat{q}_{m,i}(l)\big)^2 + \mu \sum_{(m,i) \in \mathcal{I}_{\text{m}}} \big(v_{m,i}(l) - \hat{v}_{m,i}(l)\big)^2 \Bigg], \tag{3.23}$$

where $\mathcal{I}_{\text{m}}$ is the set of index pairs (m,i) for all motorway segments and $\mathcal{L}_{\text{meas}}$ is the set of simulation steps for which measurements of the speed and the flow are available. The parameter μ is a weighting factor. In the identification cost function (3.23), the simulated flow $q_{m,i}(l)$ and the simulated average speed $v_{m,i}(l)$ in every segment are compared to the measured values $\hat{q}_{m,i}(l)$ and $\hat{v}_{m,i}(l)$ respectively. The squared error signals are summed over the samples present in the identification data set. This results in a nonlinear least-squares problem, which can be solved using a Gauss-Newton method [105, 115, 134]. For every motorway link modeled, the METANET model requires the identification of 6 parameters[4] ($v_{\text{free},m}$, $\rho_{\text{crit},m}$, τ_m, ν_m, κ_m and a_m). Since the computational complexity of the nonlinear least-squares problem increases strongly with the number of parameters, the computational complexity of the identification tends to become computationally very intensive for large traffic networks with many links and thus many link parameters to be estimated. In Chapter 5 we will keep track of the exact number of parameters that needs to be estimated for a real-life motorway stretch.

The value of $I(\theta)$ computed according to (3.23) is also a measure for the METANET model quality. The smaller the value of $I(\theta)$ is, the better the

[3] For the sake of simplicity, we used the least-squares cost criterion. However, other objective functions are possible, depending on the nature of the noise [89]. The choice of alternative cost criteria is subject to further research.

[4] The parameters for the merging and the weaving terms added to (3.3) and the on-ramp parameters in (3.9) at an origin link were not taken into account.

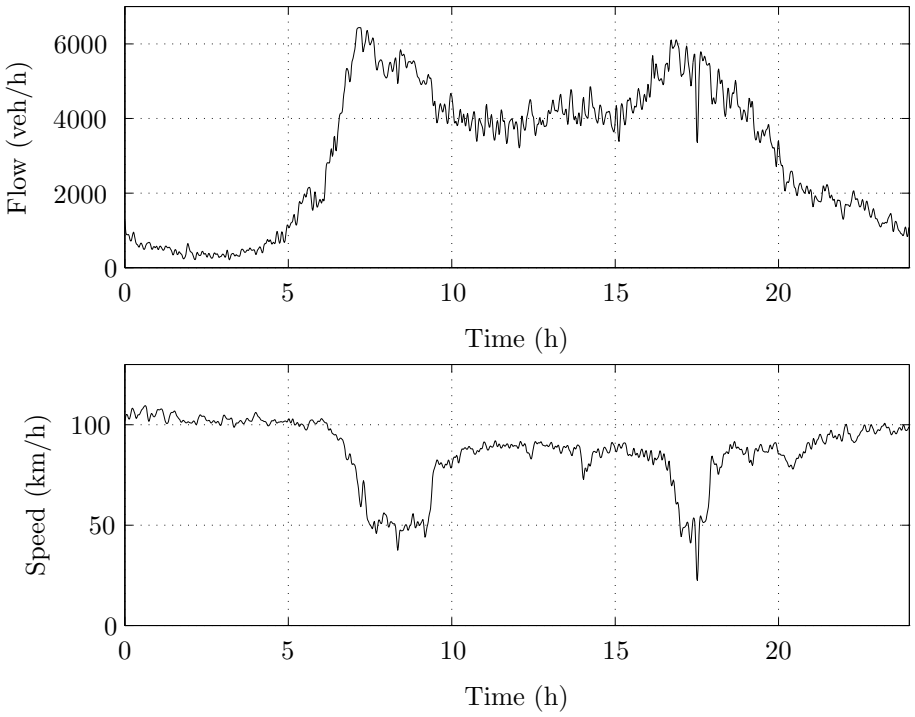


Figure 3.8: Evolution of the traffic flow and the average speed (average over a minute) as measured by camera detector CLOF on the E17 Ghent–Antwerp on Thursday, February 24, 2000. The camera is located near the beginning of the motorway stretch studied in Chapter 5 and presented in Figures 5.1 and 5.2. In the plot of the evolution of the flow over time we observe a morning and an evening traffic flow peak during the rush hours. Around 5 a.m. the traffic demand starts to increase until the maximal capacity of the motorway is reached around 7 a.m. and congestion sets in. We observe the onset of congestion as a reduction of the traffic flow on the motorway. In the plot of the average speed on the motorway over time the onset of congestion is much clearer. While the average speed is approximately constant during the whole day, the speed drastically decreases during congested traffic operation.

METANET model corresponds to the measurements of the real-life traffic system. The sensitivity of the METANET model quality to the different parameters has been investigated in the literature [32] and it was found that the difference in sensitivity of the METANET model quality $I(\theta)$ to the different parameters is large. Hence, only the subset of the parameters in θ with the highest sensitivity plays a significant role in the identification. The most important parameters of the traffic flow model were found to be $v_{\text{free},m}$ and $\rho_{\text{crit},m}$. These parameters need to be fitted to traffic data while the other parameter values can be carefully chosen from literature given the much smaller sensitivity of the METANET model quality to these parameters. This reduces the number of parameters that needs to be fitted and hence also the computational complexity of the identification.

Since not all parameters come to expression in every traffic operation mode, it is important to include traffic measurements from the different traffic operation modes in the identification dataset (i.e., the signals should be persistently exciting [89, 141]).

Even if we include data from the different traffic operation modes in our identification dataset, the resulting model is still not capable of mimicking the changing traffic behavior due to external, non-modeled influences and disturbances such as the weather, an obstruction on the motorway, ... Therefore, an adaptive framework is proposed in Chapter 6 where the traffic model is re-fitted to new traffic measurement data on a regular basis. This way, changes in the model parameters due to the non-modeled influences are tracked. By only updating the parameters with the highest sensitivity during the re-identification, the computational complexity of the re-identification process is kept as low as possible, which is important if the identification is performed in an on-line framework.

3.5 Conclusions

In this chapter, we started with the presentation of a concise history of the second-order traffic flow model. Next, we presented the METANET traffic flow model developed by Papageorgiou [96] in the non-destination oriented and in the destination oriented form. The non-destination oriented version of the METANET model can be used if traffic assignment or routing is not explicitly considered. If traffic assignment is taken under consideration, the more general destination oriented METANET model needs to be used. The METANET model is a deterministic, second-order, macroscopic traffic flow model that describes the relations between the aggregated variables traffic density, average speed and traffic flow. The METANET model is derived from the continuous second-order traffic flow model presented by Payne [117] by discretizing both

the independent variables time and space. Some extra terms were added to Paynes model by Papageorgiou in order to better model merging traffic at on-ramps and weaving traffic near lane-drops. We concluded the discussion of the METANET traffic flow model with a summary of the advantages and the disadvantages of the METANET model.

The METANET model is a grey box model that can for the main part be understood from traffic theory and from flow theory concepts such as e.g., the fundamental diagram and the conservation of vehicles. The parameters of the METANET model need to be estimated using traffic measurements. We presented a description of the identification process as well as a brief discussion of the sensitivity of the model performance to the different model parameters.

The main contribution of this chapter to the state of the art is the distinction that we make between the simulation time step ΔT_{sim} and the control time step ΔT_{ctrl}. In the literature, typically no distinction is made between the simulation time step and the control time step. However, the simulation time step ΔT_{sim}, which is imposed by the simulation model, is typically smaller than than the control time step ΔT_{ctrl}, which results from the dynamics of the traffic system. By distinguishing between the simulation time step and the control time step, the computational complexity of the controller can be reduced since the control signals need to be updated less frequently.

Chapter 4

Ramp metering controller design

Traffic control is often used as an alternative to the construction of new roads in an attempt to reduce congestion and to improve mobility. In this chapter ramp metering or admission control is discussed as a tool to improve traffic conditions on motorways. Ramp metering is a dynamic traffic responsive measure that can be implemented in the short term and relatively cheaply (unlike the construction of new motorways, extra lanes, ...) and that has been proven to yield good results in real-life situations [26, 98].

This chapter is organized as follows:
First, ramp metering as a traffic control measure is discussed in detail. Next, two ramp metering control algorithms are presented: ALINEA and model predictive control (MPC) based ramp metering control.

ALINEA is a PID-like control technique, which determines the metering rates such that the traffic state on the motorway is steered to a setpoint. Since ALINEA is widely adopted for the control of ramp metering set-ups around the world, it is used here as a benchmark to compare the performance of the MPC-based ramp metering controller to.

The major contribution presented in this chapter is the application of the MPC framework to ramp metering control. An MPC-based ramp metering controller is an advanced, model-based controller that steers the traffic state on the motorway to an optimal state that is defined by an objective function. A change in the control policy can easily be realized in an MPC framework by providing the controller with a new objective function. Furthermore, the MPC-based controller is able to take hard constraints into account when computing the

control signals. The receding horizon framework in which the MPC-based controller operates, introduces a form of feedback into the system that makes the controller more robust to measurement errors, external disturbances and changes in the environment. In the next chapter we will illustrate with some simulation results for a case study that the application of MPC-based ramp metering control results in smooth control signals (metering rates) and in an increased traffic system performance. The MPC framework also lends itself to the coordinated control of multiple on-ramps and even to the coordinated control of different traffic control measures.

4.1 Ramp metering

4.1.1 The fundamental diagram

As already discussed in Chapter 2, observations and measurements of traffic on motorways show that traffic approximately behaves according to what is known as the fundamental diagram in traffic flow theory [95]. In this section, the fundamental diagram is recapitulated in preparation of the discussion of the ramp metering concept in the next section.

A typical fundamental diagram is presented in Figure 4.1. In low traffic conditions, the traffic flow increases in a nearly proportional way with increasing traffic density. If the traffic density keeps increasing, the traffic flow starts saturating until a maximal flow is reached at the critical density ρ_{crit}. The maximal flow associated with the critical density ρ_{crit} is called the capacity q_{cap} of the motorway. A typical value of ρ_{crit} is 33.5 vehicles per kilometer and per lane [110]. The capacity q_{cap} of a three lane motorway is typically around 6000 vehicles per hour. Once the critical density is reached, traffic breakdown occurs and the traffic flow starts decreasing with further increasing traffic density. As soon as breakdown of the traffic flow at ρ_{crit} occurs, congestion sets in and traffic starts operating in congested regime. The congested regime is unstable in the sense that a perturbation that momentarily increases the density on the motorway will cause the traffic flow to decrease thus giving rise to an even larger traffic density. The traffic density in congested regime where the average traffic speed becomes zero or in other words where the traffic stalls is called the jam density ρ_{jam} (see Figure 4.1). A typical value of ρ_{jam} is 180 vehicles per hour and per lane [110]. Stable, free flowing traffic operation can only occur at densities below the critical density.

4.1.2 The ramp metering concept

A ramp metering set-up is implemented as a traffic signal that is placed at the on-ramp of a motorway as is schematically represented in Figure 4.2. The traffic

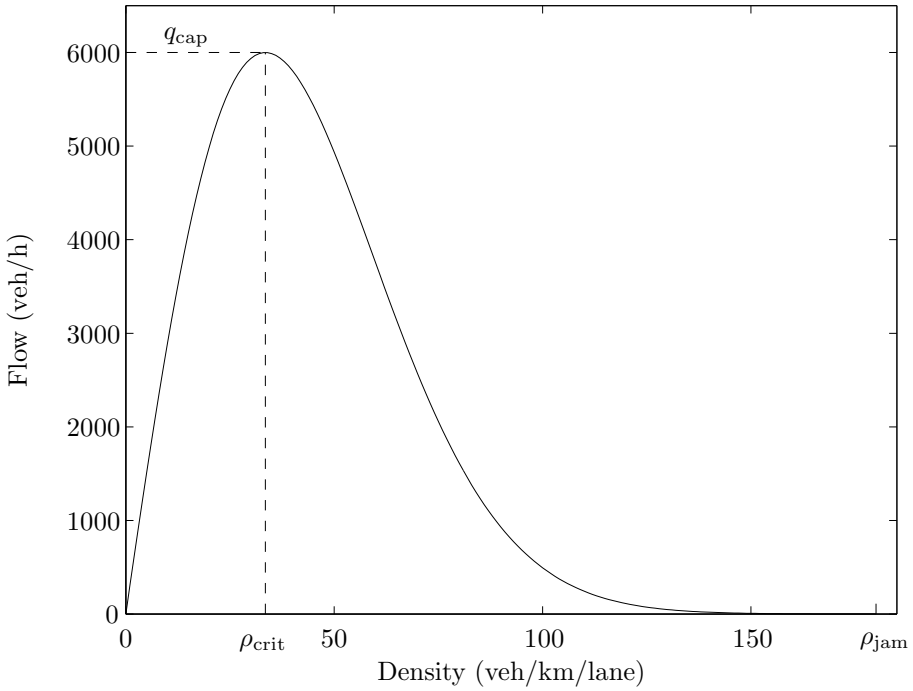


Figure 4.1: A flow-density fundamental diagram. As long as the traffic density on the motorway is smaller than the critical density ρ_{crit}, the traffic flow on the motorway increases with increasing traffic density. If the traffic density reaches the critical density, the flow is maximal and equal to the motorway capacity q_{cap}. If the traffic density further increases, the traffic flow on the motorway starts to decrease with increasing traffic density until the traffic stalls at the jam density ρ_{jam}.

signal controls the number of vehicles that is allowed to enter the motorway.

The ramp metering concept is based on the knowledge that the traffic on a motorway approximately behaves according to the fundamental diagram presented in the previous section. By keeping the traffic state on the motorway in the region of stable operation, ramp metering tries to prevent the occurrence of traffic breakdown and congestion [26, 159]. If the traffic density on the motorway tends to exceed the critical density ρ_{crit}, the ramp metering set-up can limit the flow of vehicles onto the motorway in order to keep the traffic density below the critical density, thus avoiding traffic breakdown and congestion. Note that whenever the traffic demand is larger than the number of cars that is allowed to enter the motorway, a waiting queue of vehicles is formed at the on-ramp. Hence, a trade-off exists between reducing congestion on the motor-

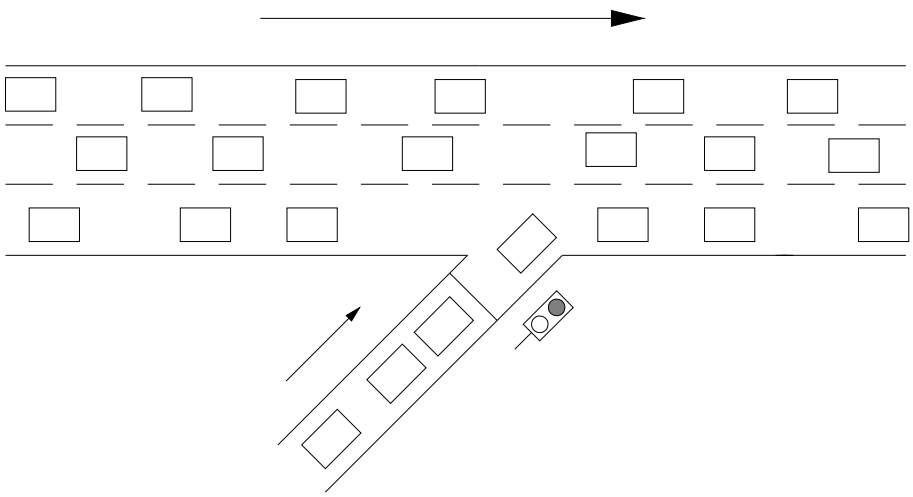


Figure 4.2: Schematic representation of ramp metering. The rectangles represent vehicles, and the arrows indicate the direction of the traffic flows. The traffic signal at the on-ramp controls the number of vehicles that enters the motorway. During ramp metering operation a queue can form at the on-ramp.

way and keeping the queue length at the on-ramp on an acceptable level. In general there is also a hard constraint on the maximal allowable queue length at the on-ramp since often it cannot be tolerated that the queue spills back and in that way blocks the urban traffic network.

The most important control objectives for ramp metering can be summarized as follows:

- maximize the throughput of the motorway,
- maximize the mean speed of the vehicles,
- improve the safety of traffic operation on the motorway by reducing shock waves,
- minimize the queue length at the on-ramps,
- minimize the waiting time at the on-ramps, and
- minimize the total time spent by all the vehicles in the network (on the motorway and in the queue at the on-ramps).

The preservation of a higher throughput on the motorway results in a smaller travel time and a higher mean speed for the vehicles on the motorway. However,

the travel time of the vehicles on the on-ramp increases due to the presence of the on-ramp queue. This illustrates that some of the control objectives can be conflicting. If the number of vehicles that is allowed to enter the motorway through the on-ramp is determined in an appropriate way (see Section 4.4), a trade-off is made between the queue length at the on-ramp and the traffic density on the motorway, and the total travel time is reduced despite the presence of the on-ramp queue.

The positive effects of ramp metering on the traffic situation are not only due to the prevention of traffic breakdown on the motorway. At least two other phenomena can also significantly contribute to a more optimal traffic operation [26, 64]:

1. Ramp metering reduces the disturbances on the motorway caused by on-ramp traffic. It is observed that vehicles often arrive in small platoons at an on-ramp [46]. These platoons can be formed, e.g., by a slow vehicle or by traffic signals near the on-ramp. Since one vehicle can merge into the motorway more smoothly than a small platoon of vehicles, ramp metering reduces the impact of the merging vehicles on the traffic flow on the motorway by only allowing one vehicle at a time to enter the motorway. This phenomenon is most relevant during traffic operation near the motorway capacity q_{cap}, where a small disturbance can result in breakdown of the traffic flow and in congestion.

2. Ramp metering can also prevent spill-back of congestion leading to blocked off-ramps (Figure 4.3). Congestion on a motorway spills back in the upstream direction. In this way an off-ramp upstream of a bottleneck could become blocked during congestion. The impact of a blocked off-ramp on the performance of a motorway increases with increasing traffic volume destined for that off-ramp. When the total time spent by all vehicles on the motorway and on the on-ramps is considered, the time gained by avoiding the occurrence of a blocked off-ramp using ramp metering generally outweighs the time lost in the queue at the on-ramp.

It has been reported in the literature [26, 98, 113] that ramp metering reduces the total travel time and improves the traffic flow in real-life situations.

4.1.3 Low-level implementation of ramp metering

A ramp metering set-up consists of a local controller that receives a metering rate and translates this metering rate to phase lengths for the traffic signal (see Figure 4.4). The metering rate for an on-ramp is a fraction of the on-ramp capacity corresponding to the maximal number of vehicles that is allowed to enter the motorway. The metering rate provided to the local controller is determined by a higher level controller. The choice of the higher level controller

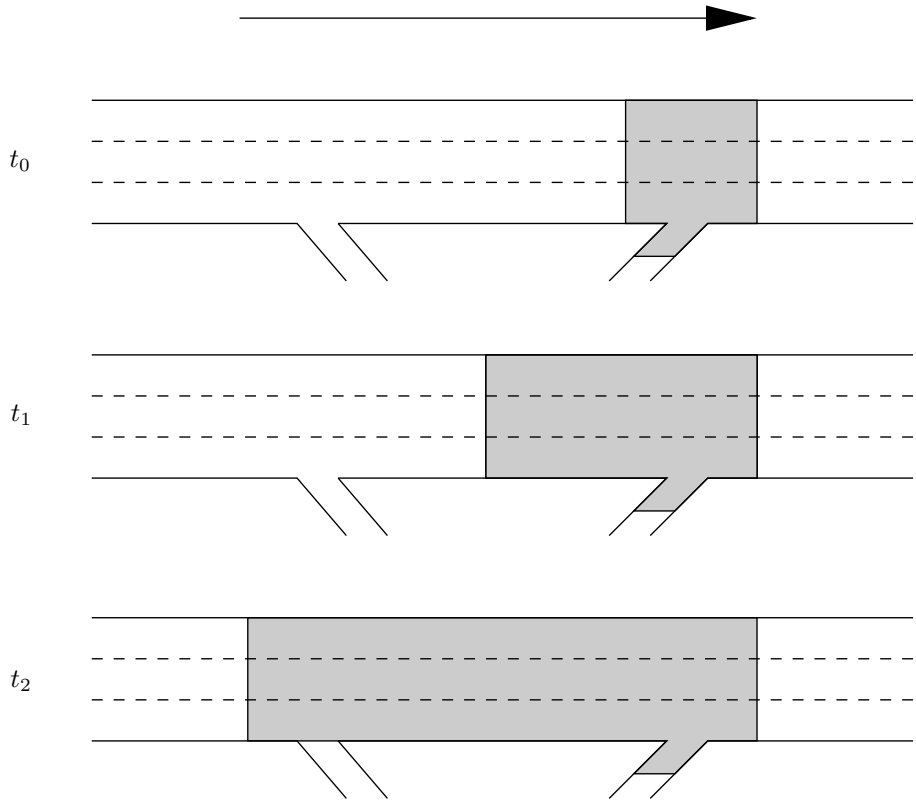


Figure 4.3: Illustration of spill-back of congestion on a motorway leading to a blocked off-ramp. The figure shows the same motorway at consecutive time instants t_0, t_1 and t_2. The arrow denotes the direction of the traffic flow. The congested (shaded) area spills further and further back as time progresses until it blocks the off-ramp at time instant t_2. Ramp metering can help to prevent or to delay this blocking by limiting the on-ramp traffic flow.

is independent of the low-level implementation of the ramp metering set-up. In this section, some issues concerning the local controller are presented while in the next section some ramp metering algorithms for the higher level controller in Figure 4.4 will be presented. In Sections 4.3 and 4.4, the ALINEA and the MPC-based controllers are discussed in more detail.

The setpoint of a local ramp metering controller is defined as a metering rate. Throughout this text a discrete-time controller is considered and the metering

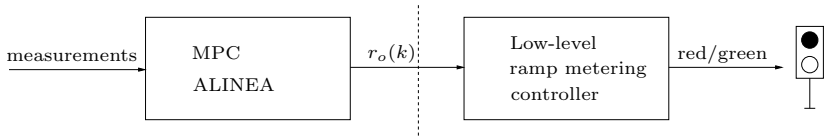


Figure 4.4: Schematic representation of ramp metering control. The high-level ALINEA or MPC-based controller (left) provides the metering rates (setpoints) to the low-level ramp metering controller (right). The low-level ramp metering controller implements the received metering rates by providing the traffic signal with the duration of the red and the green phase.

rate is defined as follows [80]:

$$r_o(k) \triangleq \frac{q_{\max,o}(k)}{Q_{\text{cap},o}} \ , \tag{4.1}$$

where k is the control sample counter, o is the on-ramp index, $q_{\max,o}(k)$ is the maximal flow of cars that is allowed to enter the motorway via on-ramp o during the period corresponding to control sample counter k and $Q_{\text{cap},o}$ is the capacity of on-ramp o. The metering rate $r_o(k)$ in (4.1) is by definition confined to the interval $[0, 1]$. However, in practice $r_o(k)$ is often limited to an interval $[r_{\min}, 1]$ in order to limit the waiting times for vehicles in the on-ramp queue during traffic operation under high traffic demand.

The type of the traffic signal used (red and green phase; red, green and amber phase; red, amber, flashing amber phase; ...) varies from country to country [76]. Depending on the traffic intensity a ramp is expected to carry, different low-level ramp metering implementations can be chosen. Three main types of ramp metering implementations can be distinguished [28]:

- Single-lane, one vehicle per green ramp metering:
  Only one vehicle is allowed to enter the motorway during the green phase. The local ramp meter realizes the required metering rate by appropriately choosing the length of the red (and amber) phase. The only current ramp metering set-up in Belgium (Figure 2.11) implements one vehicle per green ramp metering [76]. During ramp metering operation, the maximal on-ramp flow (and the maximal metering rate that can be realized) is determined by the minimal cycle length of the traffic signal. In the case of the ramp metering set-up on the E314 in Belgium, the minimal cycle length is 4 s, resulting in a maximal on-ramp flow of 900 vehicles per hour.

- Single-lane, multiple vehicles per green ramp metering:
  This strategy, also known as bulk metering, allows two or more vehicles to enter the motorway during each green phase. Bulk ramp metering

is implemented in order to increase the maximal on-ramp flow during ramp metering. The most common implementation allows two vehicles per green phase. It needs to be noted that allowing two vehicles instead of one during the green phase does not double the maximal on-ramp flow since the duration of the green phase needs to be increased in order to accommodate the second vehicle. However, despite the increased duration of the green phase, the duration of one single-lane, two vehicles per green ramp metering signal cycle is shorter than the duration of two single-lane, one car per green ramp metering signal cycles, which results in an increased maximal on-ramp flow. Allowing more than 2 vehicles per green phase is also possible but this reduces the capability of ramp metering to break up on-ramp platoons, a positive side-effect of ramp metering described in Section 4.1.2.

- Dual-lane metering:
  If sufficient room is available, an on-ramp with two lanes can be used. Both lanes are provided with a traffic signal. Depending on the implementation, the cycles of both traffic signals may or may not be synchronized. The advantage of dual lane metering is that a higher metered flow can be realized than in the single-lane, one car per green implementation.

Besides sufficient room to store the queue at the on-ramp as mentioned in Section 4.1.2, sufficient room between the head of the queue and the location where the vehicles merge into the motorway traffic should be foreseen. Indeed, the speed of the vehicles at the traffic signal is (nearly) zero and in order not to disrupt the traffic on the motorway while merging, the on-ramp vehicles need sufficient space to accelerate to the average speed on the motorway.

4.2 Ramp metering algorithms

There exist several ramp metering algorithms to determine the appropriate metering rates that are provided to the ramp metering set-up (i.e., the local controller in Figure 4.4) [54, 55]. Fixed time or non-traffic-responsive ramp metering [149] provides a fixed metering rate to the ramp metering set-up. Demand-capacity metering [108] and ALINEA [112, 113] are two local traffic responsive ramp metering methodologies which compute metering rates based on measurements of the traffic state in the vicinity of the metered on-ramp. ALINEA is a PID-like, feedback control methodology [73, 97] that is widely in use today. METALINE [111] is a coordinated feedback control methodology that can be used to coordinate control of several ramp meters. FLOW [71] is a heuristic area-wide coordinated control strategy. Optimal control based ramp metering [79, 80] and model predictive control (MPC) based ramp metering determine the appropriate metering rates by optimizing an objective function.

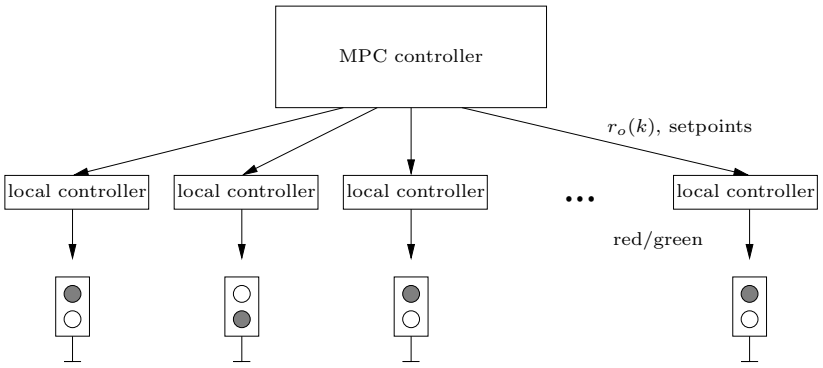


Figure 4.5: Schematic representation of MPC-based coordinated control. The MPC controller provides a metering rate $r_o(k)$ for every metered on-ramp o. The local controller transforms these metering rates into phase lengths of the traffic signal.

Neural networks and fuzzy control concepts have also been applied to the ramp metering control problem [65, 157].

In this thesis, MPC-based ramp metering control is presented as an advanced control technique to determine the appropriate metering rates for a ramp metering set-up. Since it is widely in use, the ALINEA feedback control methodology is presented in Section 4.3 in order to provide a standard control technique for ramp metering to compare the MPC-based control approach to. MPC-based control provides many attractive features. It can adapt to changes in the system and the environment and it is robust for measurement errors and disturbances. The desired behavior of the MPC-based ramp metering controller can be defined as an objective function. MPC is able to take hard constraints (e.g., on the maximum number of cars waiting at a given on-ramp) into account and, if required, MPC is able to keep the system behavior as close as possible to the constraints without violating them. As the main disadvantages of MPC-based ramp metering we note that MPC-based ramp metering control is more computationally intensive than some other methods such as e.g., ALINEA, and it requires a traffic simulation model of the motorway network that is to be controlled. A more detailed discussion of MPC-based ramp metering is provided in Section 4.4.

The MPC-based control of a single, locally controlled on-ramp can be extended to the *coordinated* dynamic control of several on-ramps in a motorway network as presented in Figure 4.5. The coordination of the metering rates of the different on-ramps ensures that the control actions taken at different locations in the network reinforce rather than cancel each other. In this way, coordination of ramp metering often leads to better results than the combination of multiple

independently locally controlled ramp metering set-ups. Coordinated MPC-based control of ramp metering will be applied to a test case in Chapter 5. The MPC-based control framework can also be applied to a combination of ramp metering and one or more of the traffic control measures presented in Section 2.3. This leads to MPC-based *integrated* traffic control. Since the focus of this thesis is on ramp metering control, integrated control of motorway traffic is beyond the scope of this text. For more information on integrated control using various traffic control measures the interested reader is referred to [22,58,79,80].

Remark 4.1 It is important to note that the control time step ΔT_{ctrl} of a discrete-time ramp metering controller will in general be different from the simulation time step ΔT_{sim} used in the traffic simulation model. This difference in control and simulation time step is often neglected in the literature. The simulation step ΔT_{sim} generally results from model specific properties such as e.g., (3.1) for the METANET model presented in Chapter 3 while the control time step ΔT_{ctrl} is typically determined by the dynamics of the controlled traffic system. In order to emphasize this difference, we will denote the simulation step counter by l and the control sample counter by k. We also define $k(l)$ as the control sample counter value corresponding to simulation step counter l, and $l(k)$ as the (first) simulation step counter value corresponding to control sample counter k i.e.,

$$k(l) = \left\lfloor l\frac{\Delta T_{\text{sim}}}{\Delta T_{\text{ctrl}}} \right\rfloor , \quad l(k) = k\frac{\Delta T_{\text{ctrl}}}{\Delta T_{\text{sim}}} ,$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to the real number x, and where we assume that ΔT_{ctrl} is an integer multiple of ΔT_{sim}. □

4.3 ALINEA

ALINEA[1] is a PID-like feedback control methodology for ramp metering developed by Papageorgiou *et al.* [112]. A schematic representation of ALINEA is given in Figure 4.6. The ALINEA controller tries to maintain the density on the motorway equal to a preset value or setpoint $\hat{\rho}$. By choosing the value of $\hat{\rho}$ smaller than or equal to the critical density ρ_{crit}, the ramp metering concept described in Section 4.1.2 is implemented. If $\hat{\rho}$ is chosen equal to the critical density ρ_{crit}, the corresponding traffic flow on the motorway is maximal and equal to the capacity q_{cap}. This results in a controller that optimizes the traffic flow on the motorway. If $\hat{\rho}$ is chosen smaller than the critical density ρ_{crit}, a more conservative controller results: the traffic flow on the motorway

[1] ALINEA is the acronym for "Asservissement linéaire d'entrée autoroutière", which could be translated as "Linear ramp metering control".

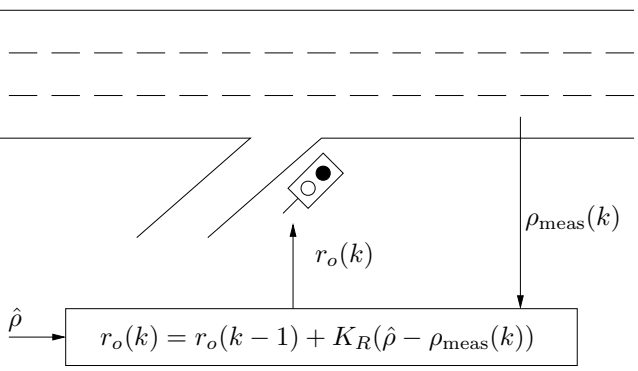


Figure 4.6: Control scheme for ALINEA-based ramp metering control. The traffic density downstream of the on-ramp is measured and provided to the ALINEA-based controller. The measured traffic density is compared to the desired traffic density $\hat{\rho}$ (the setpoint). Based on the difference between the desired and the measured traffic density on the motorway, the metering rate $r_o(k)$ that is provided to the on-ramp is adjusted. The parameter K_R is the controller gain.

corresponding to the setpoint density $\hat{\rho}$ is smaller and the queue at the on-ramp will be longer. However, due to the irregular nature of traffic flows, a more conservative controller with a safety margin between the setpoint (the desired density on the motorway) and the critical density at which breakdown of the traffic flow occurs, is often used in practice. In this way, it is avoided that small fluctuations of the traffic flow lead to unstable behavior. In discrete-time, ALINEA's control law reads:

$$r_o(k) = r_o(k-1) + K_R(\hat{\rho} - \rho_{\text{meas}}(k)), \tag{4.2}$$

where $r_o(k)$ is the metering rate at control sample step k, K_R is a positive constant, $\hat{\rho}$ is the setpoint and $\rho_{\text{meas}}(k)$ is the measured density on the motorway corresponding to controller step k. The metering rate $r_o(k)$ is by definition confined to the interval $[0, 1]$. In practice $r_o(k)$ is often limited to an interval $[r_{\min}, 1]$ in order to limit the waiting times for vehicles in the on-ramp queue during traffic operation under high traffic demand. The parameter K_R is determined heuristically. ALINEA's control results are insensitive to K_R for a wide range of values [54].

The ALINEA control law presented in (4.2) is an integrating feedback controller that steers the traffic density on the motorway towards a preset density. Indeed,

since the metering rate in (4.2) is proportional to the maximal on-ramp inflow into the motorway according to (4.1), and since a change of the traffic density on the motorway results from a net inflow into the motorway over time, (4.2) represents an integrating feedback law. If the traffic density on the motorway becomes too high (larger than $\hat{\rho}$), the metering rate is reduced and vice versa. The maximum number of vehicles allowed to enter the motorway through the metered on-ramp can be computed by (4.1). In steady state conditions, this type of integrating controller reaches the desired setpoint exactly (provided that the traffic demands are large enough to reach $\hat{\rho}$). However, if the traffic demands are very high, the metering rates provided by (4.2) may result in a very long queue at the on-ramp. Therefore, ALINEA-based ramp metering controllers are often implemented such that the provided metering rate can be overridden and set to 1 if the on-ramp queue length exceeds a threshold length [76]. If the queue length drops below the threshold, the ALINEA controller becomes active again. A main advantage of an ALINEA-based controller is its low computational complexity.

Note that the control law (4.2) is expressed in terms of the traffic density on the motorway. The traffic density $\hat{\rho}$ is used to define the desired setpoint since according to the fundamental diagram in Figure 4.1, a measurement of the traffic density uniquely defines the traffic regime on the motorway[2]. Note that in the discussion of the fundamental diagrams in Section 2.2.4 (see Figure 2.8), we already saw that for a measured value of the traffic flow on the motorway, there generally exists a traffic regime in free flow (low traffic density) and a traffic regime in congestion (high traffic density). Hence, a measurement of the traffic flow on the motorway alone does not provide sufficient information about the traffic state to the controller.

An additional advantage of defining the setpoint as a desired traffic density is the relative insensitivity of the critical density to external factors such as e.g., the weather as opposed to the sensitivity of the traffic flow [74].

The main advantages of the ALINEA ramp metering algorithm can be summarized as follows:

- ALINEA is computationally inexpensive,
- it is fairly insensitive to external influences such as e.g., the weather, and
- it requires only two parameters to tune: the setpoint $\hat{\rho}$ and the controller gain K_R.

The main disadvantages of the ALINEA ramp metering algorithm are the following:

[2]For the sake of simplicity of the exposition, hysteresis was not considered in the fundamental diagram in Figure 4.1 (see Remark 4.2).

- ALINEA is a local controller and as such only takes the traffic density on the motorway near the on-ramp into account,
- it cannot handle constraints, and
- it is hard to determine the ALINEA setpoints that achieve a certain control objective.

ALINEA is a ramp metering methodology that is widely adopted and that has been reported to improve traffic operation in terms of an increased throughput of the motorway and reduced total travel time [54, 76]. In the next section we will discuss an MPC-based approach to ramp metering. In Chapter 5 the performance of the ALINEA-based and the MPC-based ramp metering controllers will be investigated and compared for the E17 motorway case study.

Remark 4.2 For the sake of simplicity of the exposition, we have neglected hysteresis effects in the fundamental diagram used to discuss ramp metering (see Figure 4.1). It has been observed that in order to obtain free flowing traffic after congestion has set in, the traffic density must drop a little below the critical density ρ_{crit} before the motorway can operate at its full capacity q_{cap} again. In case this hysteresis effect is considered, a measurement of the traffic density does not necessarily uniquely determine the traffic state and the corresponding traffic flow as was the case without hysteresis. Although it was not explicitly considered in the exposition of ramp metering, hysteresis was taken into account in the simulation examples presented in Chapter 5 and 6. □

4.4 Model predictive control

Model predictive control (MPC) [25, 82, 94] is a very popular controller design method in the process industry. The MPC techniques known from process industry can be applied to the traffic control measures presented in Section 2.3. In this thesis, the focus is on ramp metering as a technique to control traffic on motorways. Hence, this section discusses how the MPC approach can be applied to ramp metering (see also [58]).

4.4.1 General description of MPC

Model predictive control [25,82,94,131] was pioneered simultaneously by Richalet *et al.* [122], and Cutler and Ramaker [33]. Since then, MPC has probably become the most applied advanced control technique in the process industry and many papers report successful applications [121]. The main ingredients of MPC

are that it is an on-line control approach in which a model is used to predict the future behavior of the system for a given input sequence and in which a cost criterion is optimized subject to constraints on the inputs, the outputs and/or the states. In addition, MPC uses a receding horizon strategy.

MPC provides many attractive features that can be summarized as follows:

- it can adapt to changes in the system and the environment,
- additional constraints (e.g., on the maximum number of cars waiting at a given on-ramp) can easily be taken into account and, if necessary, MPC can keep the system behavior as close as possible to the constraints without violating them,
- it is robust for measurement errors and disturbances,
- the desired behavior of the controller can conveniently be defined as an objective function, and
- since MPC is applicable to multi-variable systems, the MPC approach can easily be extended from the control of a single on-ramp to coordinated or even integrated ramp metering.

The main disadvantages of MPC-based ramp metering are:

- MPC is more computationally intensive than some other methods such as e.g., ALINEA,
- the required time for the computation of the metering rates is not fixed,
- it requires a traffic model of the motorway network that is to be controlled, and
- the inputs for the traffic model (e.g., the traffic demands, splitting rates, ...) need to be estimated for the prediction horizon.

In the MPC receding horizon framework, we define a prediction horizon N_p and at each control sample step k the control signals (metering rates) for the time period $[k\Delta T_\mathrm{ctrl}, (k + N_\mathrm{p})\Delta T_\mathrm{ctrl})$ are determined by minimizing an objective function over this period. The objective function will be discussed in Section 4.4.2. The controller time step ΔT_ctrl is the rate at which the control signals are updated. A typical value of the controller time step ΔT_ctrl for ramp metering control is 1 min [22, 110]. As already mentioned in Section 4.2, the controller rate ΔT_ctrl and the simulation rate ΔT_sim generally differ[3].

[3] A typical value for ΔT_sim for e.g., the METANET model is 10 s (see Section 3.2).

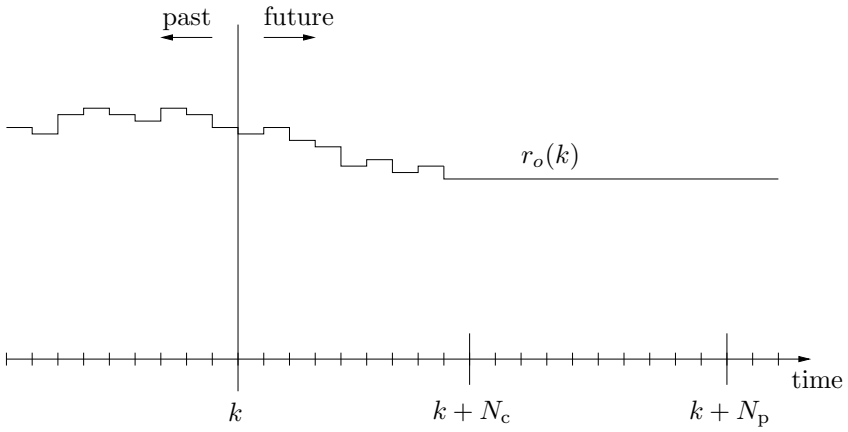


Figure 4.7: Schematic representation of the MPC concept. At control step k the metering rates for the whole prediction horizon N_p are determined. The metering rates are only allowed to change during the control horizon N_c. Only the first computed metering rate $r_o(k)$ is applied to the traffic system after which the prediction and the control horizon are shifted one control sample forward and the whole process starts all over again.

During the optimization, the objective function is evaluated based on a prediction of the future traffic behavior of the studied traffic system. This future traffic behavior is simulated using a traffic model. The initial state of the traffic model is determined by the traffic measurements at the start of the prediction horizon and the input of the traffic model is an estimation of the traffic demands during the prediction horizon[4]. The choice of the traffic model depends on the application e.g., the level of detail required by the objective function, the computational complexity, ... and is not imposed by the MPC controller design method. In Chapter 5, the METANET traffic flow model presented in Chapter 3 will be used.

There is in general a strong relation between the computational complexity of an optimization problem and the dimensionality of the parameter space. This relation is also known as the curse of dimensionality [47]. A prediction horizon of length N_p results in an optimization problem with rN_p parameters (metering rates) to be determined, where r is the number of metered on-ramps. In order to reduce the computational complexity associated with the optimization of the control signals in the MPC framework, the control horizon N_c ($N_c \leq N_p$) is defined (see Figure 4.7). The metering rate $r_o(k)$ is only allowed to change during the control horizon period $[k\Delta T_{\text{ctrl}}, (k + N_c)\Delta T_{\text{ctrl}})$, after which it is taken to remain constant. By defining the control horizon N_c, the number of

[4]In case a traffic network is simulated, some information on the routing through the network might also be required. E.g., in the case the destination oriented METANET model from Chapter 3 is used, the splitting rates are also needed as an input of the model.

parameters to be determined by the optimization algorithm is reduced from rN_{p} to rN_{c} without shortening the prediction horizon. This reduction of the number of parameters (metering rates) that needs to be optimized reduces the computational complexity.

An MPC-based controller uses a receding horizon framework. In a receding horizon framework, the prediction and the control horizon are shifted one control step ΔT_{ctrl} forward for every control sample k. For a given control sample k, the state of the traffic model is updated using measurements and the rN_{c} metering rates are optimized for the current prediction horizon. The r metering rates corresponding to control step k are applied to the traffic system and the other metering rates are discarded. The prediction and the control horizon are shifted one control step ΔT_{ctrl} forward and the whole process starts all over again. The diagram in Figure 4.8 summarizes the MPC approach to ramp metering. The frequent updates of the traffic state in the receding horizon framework introduce a form of feedback which makes the controller more robust and less sensitive to measurement errors, external disturbances and changes in the environment.

4.4.2 Objective function

The value of the MPC objective function is a measure of the performance of the traffic situation on the motorway. The cost associated with a traffic situation can be composed of several components such as, e.g., economical, social, environmental, ... terms that can be weighted according to global or local traffic policies. The choice of the objective function allows for a customization of the controller to a desired policy. By merely changing the objective function (or altering its parameters and weights), the policy implemented by the controller can be altered. The fact that it does not impose any constraints on the choice of the objective function is an important feature of the MPC framework.

An objective function that is often used in the literature [30, 56, 57, 81] is the total time spent (TTS) by all vehicles in the motorway network during a certain period (during the prediction horizon). In the simulation examples in Chapters 5 and 6, we choose the TTS as the objective function since it has been proven in the literature [114] that the minimization of the TTS results in the maximization of the network output. In order to smooth the control signal, a penalty on variations of the control signal is added to the objective function. In the receding horizon framework this leads to the following expression for the

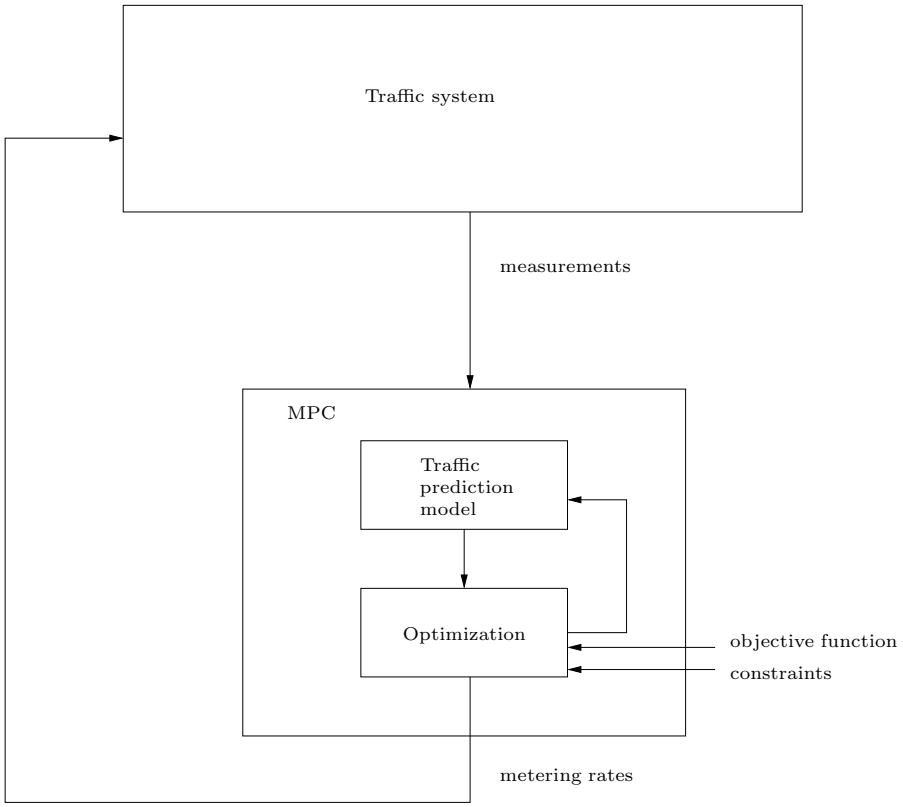


Figure 4.8: Diagram of an MPC-based ramp metering controller. The measurements of the current state of the traffic system are provided to the MPC-based controller. The optimization routine searches the metering rates that minimize the objective function over the prediction horizon taking the constraints and the prediction model into account. This process requires several simulations of the MPC traffic prediction model over the prediction horizon. The metering rate corresponding to the first control step of the prediction horizon is provided to the traffic system. The other metering rates are discarded, the prediction and the control horizon are shifted one control step forward and the whole process starts all over again to determine the metering rate that will be applied during the next control step.

MPC objective function at control step k_0:

$$\begin{aligned} J(k_0) = & \, \mathrm{TTS}_{\mathrm{motorway}}(k_0, k_0 + N_\mathrm{p} - 1) \\ & + \alpha_{\mathrm{queue}} \mathrm{TTS}_{\mathrm{queues}}(k_0, k_0 + N_\mathrm{p} - 1) \\ & + \alpha_{\mathrm{var}} \sum_{k=k_0}^{k_0+N_\mathrm{p}-1} \big(r(k) - r(k-1)\big)^2 \ , \end{aligned} \tag{4.3}$$

where $\mathrm{TTS}_{\mathrm{motorway}}(k_0, k_0 + N_\mathrm{p} - 1)$ is the TTS on the motorway during the prediction horizon that starts at controller step k and $\mathrm{TTS}_{\mathrm{queues}}(k_0, k_0+N_\mathrm{p}-1)$ is the TTS for the vehicles in the metered on-ramp queues during the prediction horizon that starts at controller step k. The TTS for all the vehicles in the studied area during the prediction horizon starting at control step k_0 and with length N_p, consists of the total time spent by all the vehicles in the motorway segments ($\mathrm{TTS}_{\mathrm{motorway}}(k_0, k_0+N_\mathrm{p}-1)$) plus the total time spent by the vehicles in the queues at the on-ramps ($\mathrm{TTS}_{\mathrm{queues}}(k_0, k_0 + N_\mathrm{p} - 1)$). The parameter α_{queue} allows putting more or less emphasis on the time spent by the vehicles in the queues. So the parameter α_{queue} corresponds to the fairness of the controller towards the motorway traffic on the one hand, and the vehicles waiting at the on-ramp on the other hand. The value of the parameter α_{var} determines the relative importance of the smoothing term that suppresses variations of the control signal. The value for α_{var} is best selected such that the variation term is about 2 to 3 orders of magnitude smaller than the other terms in the objective function. This results in a smoothing effect on the control signal without biasing the objective function too much [94]. We investigate the effect of the parameter α_{var} in Chapter 5, where we implement an MPC-based controller for the E17 case study.

4.4.3 Computing the control signals

When computing the metering rates that minimize the cost function, a trade-off between the time spent on the motorway and the time spent in the queues at the on-ramp is made. The metering rates are by definition confined to the interval $[0, 1]$. In practice, the metering rates are often constrained to the interval $[r_{\min}, 1]$ in order to limit the waiting time for the vehicles at the on-ramp. The straightforward optimization does not impose a maximal queue length on the queue at the on-ramp. Since the available capacity to store vehicles at the on-ramp is limited and since spilling back of the on-ramp queue into the underlying urban traffic network needs to be avoided, a constraint on the queue length is added to the optimization process. Besides the constraints on the queue lengths, the evolution of the traffic state over the prediction horizon is also constrained by the traffic prediction model. The resulting optimization problem to be solved is nonlinear, non-convex and has constraints. This problem can be solved using a sequential quadratic programming method [18,105,115]. This way, the MPC controller is able to take the constraint on the queue length into

account and compute the metering rates accordingly. As we have seen in Section 4.3, the ALINEA-based controller was not able to take queue constraints into account and the controller is overridden whenever the threshold length of the queue is exceeded. This will be illustrated in the case study presented in Chapter 5.

The values of N_{p} and N_{c} should be selected carefully in order to make a trade-off between the computational complexity and the accuracy of the controller. In a receding horizon framework, controller performance can be tuned using the prediction horizon and the control horizon. The larger the prediction horizon N_{p}, the further the controller looks ahead. This allows the controller to foresee certain events e.g., the onset of congestion, a queue growing too large, ... but it also increases the computational complexity. The increased computational complexity results from the MPC traffic prediction model that has to be ran over a larger prediction horizon. Since we want to implement ramp metering in a real-time framework, N_{p} is bounded from above by the time available to do the computations. As a rule of thumb, the length of the prediction horizon should be chosen in the order of the typical travel time through the network in order to obtain good results [57]. For the length of the control horizon N_{c} a similar trade-off needs to be made. The number of parameters to be optimized is proportional to the control horizon (see Section 4.4.1). Since the computational complexity of the optimization problem increases strongly with the number of parameters to be optimized, N_{c} is bounded from above by the available time to do the computations as well. In Chapter 5, we determine the prediction and the control horizon for the E17 case study.

Remark 4.3 Note that the MPC technique can be used to dynamically control all traffic control measures presented in Section 2.3, e.g., in [22] an MPC-based controller is developed for the control of speed limits. However, some of them such as e.g., peak traffic lanes, lane closures, and tidal flows lead to integer optimization problems, which are computationally very intensive. □

4.5 Conclusions

Several traffic control measures are available to reduce congestion and to improve mobility. Ramp metering, one of the traffic responsive measures presented in Chapter 2, was studied in more detail given its relevance and capabilities. Ramp metering is implementable in the short term and several successful implementations have been reported in the literature. The metering rate, which denotes how many vehicles are allowed to enter the motorway, is converted by a local controller to red and green phases of a traffic signal. The service rate of the on-ramp resulting from the metering rate depends on the state of the system (queue length and traffic density on the motorway) and on the traffic

demand.

Two techniques were presented to compute the metering rates for the ramp metering set-up: ALINEA and MPC-based control.

The ALINEA-based ramp metering controller was presented to act as a point of reference when assessing the performance of the MPC-based controller. An ALINEA-based ramp metering controller is an integrating feedback controller steering the traffic density on the motorway towards a preset density. The ALINEA-based ramp metering control is widely used in ramp metering set-ups around the world.

The main contribution to the state of the art of this chapter is the application of MPC, a well-known controller design method in the process industry, to ramp metering control. Note however that MPC is not restricted to ramp metering alone and that it can also be applied to the other traffic control measures presented in Chapter 2. In an MPC framework, a model is used to predict the future traffic states over a prediction horizon N_{p}. The choice of the traffic model is free and is not imposed by the MPC framework but rather by practical considerations such as e.g., the required level of detail, the computational complexity, ... An objective or cost function is defined to assign a cost to the computed traffic state. The choice of this objective function is free as well and is not imposed by the MPC framework. The choice of the objective function determines the traffic policy that the MPC-based controller implements. Using an optimization routine, the metering rates that minimize the objective function are determined. In this way, the MPC-based ramp metering controller steers the traffic state on the motorway to an optimal state defined by the objective function. By using constrained optimization, the MPC-based controller is able to take hard constraints into account. The optimization of the objective function makes the MPC-based controller very flexible but it also results in a higher computational complexity of the MPC-based ramp metering controller compared to ALINEA. In the MPC receding horizon framework, only the first sample of the metering rates is implemented and the prediction horizon is shifted one control sample forward after the state of the model is updated using measurements. In this way, the receding horizon approach introduces a form of feedback, which makes the controller more robust to measurement errors, external disturbances, and changes in the environment (e.g., the weather, road conditions, ...).

Chapter 5

Case study

In this chapter we discuss some simulation results of ramp metering in a real-life situation in Belgium. The performance of MPC-based ramp metering control is compared with the performance of ALINEA-based ramp metering and with the performance in the no-control case.

5.1 Case study set-up

In this chapter, we consider a part of the E17 motorway between the cities of Ghent and Antwerp in Belgium as a case study. This motorway is one of the main arteries leading to the Antwerp seaport, which makes it important from an economical point of view. The high flows of both cars and lorries on the E17 motorway during rush hours lead to structural congestion. As a case study, we focus on a stretch of 8 km of the E17 motorway in the driving direction of Antwerp where recurrent or structural congestion occurs during the morning rush hour. Besides the occurrence of structural congestion, the availability of traffic data was an important factor in the choice of this motorway stretch as the case study in this chapter. An overview of the case study area is presented in Figure 5.1.

As we can see in the more detailed schematic representation of the case study in Figure 5.2, the studied motorway stretch contains four off-ramps and five on-ramps. As shown in Figure 5.2, the on- and off-ramps are numbered as O1 - O5 and D1 - D4 respectively. The first on- and off-ramp complex a driver comes across in the case study area when traveling on the E17 motorway consists of on-ramp number O1 and off-ramp number D1, which lead to a parking lot along the motorway. The second complex (on-ramp O2 and off-ramp D2) is

located near the village Kruibeke while the third complex (on-ramp O3 and off-ramp D3) is located near the villages Burcht and Zwijndrecht. The fourth complex is located on the left-hand side of the motorway and consists of the fourth off-ramp D4 and the fourth on-ramp O4. The fourth complex is used by traffic originating from and destined to Antwerp seaport and Zelzate. The fifth complex consists only of an on-ramp 05 with traffic from Antwerp Linkeroever. Indeed, as we can observe in Figure 5.1, there is no route in the case study that leads traffic traveling on the E17 motorway in the direction of Antwerp to Antwerp Linkeroever. At the end of the case study area, the traffic enters the Kennedy tunnel underneath the river Scheldt and travels to the ring road around Antwerp, which falls outside our study area.

Special features of the motorway stretch in this case study are the presence of the Kennedy tunnel underneath the river Scheldt, and two on-ramps on the left-hand side of the motorway (O4 and 05). The last kilometer of the motorway in Figure 5.2 is a tunnel underneath the river Scheldt. Traffic drives on the right-hand side in Belgium. Hence, on-ramps are generally located on the right-hand side of the motorway and on-ramp traffic merges into the slow lane. In the study case presented here, the last two on-ramps are located on the left-hand side of the motorway, which is rather exceptional. Both the on-ramps on the left-hand side of the motorway and the tunnel influence the traffic behavior on the motorway. This behavior needs to be captured by the model parameters, which are identified based on the real-life measurement data as we discussed in Section 3.4.

5.2 METANET traffic flow model implementation

In order to be able to simulate the behavior of the E17 motorway stretch in the case study, the METANET traffic flow model presented in Chapter 3 is implemented for the study case. Since we are dealing with a motorway corridor and we do not take the urban traffic network into account in our case study, there are no alternative routes between the origins and the destinations in the case study network (see Figure 5.2). Since there are no alternative routes, we can use the non-destination oriented METANET traffic flow model for the simulation experiments for the case study. We chose the METANET model to model the case study since it provides a good trade-off between the computational complexity and the accuracy of the model. The E17 motorway stretch from Figure 5.1 is discretized as shown in Figure 5.2. The links are defined corresponding to motorway stretches without major changes in the motorway properties. Where major changes in the motorway properties occur, or at the on- or off-ramps, a node connecting the links is defined. Every link is subdivided in segments of 500 m length. We have chosen the segment length based on the configuration

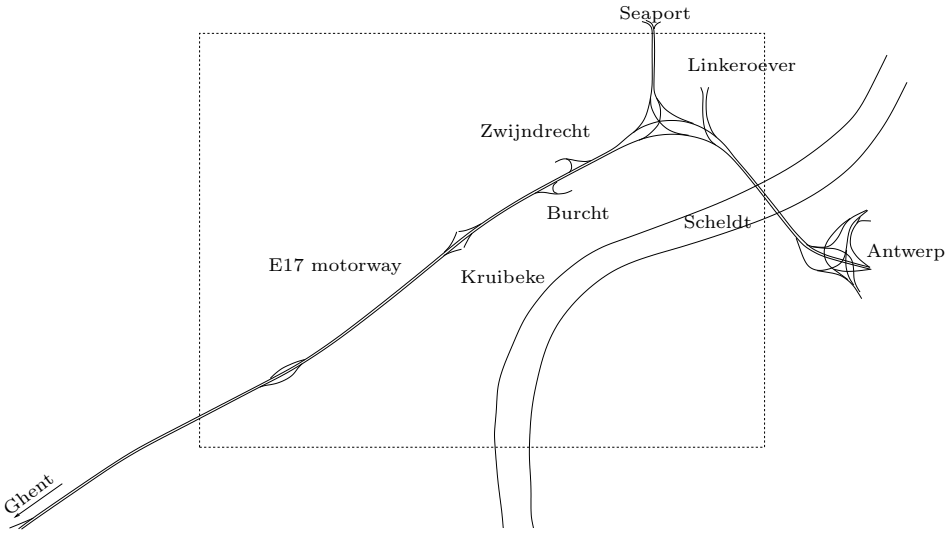


Figure 5.1: Layout of the set-up of our case study: an 8 km stretch (enclosed in the dashed box) of the E17 highway between Ghent and Antwerp. We only consider traffic driving in the direction of Antwerp.

of the case study in Figure 5.2. If we choose the segments larger, the model becomes less accurate, and if we choose the segments smaller, the computational complexity increases. The choice of a segment length of 500 m is a trade-off that is also presented in the literature [32]. Given the fact that the speed limit in Belgium is 120 km/h, we assume that the free flow speed will be smaller than 120 km/h. We take the necessary condition (3.1) for numerical stability of the METANET model into account and choose the discretization in time, or the simulation step ΔT_{sim}, equal to 10 s.

The METANET model equations of a segment in a link are given by (3.2), (3.3), (3.4) and (3.7) as presented in Chapter 3. The number of lanes n_m in a link m can be observed in Figure 5.2 and the segment length $l_{m,i}$ is 500 m. The other six link parameters $v_{\text{free},m}$, $\rho_{\text{crit},m}$, τ_m, ν_m, κ_m and a_m need to be fitted using traffic data as described in Section 3.4. The case study consists of 10 links (See Figure 5.2), which results in 60 model parameters to be determined for the links. Since the focus of this chapter is on comparing the performance of different ramp metering control strategies and not on traffic model identification, we do not go in detail about the identification process. The least significant parameters of the model for the E17 case study (i.e., all parameters except for $v_{\text{free},m}$ and $\rho_{\text{crit},m}$) are chosen from literature [81]. The parameters $v_{\text{free},m}$ and $\rho_{\text{crit},m}$ are fitted as described in Section 3.4 using traffic measurements of several days. The parameters $v_{\text{free},m}$ and $\rho_{\text{crit},m}$ for the E17 case study were found to be equal to 105 km/h and 34 veh/km/lane respectively. The quality of

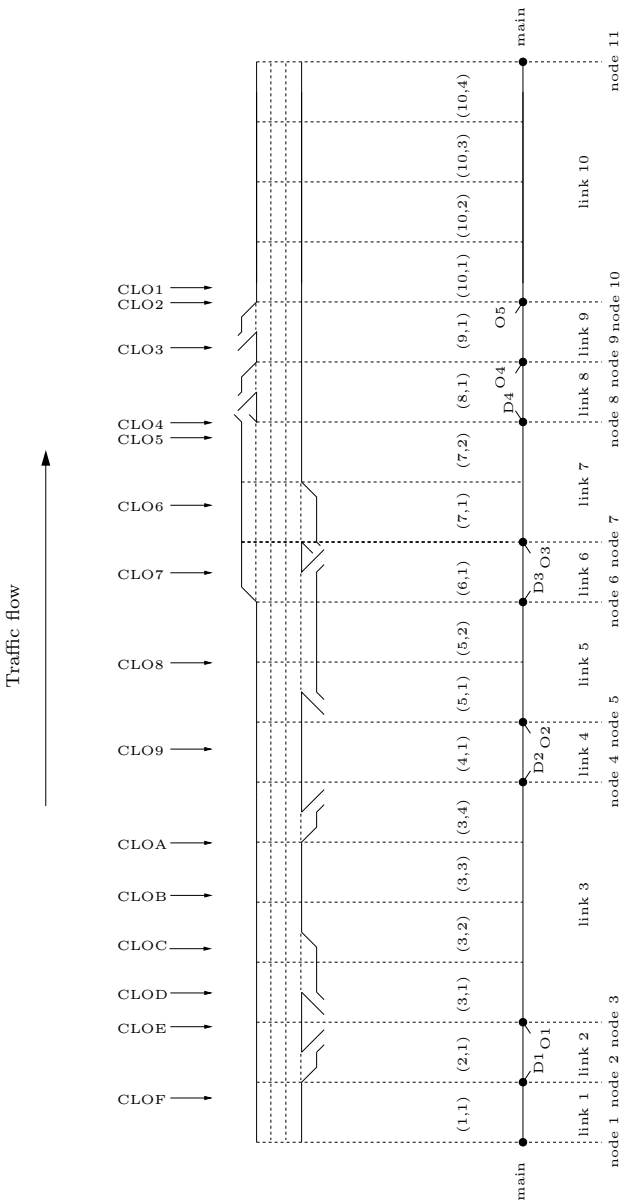


Figure 5.2: Schematic representation of the segmentation of the E17 case study motorway. The labels CLO1 to CLOF indicate the locations of the traffic measurement cameras. The labels D1 to D4 represent the off-ramp destinations and the labels O1 to O5 represent the on-ramp origins.

the obtained parameters is verified by the comparison of simulation data with real measurement data as will be presented in Section 5.3.

The on-ramps are modeled as a queue with a service rate that is dependent on the traffic density on the motorway as described in (3.8) and (3.9). The on-ramp capacity $Q_{\text{cap},o}$ of on-ramp o is the maximal number of vehicles that can enter the motorway through the on-ramp. As an upper limit for the one-lane on-ramp traffic flow on on-ramps 1, 2 and 3, we choose 2000 veh/h, which is the typical capacity of a single motorway lane. The service rate of the on-ramp queues is also determined by the properties and the state of the motorway segment the on-ramp feeds into. In order to compute the service rate according to (3.9), the jam density ρ_{jam,m_o} of the links fed by the on-ramps, needs to be determined. This results in five more parameters to be identified. We choose the jam density from the literature and equal to 180 veh/km/lane [81]. As we can observe in Figure 5.2, the fourth and the fifth on-ramps consist of two lanes. By consequence we choose the capacities for the fourth and the fifth on-ramp equal to 4000 veh/h. In the simulation experiment that will be explained in detail in Section 5.4, we simulate ramp metering on the fourth and/or on the fifth on-ramp and hence the service rate for the on-ramp queues is given by (3.10). As in (3.9) for the other on-ramps, the jam density needs to be defined for the links fed by the on-ramps. We choose the jam density also equal to 180 veh/km/lane. As described in Section 3.2, a merging term (3.5) needs to be added to (3.3) in the first segment of each link fed by the on-ramp in order to compensate for the speed drop in the segment caused by the merging vehicles. The term in (3.5) contains an extra parameter δ_{m_o}, which is chosen equal to 0.0122 [81]. For every one of the five on-ramps, the METANET model needs three extra parameters ($Q_{\text{cap},o}$, ρ_{jam,m_o}, and δ_{m_o}), which results in 15 additional parameters in the METANET flow model for the case study.

As can be seen in Figure 5.2, the number of lanes drops from 4 to 3 near the fourth off-ramp. This causes weaving effects in segment 2 of link 7, which are accounted for by (3.6). One extra tuning parameter ϕ_m for the weaving effects is added to the METANET case study model. The value of the parameter ϕ_m is chosen equal to 0.0122.

For every off-ramp, a node n with a dummy link $m_{\text{leav},n}$ leaving the motorway is defined as presented in Figure 5.3. The link $m_{\text{main},n}$ corresponds to the motorway link leaving node n. The traffic flow leaving the motorway through an off-ramp is modeled as a flow on the corresponding dummy link $m_{\text{leav},n}$. The flow on this dummy link is determined by the turning rate $\beta_n^{m_{\text{leav},n}}(l)$ and can be computed using (3.12). The turning rate $\beta_n^{m_{\text{main},n}}(l)$ defines the flow that leaves node n through the motorway. From (3.12) it follows that the following relation holds in an off-ramp node n:

$$\beta_n^{m_{\text{leav},n}}(l) + \beta_n^{m_{\text{main},n}}(l) = 1.$$

In Figure 5.2 we see that we need to define leaving rates corresponding to off-

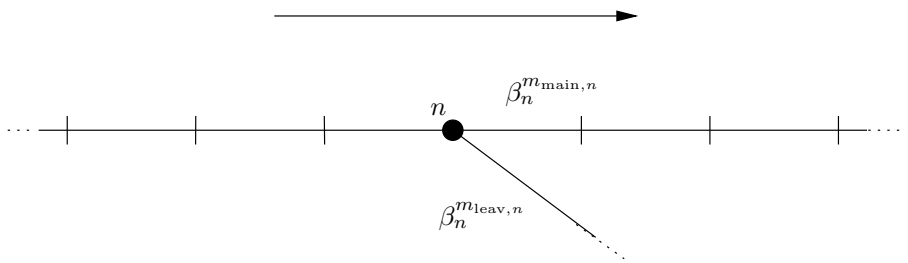


Figure 5.3: Schematic representation of a node n with a link $m_{\text{leav},n}$ corresponding to the off-ramp and with a motorway link $m_{\text{main},n}$. The leaving rate $\beta_n^{m_{\text{leav},n}}$ is the fraction of the total traffic flow entering node n that leaves the motorway through the off-ramp in node n. The turning rate $\beta_n^{m_{\text{main},n}}$ is the fraction of the traffic flow entering node n that continues its journey on the motorway. The sum of both fractions $\beta_n^{m_{\text{main},n}}$ and $\beta_n^{m_{\text{leav},n}}$ always equals 1. The arrow denotes the direction of the traffic flow.

ramps at nodes 2, 4, 6 and 8. The turning rates are provided to the simulation model as inputs. We discuss the estimation of the turning rates based on traffic measurements in Section 5.3.

Finally, we need to take the initial and the boundary conditions into account. The initial conditions are determined by simulating the traffic model with a constant demand until a steady state is reached. These steady states are used as the initial conditions. At the mainstream origin, a virtual upstream segment is considered in order to account for the downstream propagation of the speed as was discussed in Chapter 3. The average speed in the virtual segment is chosen equal to the average speed in the first segment of the case study which results in a convection term equal to zero for the first segment of the case study. Since the traffic demands at the mainstream origin can be larger than the number of vehicles that can enter the first motorway segment under congested traffic operation, a queue is defined at the entrance of the motorway. The service rate of this queue is determined by the capacity of the motorway and by the traffic state in the first segment of the motorway. We choose a capacity of 6000 veh/h at the entrance of the three-lane motorway. During high density operation of the motorway ($\rho_{1,1} > \rho_{\text{crit},1,1}$), the density-dependent service rate of the entrance queue is chosen equal to the flow in the first segment. The difference in the computation of the service rate of the entrance queue and the computation of the service rate at the on-ramps results from the different phenomena that occur at these traffic origins. At the on-ramp origins merging occurs, which results in a density dependent service rate. At the entrance of the modeled motorway, we assume that there are motorway segments upstream of the first segment of the motorway model, which fall outside the case study model. Hence, the entering vehicles already travel at a high speed and no merging nor

convection effects occur. During all simulations, we carefully observed that the queue at the motorway entrance remained empty since a non-empty queue at the motorway entrance is an indicator for congestion that spills back outside of the modeled network. If this happens, the case study network needs to be enlarged to capture all effects of congestion. In our simulation experiments, the queue length at the motorway entrance remained zero at all times.

Another boundary condition occurs at the last segment of the simulated motorway. In order to take the upstream influence of the traffic density through anticipation into account in the last segment, a traffic density needs to be defined downstream of the last segment. We chose the downstream traffic density equal to the critical density of the last motorway link. In this way, it is prevented that the vehicles are 'sucked' out of the motorway due to anticipation effects during congested traffic operation. However, during free flow traffic operation, this boundary condition results in a decrease of the average speed due to the anticipation term. We observed that this effect was quickly dampened in the upstream segments. In Figure 5.2, the two last segments were added to dampen this effect without distorting the simulation results. Note that these two extra segments are only relevant to obtain a good boundary condition, they are not taken into account in the objective function.

Given the traffic demands and the turning fractions, the parameters of the traffic flow model can be fitted using identification techniques [89]. The parameters are determined such that a least-squares distance criterion between the measured and the simulated traffic states on the motorway is minimal as described in Section 3.4. The total number of parameters that has to be determined for the case study model is 79. Since the main goal of this chapter is to compare the performance of MPC-based and ALINEA-based ramp metering control, we chose most of the parameters from the literature. Only the most sensitive parameters were fitted to traffic measurement data as described above. The validity of this choice is investigated in the next section using simulation and traffic measurements.

5.3 Traffic demands and measurements

As most motorways in Belgium, the E17 motorway in the case study is equipped with inductive loop detectors before each off-ramp and after each on-ramp. The traffic detectors were originally installed in the seventies for monitoring the traffic conditions on the motorways. The combination of the measurements before and after the complex provides an estimate of the net number of vehicles leaving (or entering) the motorway at the complex. The traffic detectors only measure the traffic parameters for the traffic flows on the motorway and not for the traffic flows on the off-ramps or on the on-ramps. If we want to

simulate the METANET model for the study case as presented in the previous section, we need to know the number of vehicles leaving the motorway through the off-ramp and the number of vehicles entering the motorway using the on-ramp. In order to estimate these on-ramp and off-ramp flows, at least one additional traffic flow measurement per on-ramp/off-ramp complex is required. These additional measurements are in general not implemented on the Belgian motorways out of cost considerations. However, for the E17 motorway stretch presented in our case study, there were additional cameras placed in the scope of a European project. The distance between the cameras is approximately 500 m and the exact camera locations are presented in Figure 5.2. The available traffic measurements for the E17 motorway stretch consist of the traffic flow, the average speed, and the occupancy for every of the three lanes. The camera images are fed to an image processing algorithm, which provides new values of the traffic flow, the average speed and the occupancy for every lane every minute. The availability of these extra measurements is, besides the economical relevance of the E17 motorway and the occurrence of recurrent congestion, an additional argument to choose the E17 motorway as our case study in this thesis.

Advanced techniques exist to estimate the traffic flows between the origins and the destinations of a traffic network based on traffic measurements in the network. However, due to the corridor layout of the traffic system that we consider in the case study (see Figure 5.2) and due to the presence of many traffic cameras, we can estimate the on-ramp demands and the splitting rates for the case study in a simplified manner as we will discuss below. We refer the reader interested in origin-destination estimation to the literature [4,84,85, 138, 143].

When we observe the locations of the cameras in the study case as shown in Figure 5.2, we see that at every complex there are cameras available before the off-ramp, between the off-ramp and the on-ramp, and after the on-ramp. This situation is schematically represented in Figure 5.4. The traffic flow leaving the motorway through the off-ramp in Figure 5.4 is given by the difference between the flow $q_A(p)$ measured by sensor A and the flow $q_B(p)$ measured by sensor B where p is the measurement step counter for the sensors. In the case study, the measurement step ΔT_{meas} is equal to 1 min. Since the sensors are only 500 m apart, we neglect the travel time between sensor A and sensor B and compute the flow on the off-ramp as $q_A(p) - q_B(p)$. Using (3.12), the turning rate for the off-ramp can be computed as $\dfrac{q_A(p) - q_B(p)}{q_A(p)}$. In a similar way, the on-ramp traffic demands can be computed based on the measurements of the traffic sensors B and C (see Figure 5.4). The on-ramp traffic flow is equal to the difference between the traffic flow $q_C(p)$ and the traffic flow $q_B(p)$. Once again, we neglect the travel time between the sensors B and C.

Based on the available traffic flow measurements and based on the reasoning

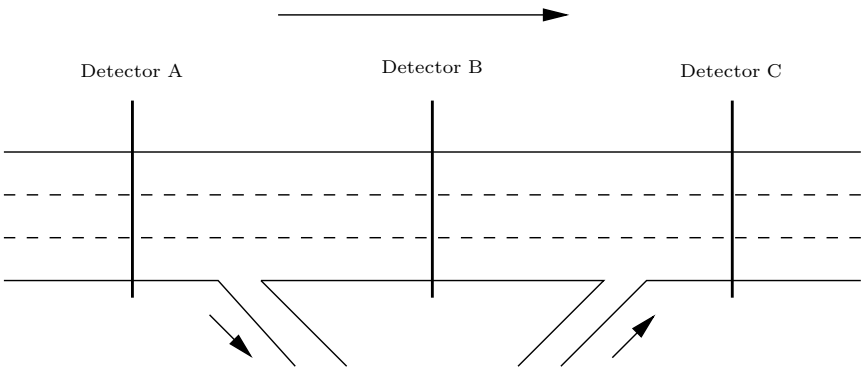


Figure 5.4: Schematic representation of a traffic complex consisting of an off-ramp followed by an on-ramp. Three detectors are present: detector A, detector B, and detector C. Using the layout information of the motorway and the location of the sensors, the off-ramp flow (and the associated leaving rate) and the on-ramp flow can be computed. The arrow denotes the direction of the traffic flow.

presented above, an estimation of the traffic flows on the on- and the off-ramps can be made.

The evolution of the traffic demands and the turning fractions used in the METANET-based simulations in this chapter are based on traffic measurements during a test period ranging from Tuesday, February 22, 2000 up to and including Friday, February 25, 2000. The measurements considered in this study originate from four ordinary working days. There is a large correspondence between the measured evolutions of the traffic flows over time for the four days considered in the test period. Hence, for every camera location, the evolution of the expected traffic flow measurements can be calculated by averaging the measurements of the four days in the test period. In the upper two plots of Figure 5.5 the average evolution over time of the traffic flow measured by the traffic detectors CLO2 and CLO3 is presented (see Figure 5.2 for the exact locations of the cameras). The plotted average is an average evolution of the four working days in our test period. By combining the knowledge about the locations of the cameras and the motorway topology presented in Figure 5.2 with the average evolutions of the traffic flows on the motorway during a typical working day, an estimation of the average evolution of the traffic flows and the turning fractions over time can be obtained. In the lower plot of Figure 5.5 the traffic flow on the fifth on-ramp is plotted in diamonds. Since the traffic flow measurements of the detectors CLO2 and CLO3 fluctuate from minute to minute, as can be seen in the upper two plots of Figure 5.5, the computed flow for the fifth on-ramp is averaged over 15 minutes. In this way, the impact of the neglected travel time from detector CLO3 to detector CLO2, which is of the

order of magnitude of 1 minute under congested traffic operation, is reduced even further.

There is a fundamental difference between the on-ramp flows that we computed as presented above and the traffic demands at the different on-ramps. Indeed, the traffic demands cannot be measured directly. The traffic flows that are measured, or computed, are the result of a traffic demand applied to the traffic system. This can be clarified by an example. Consider the traffic flow and the average speed on the E17 motorway near camera CLOF (i.e., the traffic demand at the main motorway entrance of the study area) during the morning rush hour on Thursday, February 24, 2000 (see Figure 3.8). During congested traffic operation, a decrease in the traffic flow on the motorway can be observed. This decrease in the traffic flow is not caused by a decreased traffic demand but it is induced by the properties of the motorway. The observed congestion and the corresponding reduced traffic flow in Figure 3.8 occur due to a traffic demand that is higher than the traffic capacity of the motorway. The traffic demand that will be used as the input for the simulation model is the number of vehicles that *want to travel* on the motorway over a period of time while the measured flow is the number of vehicles that *travel* on the motorway over a period of time. For simplicity, we model the evolution of the traffic demands over time as a piecewise affine function. In Figure 5.5 the solid line in the lower plot represents a piecewise affine approximation of the evolution of the traffic demand at the fifth on-ramp over time while the diamonds represent the on-ramp flows that were computed based on measurements. For every on-ramp we model the demands as a piecewise affine function that is determined based on the measurements. The estimated evolution of the traffic demands at the on-ramps is plotted in Figure 5.6. For simplicity, we assume that the turning fractions for the case study are invariant in time. The turning fractions are computed based on the measurements as presented above.

In order to verify the accuracy of the METANET model with the parameters chosen as presented in the previous section and with the traffic demands and the splitting rates computed as presented above, we simulated the traffic model and compared the evolution of the measured traffic flows and the simulated traffic flows. The result for the camera CLO2 is shown in Figure 5.7. After a comparison of the simulated traffic flows and average speeds with the measured traffic flows and average speeds, we concluded that the estimations of the parameters, of the turning fractions and of the demand patterns are sufficiently accurate to allow for a comparison of ramp metering strategies for the case study. The description of the simulation experiment set-up to compare ramp metering controllers is presented in the next section.

Remark 5.1 It can be observed in Figure 5.7 that during the early morning hours, the measured average speed of the traffic does not correspond with the simulated average speed. The registered speed by the camera is approximately

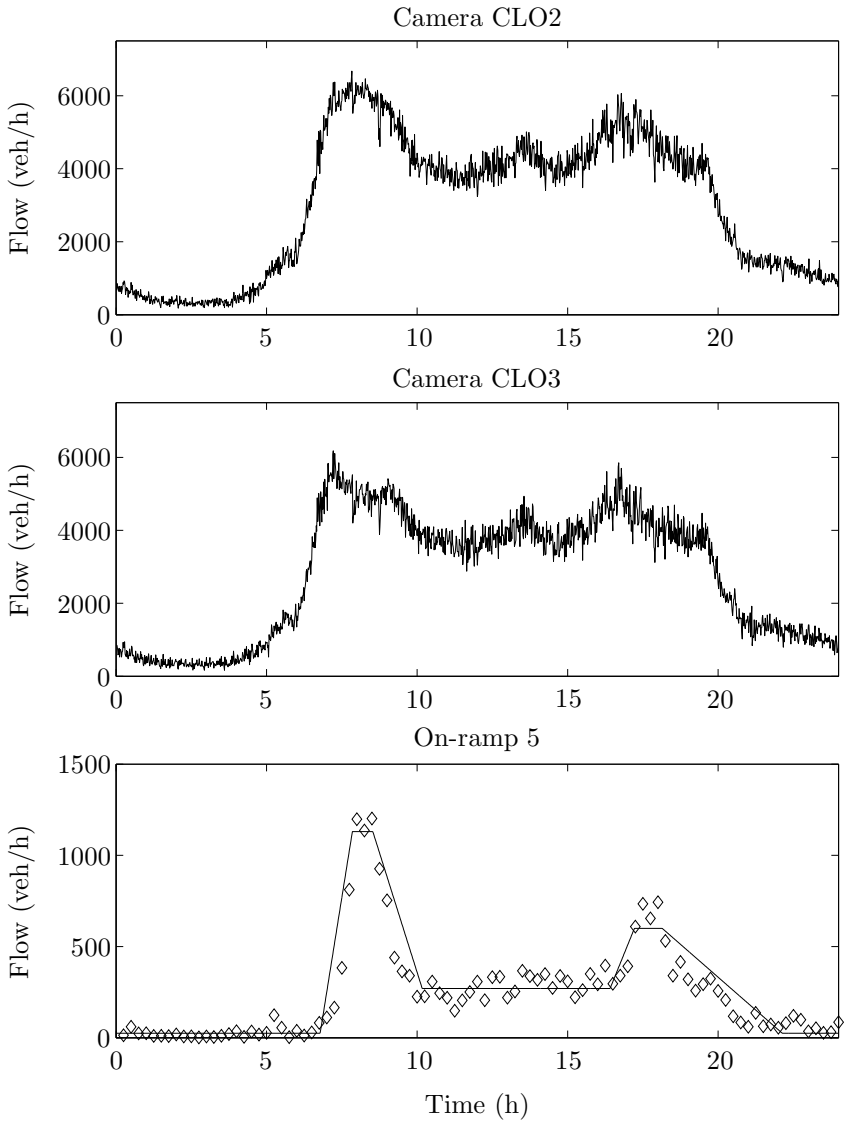


Figure 5.5: Average evolution of the traffic flow over a day as measured by the cameras CLO2 and CLO3 on the E17 Ghent–Antwerp. The values are the average of the values measured during the four day test period from Tuesday, February 22, 2000 up to and including Friday, February 25, 2000. In the lower plot, the average over 15 minutes of the difference between the flows measured by CLO2 and CLO3 is represented by the diamonds. The estimation of the evolution of the on-ramp demand at the fifth on-ramp is plotted in the lower plot in a solid line.

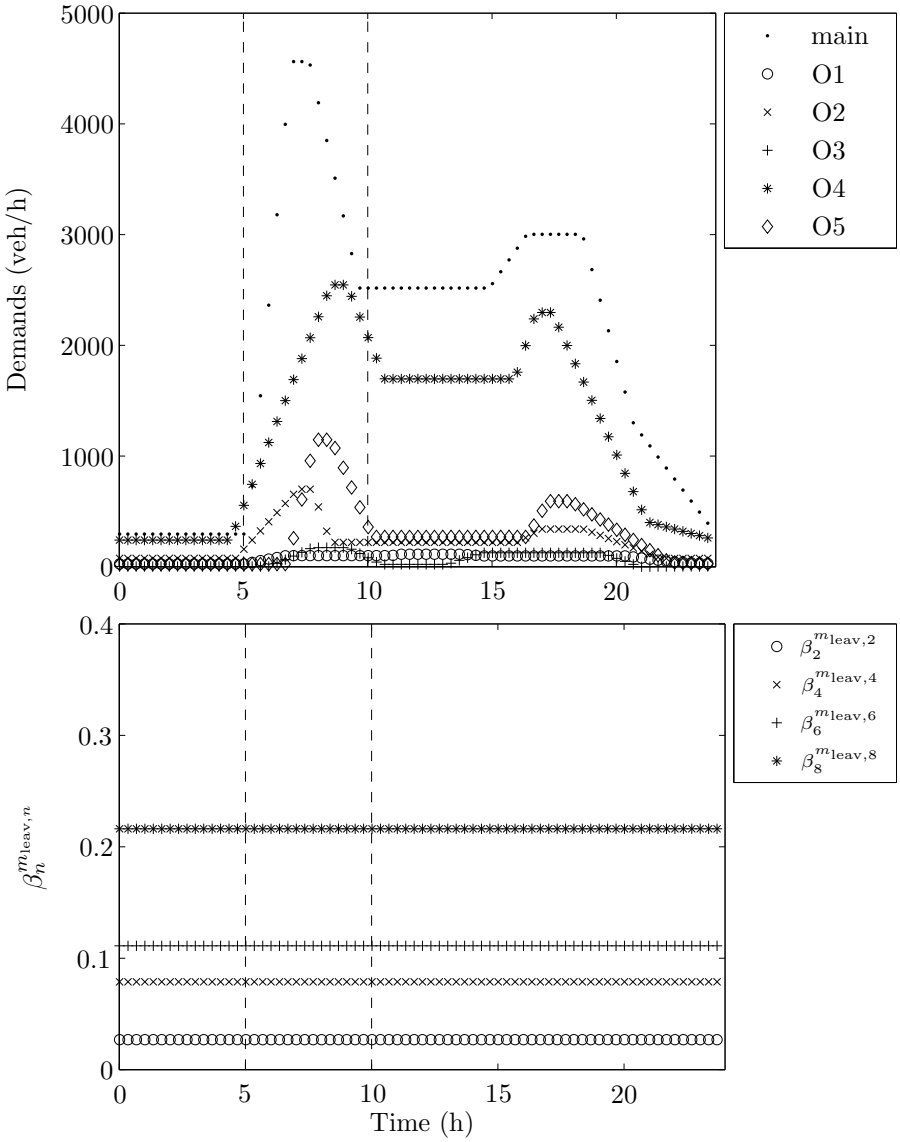


Figure 5.6: Piecewise affine approximation of the traffic demands and the leaving rates $\beta_n^{m_{\text{leav},n}}$ for the E17 motorway Ghent–Antwerp case study on a typical working day. The period ranging from 5 a.m. to 10 a.m. is the morning rush hour period and is marked by the vertical dashed lines. The traffic demands and the leaving rates during this morning rush hour period will be used in the simulation experiments below.

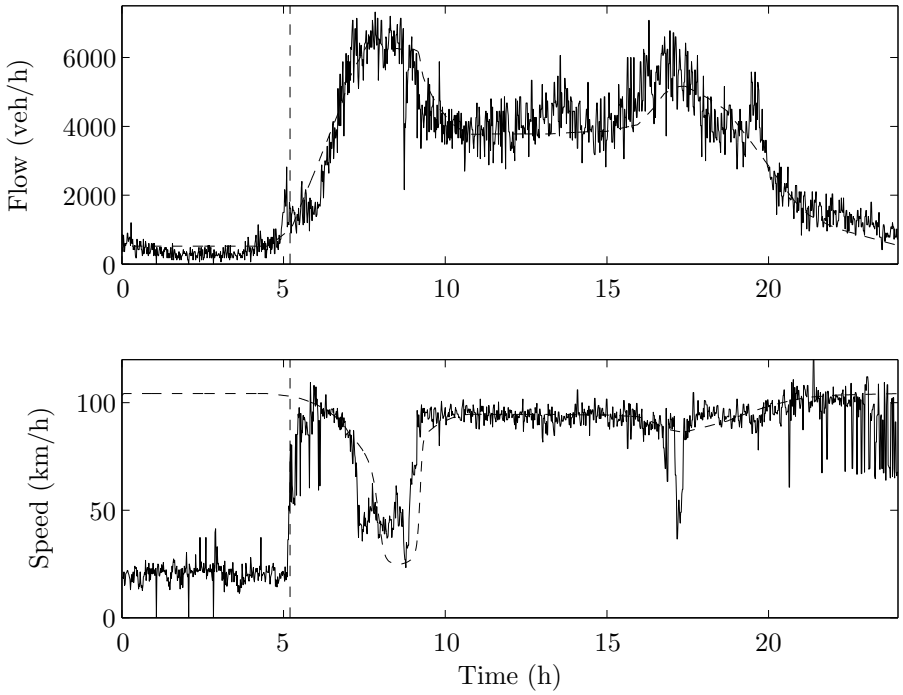


Figure 5.7: Plot of the flow (upper plot) and the average speed (lower plot) measured by the sensor CLO2 on the E17 motorway on Wednesday, February 23, 2000. The dashed line represents the simulated values using the estimated demands, the estimated leaving fractions, and the METANET model parameters. The traffic data at the left-hand side of the dashed vertical lines have been discarded since they probably correspond to sensor failures as can be seen from the very low speed measurements (see Remark 5.1).

25 km/h, which is unrealistically low, especially since no incidents were reported during this period. These data probably correspond to measurement errors or sensor failures and they were discarded from the identification data set. □

5.4 Experiment description

Recurrent congestion originates on the studied motorway near the entrance of the tunnel underneath the river Scheldt, i.e., near the fourth and the fifth on-ramp in Figure 5.2. Therefore, ramp metering will be investigated at these two large volume (two lane) on-ramps. Two ramp metering control techniques will be implemented and their performance will be assessed and compared. The METANET model presented before is used to simulate the case study motorway. This simulation is performed to investigate two ramp metering control techniques: ALINEA and MPC-based ramp metering. In order to be able to assess the performance of the traffic operation realized by the controllers, this section presents a performance measure for the simulated traffic situation on the motorway and on the on-ramps in the case study.

First, the model is simulated without any control in order to have a reference value of the performance measure, which can subsequently be compared to the performance values resulting from the simulations with ALINEA and MPC-based control.

The simulations in this chapter are ran over a period from 5 a.m. to 10 a.m., which corresponds to a morning rush hour. The total time spent (TTS) by all the vehicles in the network during the simulated five hour period is defined as a performance measure P_{TTS}. The TTS by the vehicles in the traffic network during the simulation period is given by:

$$P_{\mathrm{TTS}} = \sum_{l=0}^{N_{\mathrm{sim}}-1} \left[\sum_{(m,i)\in\mathcal{I}_{\mathrm{m}}} \rho_{m,i}(l) l_{m,i} n_m + \sum_{o\in\mathcal{I}_{\mathrm{o}}} w_o(l) \right] \Delta T_{\mathrm{sim}}, \tag{5.1}$$

where N_{sim} is the number of simulation steps in the simulated period, $\mathcal{I}_{\mathrm{m}}$ is the set of index pairs (m,i) for all motorway segments, and $\mathcal{I}_{\mathrm{o}}$ is the set of on-ramp indices. The simulation sample step of the model ΔT_{sim} was chosen to be 10 s. As a result there are $N_{\mathrm{sim}} = 1800$ simulation steps in the simulated morning rush hour. The lower the TTS during the simulated five hour period, the higher the performance of the motorway system. Note that there is a difference between the performance measure P_{TTS} in (5.1), which provides a performance measure for the simulated morning rush hour, and the MPC objective function in (4.3), which is used in the MPC rolling horizon framework to compute the metering rates.

In the next section, the no-control case is simulated. The simulation results are

discussed and the performance criterion P_{TTS} is computed as a reference for the simulations with ramp metering control enabled. Since congestion occurs near the fourth and the fifth on-ramps in the study case, the effect of ramp metering at these on-ramps on the motorway performance will be investigated. Ramp metering will be investigated on the fourth on-ramp, on the fifth on-ramp, and on both on-ramps simultaneously. The performance of each configuration is assessed using (5.1). Since the length of the queue resulting from ramp metering operation needs to be acceptable, we impose a hard constraint on the queue lengths at the on-ramps. The maximal allowed queue length is 100 vehicles.

In the following section, we discuss the implementation of the widespread ALINEA ramp metering algorithm for the fourth and the fifth on-ramps in the case study. The performance of the case study motorway system is assessed for ALINEA-based ramp metering on the fourth, on the fifth, and on both the fourth and the fifth on-ramps simultaneously. The ALINEA controller parameters are varied in order to investigate their impact on the system performance.

After the study of the ALINEA technique, we investigate the performance of the MPC-based approach to ramp metering that was presented in Chapter 4. MPC-based ramp metering is implemented on the fourth on-ramp, on the fifth on-ramp, and on both the fourth and the fifth on-ramp simultaneously. The experiments discussed in Section 5.7 include the tuning of the MPC controller parameters N_{p} and N_{c}. The influence of a varying weighting parameter α_{var} in the control objective function is investigated as well. Since the MPC framework allows for coordinated control of ramp metering set-ups, we investigate the difference in the performance of the system between non-coordinated MPC-based ramp metering and coordinated MPC-based ramp metering on the fourth and the fifth on-ramps. It will be shown that, in contrast to MPC, ALINEA is not able to guarantee a *hard* constraint on the queue length.

5.5 The no-control case

In order to have a point of reference, a simulation of the morning rush hour without control is performed. The METANET traffic flow model described in Section 3.2 and in Section 5.2 is simulated using the traffic demands determined in Section 5.3. The evolution of the traffic density, the average speed and the traffic flow in every segment and for every simulation step is plotted in Figure 5.8.

In the upper plot in Figure 5.8, for every of the 18 segments in the case study model (see Figure 5.2), the evolution of the traffic density during the morning rush hour is presented. The traffic density in the first segment increases as the traffic demand increases and decreases again after the peak traffic demand is

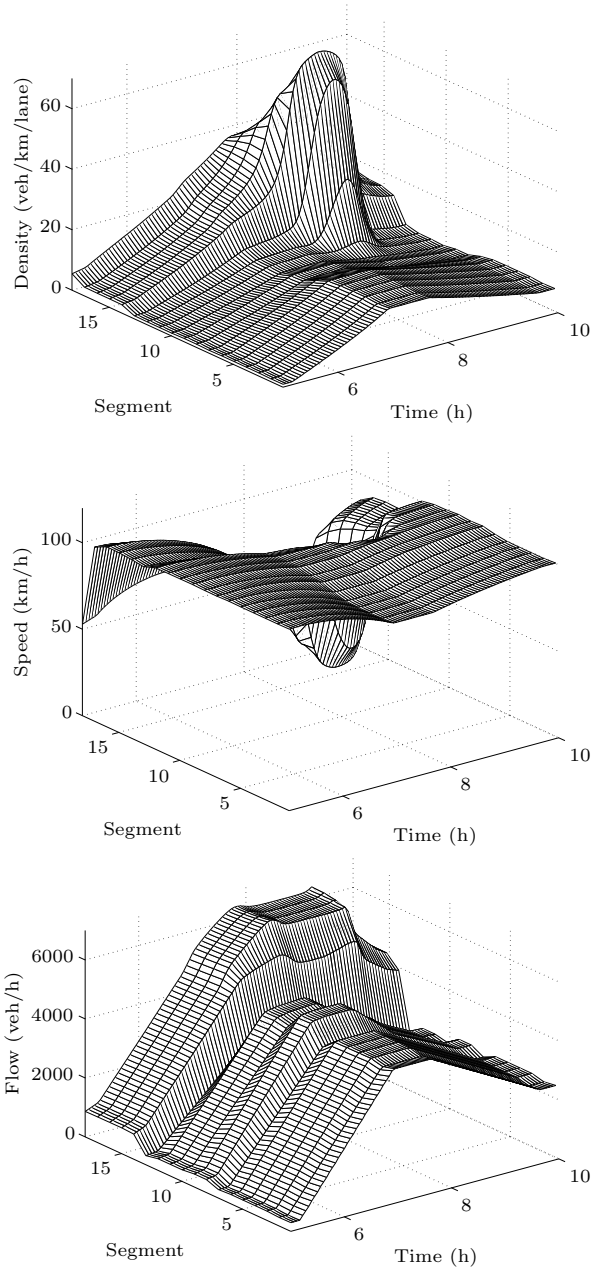


Figure 5.8: Simulation results of the no-control case. Plots of the traffic density (top), the average speed (middle) and the traffic flow (bottom) for every segment and for every simulation step of the METANET E17 case study model. The simulation period is the morning rush hour defined in Figure 5.6.

over. This plot also illustrates the occurrence of congestion in the last segments of the motorway. E.g., in segment 14 (corresponding to segment 1 from link 9 in Figure 5.2), the traffic density grows larger than the critical density ($\rho_{\text{crit}} = 34$ veh/km/lane) as the rush hour progresses. Congestion sets in. Some time after congestion sets in in segment 14, the density starts to increase in segment 13. The congestion spills back in the upstream direction. The onset of congestion can also be observed in the middle plot in Figure 5.8 where the evolution of the average speed in the segments over time is plotted. The average speed in segment 14 starts decreasing drastically when the traffic density in the segment is larger than ρ_{crit} and congestion sets in. Due to the spill-back of congestion, the speed in the upstream segments decreases as well but with a delay. In the lower plot of Figure 5.8, the evolution of the traffic flow in every segment over time is shown. The typical shape of the traffic demand patterns from Figure 5.6 can still be distinguished in the traffic flows in the segments. The large differences in the traffic flows in adjacent segments are caused by the presence of on-ramps and off-ramps, which result in a flow of vehicles entering or leaving the motorway.

The TTS in the case study traffic network without control, computed according to (5.1) is $P_{\text{TTS}} = 2038$ vehicle hours (veh h). In the next sections we will present ramp metering controllers for the large volume fourth and fifth on-ramps and compare the value of P_{TTS} for the no-control case with the values of P_{TTS} realized by the ramp metering controllers.

5.6 ALINEA

Using (4.2) and the traffic flow model of the motorway, a simulation can be performed to assess the performance of ALINEA as a traffic control measure on the E17 motorway.

5.6.1 Experiment set-up

Since during ramp metering a queue can form and since this queue cannot be allowed to grow larger than the available storage capacity of vehicles at the on-ramp, the queue length must be limited. ALINEA does not take the queue length into account in its control law. The constraint on the queue length can be imposed by overriding the metering rate provided by ALINEA once a threshold queue length is exceeded. The metering rate is in practice set equal to 1 for as long as the queue length is above the threshold [76]. Once the queue length becomes again smaller than the threshold, the metering rates provided by ALINEA are implemented again.

ALINEA's integrating control law (4.2) contains two parameters $\hat{\rho}$ and K_R.

Some simulations of the morning rush hour period were performed to determine appropriate values for these parameters. The performance of ALINEA was found not to be very sensitive to the choice of the parameter $\hat{\rho}$ for the simulated traffic demands. Since $\hat{\rho}$ is in fact a setpoint expressing the desired traffic density on the motorway, we chose $\hat{\rho} = \rho_{\text{crit}}$ since the motorway operates at full capacity at the critical density. In Remark 4.2 it was stated that the traffic flow associated with the critical density could be lower than the motorway capacity due to hysteresis effects. The effects of hysteresis on the performance of the ALINEA controller are taken into account in the simulations presented below since hysteresis can be observed in the traffic states simulated by the METANET model.

5.6.2 ALINEA-based ramp metering control simulation results

We simulated the following configurations of ALINEA-based ramp metering control on the fourth and the fifth on-ramps of the E17 case study motorway stretch:

- ALINEA-based ramp metering control on the fourth on-ramp,
- ALINEA-based ramp metering control on the fifth on-ramp, and
- ALINEA-based control on the fourth and on the fifth on-ramps.

Let us now investigate the simulation results of each of these configurations in more detail.

ALINEA-based control on the fourth on-ramp

In this section we investigate the impact of the parameter K_R on the performance P_{TTS} of the ALINEA-based ramp metering controller on the fourth on-ramp. We simulated the case study motorway with ALINEA-based ramp metering control on the fourth on-ramp for different values of K_R. We now discuss the performance P_{TTS} and the behavior of the resulting controller for the morning rush hour shown in Figure 5.6.

The parameter K_R was allowed to vary from 0.0001 km lane/veh to 1 km lane/veh. The experiments with varying value of K_R are summarized in Table 5.1. For $K_R = 0.0001$ km lane/veh, the gain of the feedback law is too small and the ramp metering set-up does not become active. The performance P_{TTS} is 2038 veh h, which is exactly the value found in the no control case. From $K_R = 0.0005$ km lane/veh and up the value of P_{TTS} decreases with increasing K_R. Figure 5.9 shows the simulation results of two ALINEA controllers.

K_R (km lane/veh)	P_{TTS} (veh h)
0.0001	2038
0.0005	2034
0.001	2029
0.005	1999
0.01	1978
0.05	1942
0.1	1942
0.5	1942
1	1943

Table 5.1: Overview of the TTS (P_{TTS}) by the vehicles on the E17 motorway and on its on-ramps during the simulated morning rush hour, realized by an ALINEA ramp metering controller on the fourth on-ramp. The ALINEA controller gain K_R was varied. We observe that the performance of the ALINEA-based controller increases (P_{TTS} decreases) with increasing values of the controller gain K_R.

The plots on the left-hand side result from an ALINEA controller with $K_R = 0.0005\,\mathrm{km\,lane/veh}$ and the plots on the right-hand side result from an ALINEA controller with $K_R = 0.5\,\mathrm{km\,lane/veh}$. We can see in Table 5.1 that the controller with the larger gain ($K_R = 0.5\,\mathrm{km\,lane/veh}$) outperforms the controller with the smaller gain ($K_R = 0.0005\,\mathrm{km\,lane/veh}$) considering the TTS during the simulated morning rush hour.

In the upper plots in Figure 5.9, we see the traffic demand at the fourth on-ramp in a solid line and the real on-ramp flow (resulting from traffic demand, metering rate and on-ramp capacity according to (3.10)) in a thick dotted line. In the next plots the evolution of the metering rate, the queue length, the traffic density on the motorway and the average speed on the motorway are presented respectively. The controller with the smaller gain (left) has a much smoother metering rate and resulting on-ramp flow than the controller with the larger gain (right). Note that according to (3.10) a metering rate lower than 1 does not necessarily mean that a queue is formed. A queue is only formed if the metering rate is small enough to reduce the service rate (3.10) below the on-ramp traffic demand. In the example of the ALINEA controller with the smaller gain in Figure 5.9, this can be observed as a delay between the reduction of the metering rate and the start of the build-up of the queue.

In practical implementations (e.g., in the set-up in Kessel-Lo shown in Figure 2.11), the ALINEA controller is overridden and the metering rate is set to 1 once the queue length exceeds the threshold length. If the ALINEA controller is overridden, the metering rate suddenly jumps towards 1 and the accompanying on-ramp flow immediately increases. This can e.g., be observed as the sudden jump of the metering rate and the on-ramp traffic for the controller

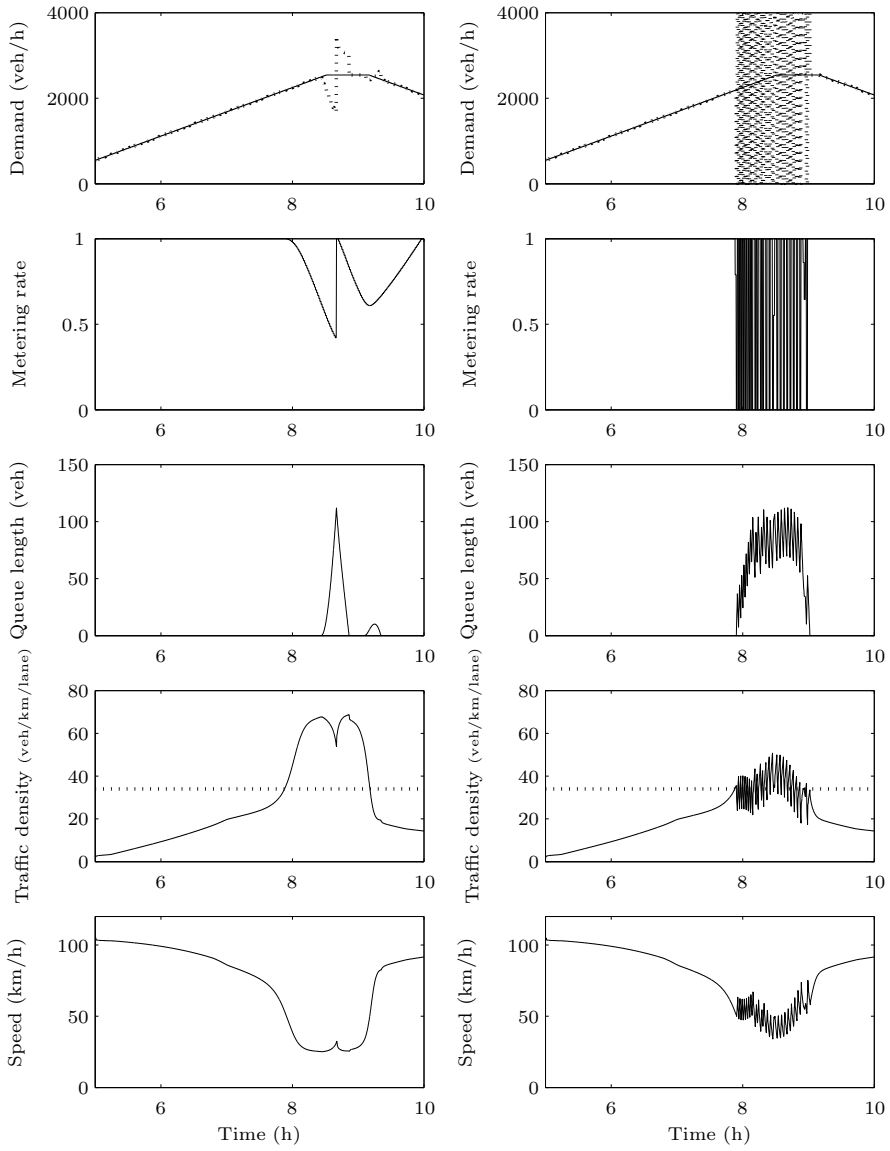


Figure 5.9: ALINEA ramp metering control for $K_R = 0.0005\,\mathrm{km\,lane/veh}$ (left) and $K_R = 0.5\,\mathrm{km\,lane/veh}$ (right). The dotted line in the upper plots represents the real on-ramp traffic flow. The dotted line in the density graphs represents the critical density ρ_{crit}. As can be seen in the plots for $K_R = 0.5$ the ALINEA controller oscillates if the controller gain is chosen too large.

with the smaller gain in Figure 5.9. If the queue length becomes smaller than the threshold length, ALINEA takes over again. Overriding the ALINEA controller generally occurs if the traffic density on the motorway is high (if the metering was active, which resulted in a queue that grew too large) and when ALINEA becomes active again this generally results in an immediate decrease of the metering rate as can be observed in the plots of the metering rate in Figure 5.9. The larger the gain K_R, the faster the decrease in metering rate after the controller has been overridden, but also the faster the switching between ALINEA control and the overriding of the control under high on-ramp demand operation.

Both ALINEA controllers in Figure 5.9 start adapting the metering rate at the same moment in time. This can be observed in the plots of the density in the segment fed by the on-ramp in Figure 5.9. The ALINEA controllers both start metering the on-ramp at the moment the traffic density on the motorway grows larger than $\hat{\rho}$ (represented by the dotted line). However, the controller with the smaller gain reacts much more slowly to the increased traffic density than the controller with the larger gain. Hence, the traffic density on the motorway is allowed to increase substantially above the critical density ρ_{crit} before the ALINEA controller with the smaller gain starts limiting the on-ramp flow by sufficiently reducing the metering rate (see Figure 5.9). The increased motorway traffic density is accompanied with a reduced average speed on the motorway as presented in the lower plot of Figure 5.9. This reduced speed on the motorway results in a reduced performance (higher value of P_{TTS}) for the controller with the smaller gain.

We observe that the traffic density on the motorway realized by the controller with the larger gain remains substantially lower than the traffic density realized by the controller with the smaller gain. The traffic density realized by the ALINEA controller with the higher gain oscillates around the setpoint $\hat{\rho}$. The higher the gain of the controller is chosen, the larger the frequency of the oscillations. This is a known phenomenon in control theory [43]. If the gain of the controller is chosen extremely large, the resulting ALINEA controller implements bang-bang control of the on-ramp. This phenomenon can already be observed in the large-gain example in Figure 5.9 where nearly all metering rates provided by the controller are either 0 or 1.

The presence of oscillations of the metering rate has an influence on the traffic conditions on the motorway. In the lower two plots in Figure 5.9, the traffic density and the average speed in the segment fed by the fourth on-ramp are presented. The larger the gain of the ALINEA controller, the stronger the controller responds to densities larger than $\hat{\rho}$ and the better the P_{TTS}. But the oscillations in the average speed and the traffic density and the frequency of the oscillations increase as well. These oscillations in the average speed on the motorway need to be suppressed as much as possible since they can result in dangerous traffic situations if their amplitude and frequency becomes too large.

The frequency of the oscillations is also observed to increase with increasing traffic demand at the on-ramp. Hence, in order to determine a value of K_R that is safe under all circumstances, the simulations need to be performed with a traffic demand that is equal or higher than the highest expected traffic demand.

A trade-off needs to be made between the performance P_{TTS} of the ALINEA controller and the suppression of oscillations in the metering rate. An ALINEA controller with a controller gain of $K_R = 0.05\,\text{km lane/veh}$ seems to be a good trade-off for the fourth on-ramp of the E17 motorway. This is also confirmed in [112], where a value of $0.035\,\text{km lane/veh}$ (value after conversion) is proposed. The resulting on-ramp flow, metering rate, queue length, and traffic density and average speed on the motorway for the ALINEA controller with $K_R = 0.05\,\text{km lane/veh}$ are presented in Figure 5.10.

As can be observed in the second plot of Figure 5.10, the ALINEA-based controller with $K_R = 0.05\,\text{km lane/veh}$ is overridden several times during the simulation period. The jumps of the metering rate to 1 associated with overriding the controller occur instantaneously while the change in the metering rate due to the ALINEA control law in (4.2) is much smoother. The smooth transition of the metering rate after the controller has been overridden results in a period during which the metering rate is rather high. During this high metering rate transition period, the queue at the on-ramp partially dissolves as can be seen in the third plot in Figure 5.10. When comparing the duration of the transition period after the controller has been overridden for the ALINEA controller with the smaller gain ($K_R = 0.0005\,\text{km lane/veh}$) in Figure 5.9 and for the ALINEA controller presented in Figure 5.10, we observe that the transition period becomes smaller with increasing gain K_R of the controller. Since during the overriding of the controller the metering rate is set equal to 1, the service rate of the long on-ramp queue is determined by the state of the traffic on the motorway according to (3.10). Indeed, with $r_o(k) = 1$, (3.10) becomes equivalent to (3.9), i.e., we get the service rate of the on-ramp without control. As a result of the increased service rate of the on-ramp, the density on the motorway increases. This effect can be observed in the fourth plot of Figure 5.10. The increase of the traffic density on the motorway results in a decrease of the average speed in the segment. As can be seen in the last plot in Figure 5.10, the decrease in the average speed due to overriding the controller occurs in a time frame of approximately 5 minutes and the reduction in average speed is $10\,\text{km/h}$. From a safety point of view, the impact of the controller on the traffic operation on the motorway does not result in dangerous traffic situations. From Table 5.1, we know that the performance of the ALINEA controller with $K_R = 0.05\,\text{km lane/veh}$ is $P_{\text{TTS}} = 1942\,\text{veh h}$ for the simulated morning rush hour.

Remark 5.2 As we presented in the experiment description in Section 5.4, the maximal queue length was chosen equal to 100 vehicles. From the simu-

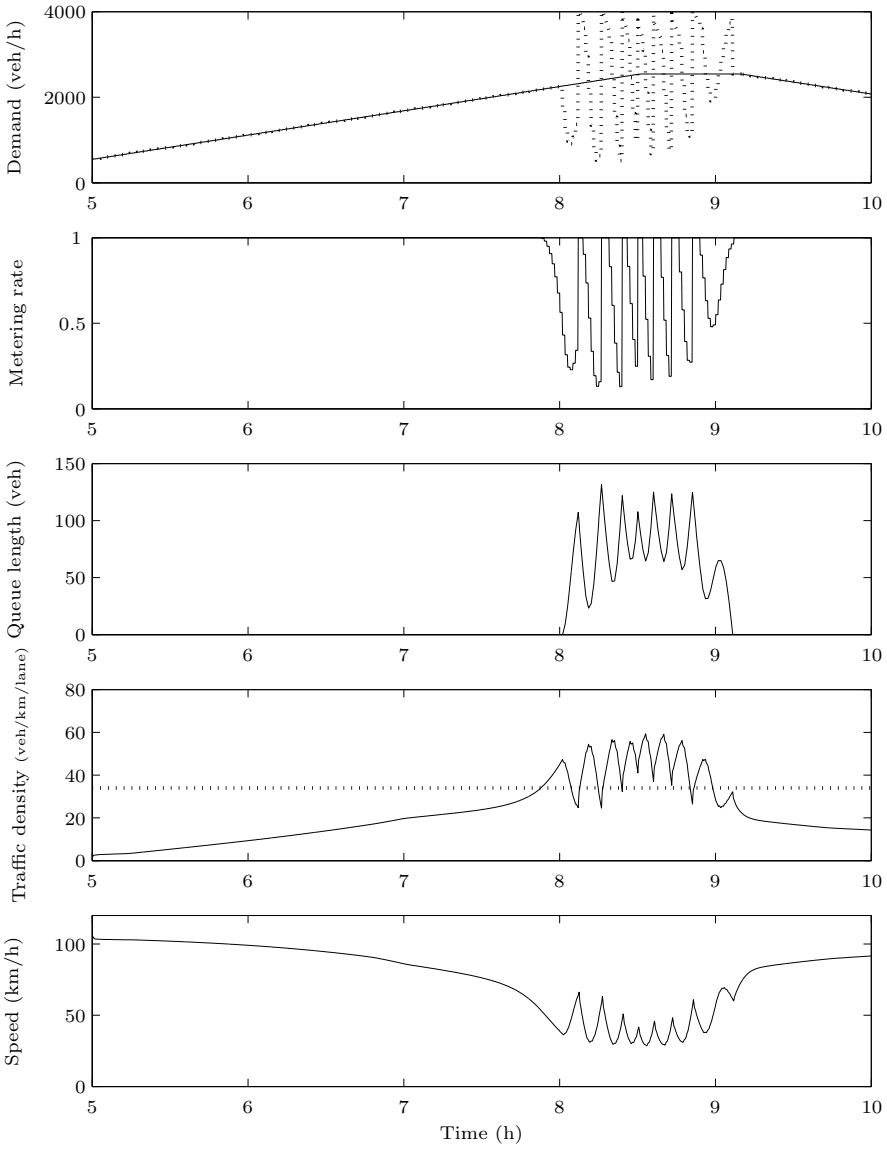


Figure 5.10: ALINEA-based ramp metering control on the fourth on-ramp with $K_R = 0.05\,\mathrm{km\,veh/lane}$. The dotted line in the upper plot represents the real on-ramp traffic flow while the solid line represents the on-ramp demand. The dotted line in the density graphs represents the critical density ρ_{crit}.

K_R (km lane/veh)	P_{TTS} (veh h)
0.0001	2038
0.0005	2039
0.001	2035
0.005	2010
0.01	2003
0.05	1977
0.1	1979
0.5	1967
1	1969

Table 5.2: Overview of the TTS (P_{TTS}) by the vehicles on the E17 motorway and on its on-ramps during the simulated morning rush hour, realized by an ALINEA ramp metering controller on the fifth on-ramp. The ALINEA controller gain K_R was varied.

lation results of ALINEA-based ramp metering control on the fourth on-ramp presented in Figures 5.9 and 5.10, we observe that this queue length constraint is violated by the ALINEA controllers. In order to avoid this queue length constraint violation the threshold at which the ALINEA controller is overridden must be chosen conservatively. However, a more conservative choice of this threshold results in a loss of ALINEA controller performance. □

ALINEA-based ramp metering control on the fifth on-ramp

In a similar way the effect of ALINEA-based ramp metering on the fifth on-ramp of the E17 motorway (Figure 5.2) was simulated. The phenomena that occur and the reasoning are the same as the ones described for the fourth on-ramp. The simulation results are summarized in Table 5.2.

Although the peak traffic volume on the fifth on-ramp is smaller than on the fourth (Figure 5.6), the parameter K_R that realizes a trade-off between the performance of the controller (Table 5.2) and the suppression of oscillations is also found to be 0.05 km lane/veh. The simulation results of the controller are presented in Figure 5.11.

As presented in Table 5.2, a value K_R of 0.05 km lane/veh results in a controller with a performance of $P_{\mathrm{TTS}} = 1977$ veh h which is slightly better than a controller with $K_R = 0.01$ km lane/veh ($P_{\mathrm{TTS}} = 2003$ veh h). As can be observed in the second plot of Figure 5.11, the ALINEA-based controller with $K_R = 0.05$ km lane/veh is able to keep the traffic density on the motorway approximately equal to the critical density as long as the constraint on the queue length does not become active. If the constraint on the queue length becomes active, the metering rate oscillates between $r_{\min} = 0$ and 1. The period of the oscillation is approximately 8 minutes and the amplitude of the oscillations in

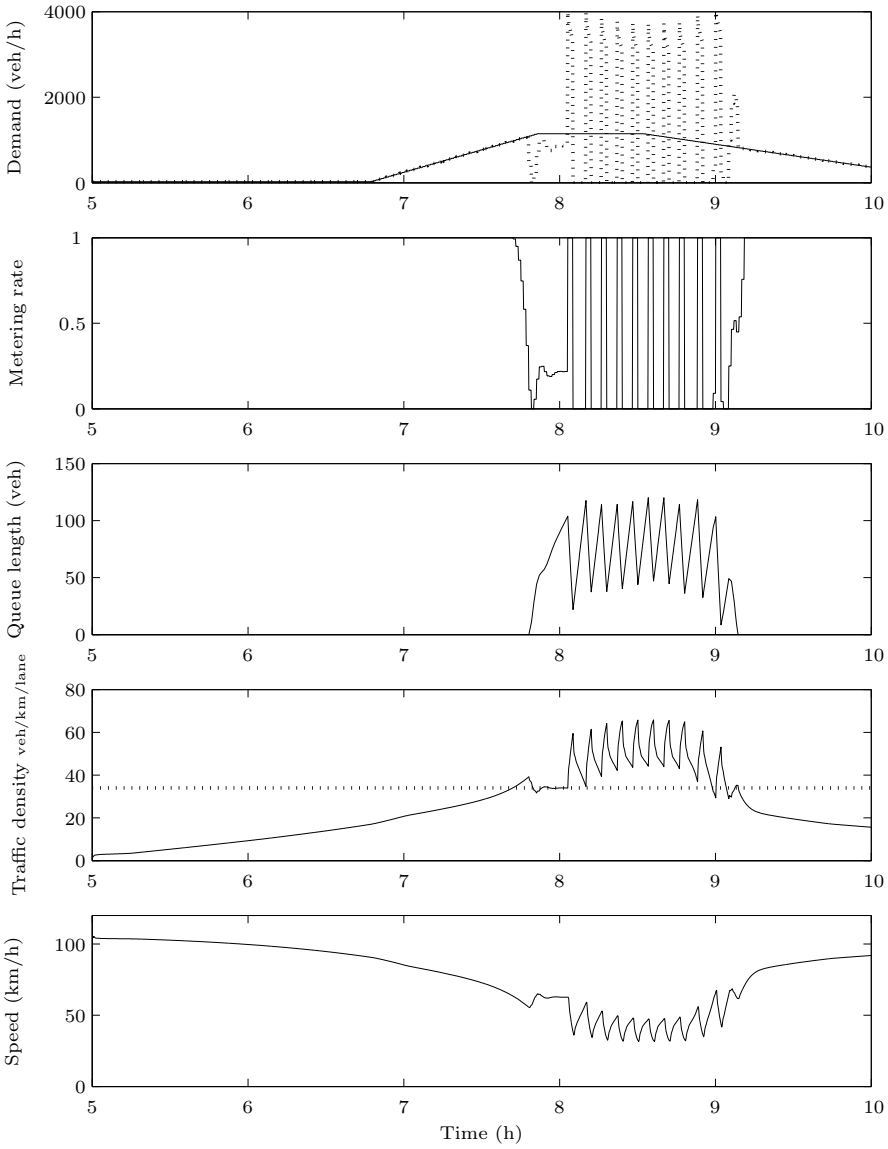


Figure 5.11: ALINEA-based ramp metering control on the fifth on-ramp with $K_R = 0.05\,\mathrm{km\,lane/veh}$. The dotted line in the upper plot represents the real on-ramp traffic flow while the solid line represents the on-ramp demand. The dotted line in the density graphs represents the critical density ρ_{crit}.

the average speed is 10 km/h. From a safety point of view, the impact of the controller on the traffic operation on the motorway does not result in dangerous traffic situations.

ALINEA-based ramp metering control on the fourth and on the fifth on-ramps

After simulations of ALINEA-based ramp metering control on the fourth and on the fifth on-ramps, an experiment was conducted with ALINEA-based ramp metering control on the fourth and on the fifth on-ramp simultaneously. Both ALINEA controllers found for the fourth and the fifth on-ramp were implemented and a simulation of the morning rush hour was performed. The simulation results are presented in Figure 5.12.

Both controllers are overridden due to the constraint on the queue length as was the case in the simulations of the ALINEA-based controllers on the fourth or the fifth on-ramp. The performance of the traffic system with ALINEA-based traffic control on the fourth and the fifth on-ramps is $P_{\mathrm{TTS}} = 1956\,\mathrm{veh\,h}$. The performance of the simultaneous application of the ALINEA-based controllers with $K_R = 0.05\,\mathrm{km\,lane/veh}$ on the fourth and on the fifth on-ramps lies in between the performance of the ALINEA-based ramp metering controller with $K_R = 0.05\,\mathrm{km\,lane/veh}$ on the fifth on-ramp alone ($P_{\mathrm{TTS}} = 1977\,\mathrm{veh\,h}$) and the performance of the ALINEA-based ramp metering controller with $K_R = 0.05\,\mathrm{km\,lane/veh}$ on the fourth on-ramp alone ($P_{\mathrm{TTS}} = 1942\,\mathrm{veh\,h}$). As can be observed by comparing Figures 5.10 and 5.11 with Figure 5.12, the amplitude of the oscillations in the density and the average speed in the segments fed by the on-ramps increases if ALINEA control is applied to both on-ramps simultaneously. The disturbances caused by overriding one controller can propagate on the motorway and enhance the disturbances caused by the other controller.

5.6.3 Summary of the simulation results for ALINEA-based control

ALINEA-based ramp metering was applied to the fourth on-ramp, to the fifth on-ramp, and to the fourth and to the fifth on-ramps simultaneously. From Tables 5.1 and 5.2 it was observed that the performance of ALINEA-based control on the fourth on-ramp systematically resulted in a larger performance increase (for a given value of K_R) than ALINEA-based control on the fifth on-ramp. For the simulated morning rush hour and for our choice of the ALINEA controller gain K_R, the performance P_{TTS} of the ALINEA-based (non-coordinated) ramp metering control of on-ramps 4 and 5 simultaneously was found to be worse than the performance of ALINEA-based ramp metering control on on-ramp 4 but better than ALINEA-based ramp metering control on on-ramp 5.

Due to the high traffic density on the motorway and the high on-ramp demands,

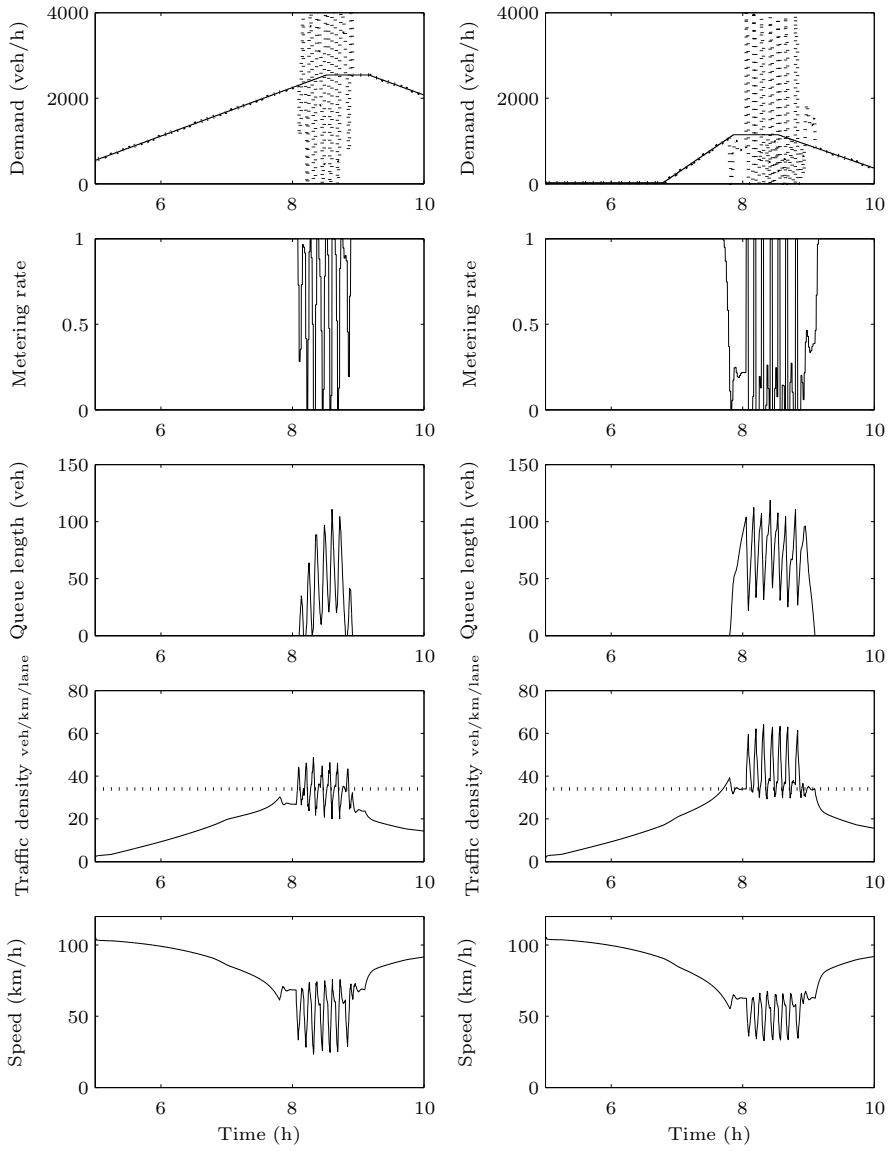


Figure 5.12: ALINEA-based ramp metering control on the fourth and the fifth on-ramps with $K_R = 0.05\,\mathrm{km\,lane/veh}$. The dotted line in the upper plot represents the real on-ramp traffic flow while the solid line represents the on-ramp demand. The dotted line in the density graphs represents the critical density ρ_{crit}.

the queues at the on-ramps grew rapidly resulting in the controllers being overridden. It was observed that if we choose the threshold on the queue length at which the ALINEA controller is overridden equal to the maximal allowed queue length, the queue length constraint is violated. The larger K_R, the faster the response of the metering rate to the traffic density on the on-ramp and the faster the oscillations that occur. These oscillations need to be suppressed such that dangerous traffic conditions cannot occur. In the examples above, a value of K_R was chosen which trades off the reduction of oscillations for the given demands and the performance (TTS) of the controller. Since the frequency of the oscillations increases with increasing traffic demand at the on-ramp, the value of K_R needs to be determined using the highest expected traffic demand. The amplitude of the oscillations of the traffic density and the average speed in the motorway segments fed by the on-ramps was found to be larger for the simulations with two simultaneous ALINEA-based controllers than for the simulations with ALINEA-based control of either the fourth or the fifth on-ramp.

5.7 Model predictive control

5.7.1 Experiment set-up

In a model predictive control framework as described in Section 4.4, the metering rates for the controlled on-ramps are computed using optimization. For every metered on-ramp, N_c metering rates are computed such that an objective function is minimized for the prediction horizon N_p. Only the first metering rate found for every on-ramp is applied to the system. Next, the prediction and the control horizon are shifted one sample forward, and the traffic states are updated after which the whole process is started all over again during the next iteration step.

Based on the METANET model implementation for the case study motorway stretch presented in Figure 5.2, the expression for the MPC objective function at control step k_0 in (4.3) can be rewritten as:

$$J(k_0) = \sum_{l=l(k_0)}^{l(k_0+N_\mathrm{p})-1} \left[\sum_{(m,i)\in\mathcal{I}_\mathrm{m}} \rho_{m,i}(l) l_{m,i} n_m + \alpha_\mathrm{queue} \sum_{o\in\mathcal{I}_\mathrm{o}} w_o(l) \right] \Delta T_\mathrm{sim} + \alpha_\mathrm{var} \sum_{k=k_0}^{k_0+N_\mathrm{p}-1} \big(r(k) - r(k-1)\big)^2 \ , \tag{5.2}$$

where $\mathcal{I}_\mathrm{m}$ denotes the set of all index pairs (m, i) for motorway segments, and $\mathcal{I}_\mathrm{o}$ is the set of all on-ramps indices.

The optimization of the objective function over the prediction horizon is conducted using the *fmincon* routine from Matlab's optimization toolbox [134]. This optimization routine uses sequential quadratic programming [18,105,115] and allows for an easy incorporation of (nonlinear) constraints in the optimization. The fact that a queue at an on-ramp cannot grow larger than the physically available space is added as a hard constraint to the optimization by limiting the queue lengths at the on-ramps to 100 vehicles or less.

In a receding horizon framework, an optimization of the metering rates is conducted at every iteration step. During the optimization of the metering rates, the optimization algorithm can get stuck in a local minimum and return locally optimal values of the metering rates. This problem can be countered by restarting the optimization process several times with different starting values of the parameters to be optimized. The necessary number of restarts depends on the problem at hand. In the first iteration step, the number of optimization restarts with random starting parameters is chosen fairly large (e.g., 10 or even 20). During the next iterations, the number of restarts, and also the computational complexity, can be reduced by carefully choosing the starting values of the optimization based on the optimization results of the previous iteration step. If the optimization process does not return a feasible solution for any of the restarts, additional restarts can be tried (if the time is available) or the controller can be overridden by setting the metering rates equal to 1. For the simulations presented here, it was found that 3 restarts suffice. This was verified by comparing the values of the objective function for the different restarts. The following three sets of starting values for the optimization of the metering rates were chosen:

1. A set of initial parameters based on the optimization results from the previous controller step. The first sample of the metering rates found during the previous controller step was applied to the system in the previous iteration. In the current controller step, the discarded metering rates from the previous controller step can be used as initial values for the parameters. If

   $$\{\tilde{r}(k-1|k-1), \ldots, \tilde{r}(k+N_c-2|k-1)\}$$

   is the optimal metering rate sequence for control step $k-1$, then

   $$\{\tilde{r}(k|k-1), \ldots, \tilde{r}(k+N_c-2|k-1), \tilde{r}(k+N_c-2|k-1)\}$$

   can be used as the starting point for the optimization at the current control step k. The last value is repeated in order to obtain the required N_c initial values per metered on-ramp.

2. The second set of initial metering rates consists of the minimal metering rates for every metered on-ramp. This way, the optimization starts from

the state where the on-ramp traffic is maximally restricted at all metered on-ramps. This will generally cause queues to form at the on-ramps and force the optimization to relax the metering rates to a higher value.

3. Finally, also a random sequence of initial parameters uniformly distributed over the interval $[r_{\min}, 1]$ is used.

After completion of all the optimizations, the set of parameters resulting in the lowest value of the objective function is chosen.

5.7.2 MPC tuning parameters

In the MPC framework of Section 4.4, four tuning parameters were defined: N_{p}, N_{c}, α_{queue} and α_{var}. We now describe how the parameter values used in the simulation examples in this section were determined.

1. The parameter α_{queue} in (5.2) assigns a weight to the total time spent by the vehicles in the on-ramp queues relative to the total time spent by the vehicles on the motorway. The choice of the value of α_{queue} is a policy decision. In this text α_{queue} is chosen equal to 1 i.e., the same importance is assigned to the time spent by the vehicles in the motorway as to the time spent by vehicles in the on-ramp queues.

2. As we discussed in Section 4.4.1, the values of the parameters N_{p} and N_{c} are a trade-off between the computational complexity and the performance of the controller. In order to determine appropriate lengths of the prediction horizon N_{p} and of the control horizon N_{c}, several simulations were performed. MPC-based ramp metering was simulated on the fifth on-ramp for various lengths of the prediction horizon. We set the parameter α_{var} in the objective function (5.2) equal to 0 in order to eliminate the effects of the smoothing term during the tuning of N_{p} and N_{c}. The performance of the MPC-based controllers was found not to be very sensitive to the length of the prediction horizon. We summarized the performance of the MPC-based controller for the different lengths of the prediction horizon N_{p} in Table 5.3. Figure 5.13 presents an indication of the CPU-time required by the MPC-based controller as a function of the length of the prediction horizon. The presented CPU-times in Figure 5.13 are the CPU-times needed to simulate the case study over a five hour morning rush hour period with an MPC-based ramp metering controller on the fifth on-ramp and with $N_{\mathrm{p}} = N_{\mathrm{c}}$. Since the major part of the CPU-time during the simulations is spent on the optimization in the MPC-based controllers, the CPU-times presented in Figure 5.13 are an indication of the CPU-times required by the MPC-based controllers to compute the metering rates for the simulated five hour period. Due to the increasing

$N_\mathrm{p} = N_\mathrm{c}$	P_TTS (veh h)
2	1929
3	1911
4	1913
5	1912
6	1912
7	1912
8	1911
9	1911
10	1911

Table 5.3: The performance P_TTS (veh h) of an MPC-based ramp metering controller on the fifth on-ramp of the case study motorway for different values of the prediction horizon length N_p. In this experiment the control horizon length N_c is chosen equal to the prediction horizon length N_p and the objective function parameter α_var is set to 0. If the prediction horizon is chosen too short, the performance of the MPC-based controller decreases by the inability to take the future traffic behavior into account. If the prediction horizon becomes longer, the performance of the MPC-based ramp metering controller improves until it stabilizes and does not improve with further increasing prediction horizon although the computational complexity of the controller increases drastically (see Figure 5.13).

number of dimensions in the parameter space, the required computation time for the optimization step in the MPC-based controller increases with increasing length of the prediction horizon. We chose the length of the prediction horizon N_p equal to 10 minutes, which is roughly the travel time through the studied motorway at congested traffic operation as was suggested in Section 4.4.3.

3. In order to limit the computational complexity of the optimization, the control horizon N_c was chosen to be 5 minutes, reducing the number of parameters to be optimized per metered on-ramp from 10 to 5, while preserving the ability of the controller to take the simulated future traffic states of the next 10 minutes into account. We observed in our simulations that the computational complexity of the MPC-based ramp metering controller with a prediction horizon $N_\mathrm{p} = 10\,\mathrm{min}$ and a control horizon $N_\mathrm{c} = 5\,\mathrm{min}$ is sufficiently low in order to be able to apply on-line MPC-based ramp metering control to the case study motorway.

4. The objective function (5.2) is the total time spent with an additional penalty term on variations in the metering rate. The penalty term is weighted by a factor α_var, which allows to put more or less emphasis on a smooth control signal. In order to determine an appropriate value for α_var, MPC-based ramp metering was simulated on the fourth on-ramp,

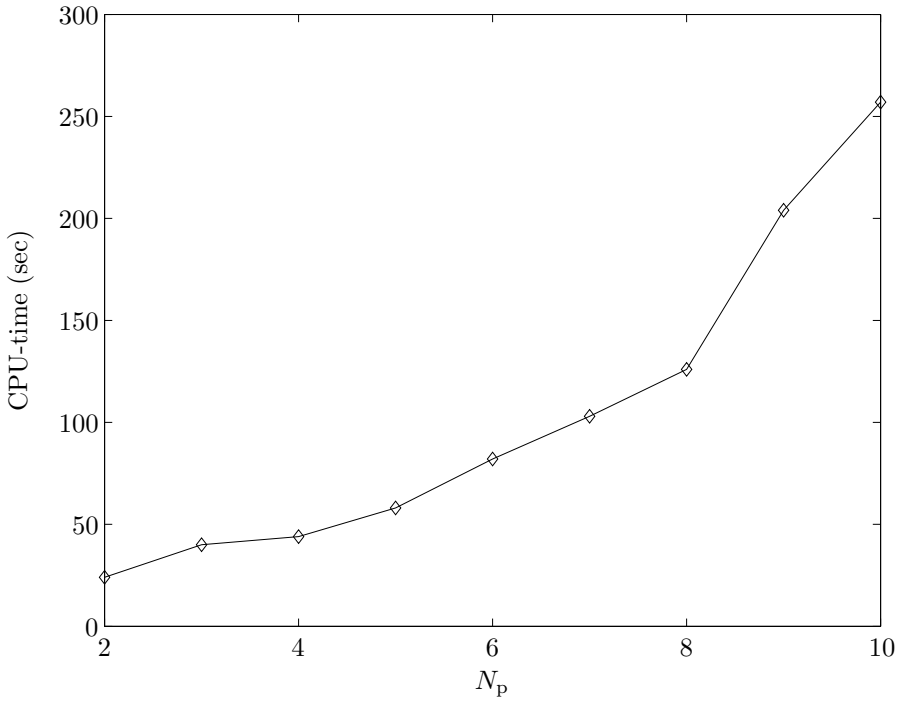


Figure 5.13: The required CPU-times for a simulation of the case study motorway over a five hour morning rush hour period with an MPC-based ramp metering controller on the fifth on-ramp as a function of N_p ($N_\mathrm{p} = N_\mathrm{c}$). The case study METANET model and the MPC-controller are implemented in C and in Matlab version 6.5 and the simulation is performed on a Pentium IV (1 GHz) computer. Since the major part of the CPU-time is spent on the optimization in the MPC-based controller, these CPU-times are an indication of the CPU-times required by the MPC-based controller to compute the metering rates for a five hour period.

on the fifth on-ramp, and on both the fourth and on the fifth on-ramps with varying values of α_{var}.
It was found that the value of α_{var} had very little impact on the simulation results for MPC-based ramp metering on a single on-ramp (fourth or fifth on-ramp) for the simulated morning rush hour. The performance of the MPC-based ramp metering controller for different values of α_{var} is summarized in Table 5.4. Only when an excessively large value of α_{var} was chosen, e.g., $\alpha_{\text{var}} = 1000$, the performance P_{TTS} of the resulting MPC-based controller started to decrease. For values of $\alpha_{\text{var}} = 10000$ and higher, the cost of variations in the metering rate became so high that the metering rate did not change over time anymore and the MPC-based controllers remained inactive (metering rate equal to 1). The control signals found were smooth, even for $\alpha_{\text{var}} = 0$.
In the case non-coordinated ramp metering was implemented on the fourth and on the fifth on-ramps, it was observed that due to the interaction between the non-coordinated MPC-based controllers, the metering rates could oscillate. A value of $\alpha_{\text{var}} = 40$ has a smoothing effect on the metering rates as will be illustrated in the next section. In the case coordinated MPC-based ramp metering control was simulated, the choice of a parameter value of $\alpha_{\text{var}} = 40$ was found to improve the convergence of the optimization resulting in a good performance (see Table 5.4). Hence, the parameter α_{var} is chosen to be equal to 40 for all the MPC-based ramp metering controllers on the fourth and on the fifth on-ramps.

Based on the reasoning above, we choose $N_{\text{p}} = 10$, $N_{\text{c}} = 5$, $\alpha_{\text{queue}} = 1$ and $\alpha_{\text{var}} = 40$ for the simulation experiments in the next section.

5.7.3 MPC-based ramp metering control simulation results

We simulated four configurations of MPC-based ramp metering control on the fourth and on the fifth on-ramps of the case study motorway stretch:

- MPC-based ramp metering on the fourth on-ramp;
- MPC-based ramp metering on the fifth on-ramp;
- Non-coordinated MPC-based ramp metering on the fourth and on the fifth on-ramps;
- Coordinated MPC-based ramp metering on the fourth and on the fifth on-ramps.

Let us now investigate the simulation results of each of these configurations in more detail.

α_{var}	ramp 4	ramp 5	ramp 4 & 5 (nc)	ramp 4 & 5 (c)
0	1911	1915	1923	1995
1	1911	1915	1923	1985
10	1911	1915	1924	1953
20	1911	1915	1925	1911
30	1911	1916	2037	1911
40	1911	1915	1927	1911
50	1911	1915	1929	1911
100	1911	1917	1935	1911
500	1912	2018	1913	1912
1000	1988	2038	1988	1987
10000	2038	2038	2038	2038

Table 5.4: Impact of the objective function parameter α_{var} on the performance P_{TTS} (veh h) of an MPC-based controller. The parameter α_{var} in (5.2) weighs the penalty term of variations in the metering rate. The case study network is simulated over a morning rush hour and four ramp metering configurations are investigated: ramp metering on the fourth on-ramp, ramp metering on the fifth on-ramp, non-coordinated ramp metering on the fourth and on the fifth on-ramps (nc), and coordinated ramp metering on the fourth and on the fifth on-ramps (c).

MPC-based ramp metering on the fourth or on the fifth on-ramp

Figure 5.14 shows the results of a simulation of the E17 motorway case study with MPC-based ramp metering control on the fourth on-ramp. The values of the controller parameters N_{p} and N_{c} and the objective function parameters α_{queue} and α_{var} are chosen as presented above. The solid line in the first plot from the top in Figure 5.14 represents the traffic demands at the fourth on-ramp during the simulated period. The dotted line in the first plot represents the service rate or the traffic flow that is allowed to enter the motorway through the fourth on-ramp according to (3.10). If the traffic demand is larger than the service rate, a queue is formed at the on-ramp. If the traffic demand is smaller than the service rate, the on-ramp queue is dissolved. The evolution of the queue length over time at the fourth on-ramp is represented in the third plot in Figure 5.14. The second plot represents the evolution of the metering rate for the fourth on-ramp over time. The fourth and the fifth plot in Figure 5.14 represent the traffic density $\rho_{9,1}$ and the average speed $v_{9,1}$ in segment 1 of link 9 of the case study motorway stretch. It can be seen in Figure 5.2 that this is the first segment of the link fed by the metered fourth on-ramp.

The metering rate in Figure 5.14 remains 1 until the MPC-based controller predicts that the traffic density on the motorway will reach the critical density ρ_{crit} and then the metering rate drops, thus limiting the number of vehicles that is allowed to enter the motorway through the fourth on-ramp. A queue

starts to form at the fourth on-ramp until the maximal queue length is reached. It can be seen in Figure 5.14 that the maximal queue length of 100 vehicles is never exceeded as opposed to the ALINEA case from the previous section. If the maximal queue length of 100 vehicles is reached, the metering rate must be adjusted such that the service rate of the queue equals the demand of the on-ramp[1]. This behavior of the metering rate can be explained by looking at the equation for the service rate of the on-ramp (3.10). The metering rate cannot be larger than the value corresponding to the demand since this would increase the traffic density on the motorway and thus decrease the throughput. The metering rate can also not be smaller than the value corresponding to the demand (although it would probably correspond to a smaller value of the objective function) since this would result in a queue longer than the maximal allowed queue length. Looking at the traffic density in Figure 5.14, it is observed that the ramp metering controller is able to avoid the occurrence of too high a traffic density with the corresponding lower traffic flow in the segment fed by the fourth on-ramp. Near the end of the simulated rush hour period, the traffic demands start to decrease and the on-ramp queue starts to dissolve (see Figure 5.14). We observe that the metering rate remains approximately 0.75 after the peak traffic demand is over. However, the metering rate set-up does not restrict the access of vehicles through on-ramp 4 to the motorway. This can be observed in Figure 5.14 since the traffic demand equals the service rate of the queue given by (3.10) and the queue length remains constantly 0 after 9 a.m. Since the ramp metering set-up does not restrict the access to the motorway anymore, the metering rate can be set to 1, which comes down to switching the ramp metering set-up off after the rush hour. Note that the decision to swith the ramp metering set-up off under these circumstances can be taken by the low-level controller (see Figure 4.4).

Comparing the two lower plots in Figure 5.14, which show the traffic density and the average speed in the segment fed by the fourth on-ramp respectively, with the plots of the traffic density and the average speed in the same segment using ALINEA control (see Figure 5.10), it is clear that the metering rates and the evolution of the average speed and the traffic density on the motorway are much smoother in the MPC case than in the ALINEA case. This is a major argument for the use of MPC-based ramp metering over ALINEA-based ramp metering despite the higher computational complexity of MPC-based ramp metering.

By adding the queue length constraints as hard constraints to the optimization problem, the queue length constraints are strictly respected using MPC as opposed to the ALINEA case presented in the previous section. In case the constraints on the queue lengths are not added as hard constraints to the opti-

[1]Note that under extreme circumstances, e.g., for persistent high demands, the constrained optimization problem in the MPC controller can become infeasible. Under these extreme circumstances, a constraint relaxation can be done in order to obtain a feasible solution [25].

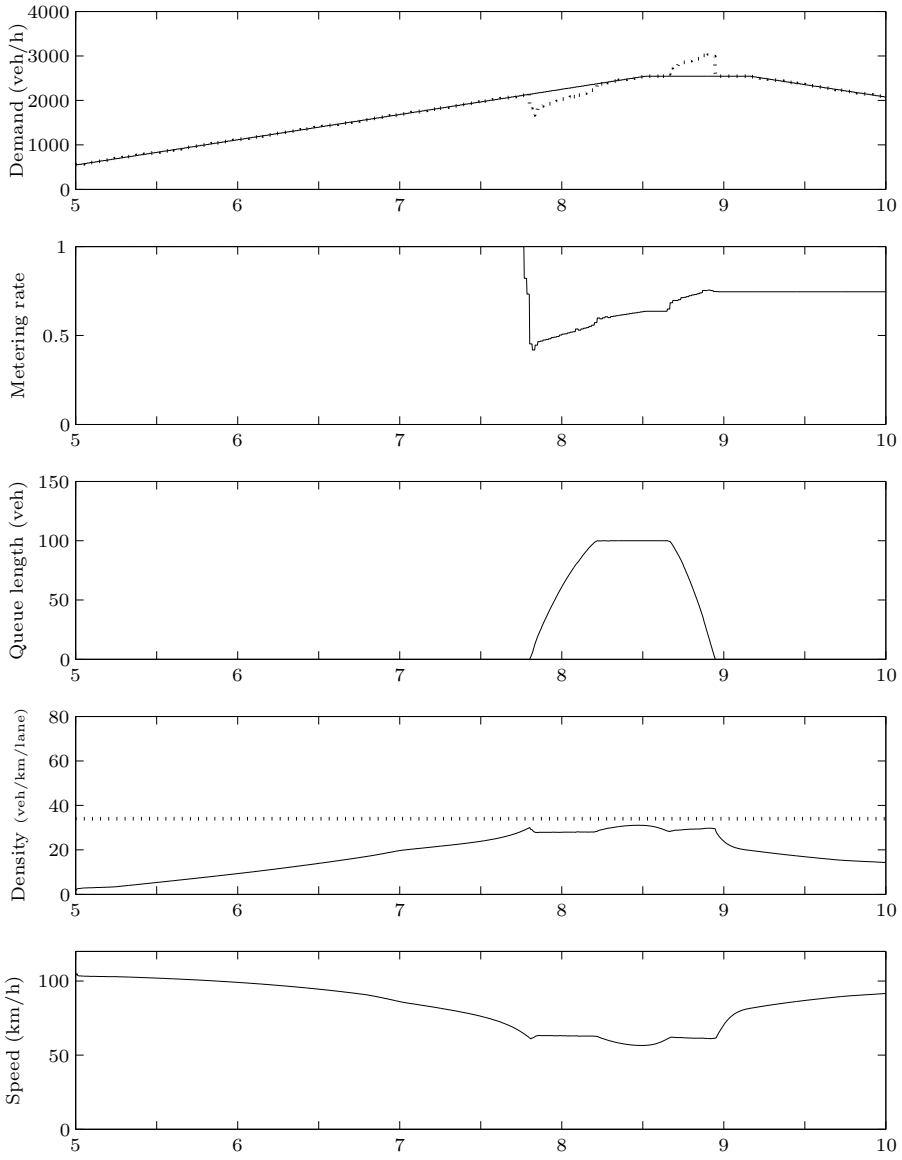


Figure 5.14: Simulation results of MPC-based ramp metering on the fourth on-ramp of the E17 Ghent–Antwerp. The dotted line in the upper plot represents the real on-ramp traffic flow while the dotted line in the traffic density plot represents the critical density ρ_{crit}. The lower two plots present the traffic density $\rho_{9,1}$ and the average speed $v_{9,1}$ in the segment fed by the fourth on-ramp respectively.

mization problem but as an additional penalty term in the objective function as was done in [79], the compliance with the constraints cannot be guaranteed. Adding the constraints to the optimization problem as was done in that paper guarantees compliance of the MPC-based controller with the queue length constraints.

We also simulated MPC-based ramp metering at the fifth on-ramp of the case study motorway presented in Figure 5.2 in a similar way as at the fourth on-ramp. The reasoning and the phenomena that occur are similar to these presented above. However, due to the different location of the fifth on-ramp and due to the different on-ramp demands, the performance for the simulated morning rush hour that can be attained by the MPC-based on-ramp controller at the fifth on-ramp ($P_{\mathrm{TTS}} = 1915\,\mathrm{veh\,h}$) is slightly lower than the performance that can be attained by MPC-based ramp metering at the fourth on-ramp ($P_{\mathrm{TTS}} = 1911\,\mathrm{veh\,h}$). This phenomenon can be observed for different values of α_{var} as shown in Table 5.4.

Non-coordinated MPC-based ramp metering on the fourth and on the fifth on-ramps

Some simulation experiments with non-coordinated ramp metering on the fourth and on the fifth on-ramps were conducted for varying values of the objective function parameter α_{var}. As can be seen in Table 5.4, the performance of non-coordinated ramp metering on both on-ramps is in general worse than a single MPC-based ramp metering set-up on the fourth on-ramp.

Both controllers on the fourth and on the fifth on-ramp use the same initial conditions in the prediction model but they do not exchange information about the evolution of the metering rates during the prediction horizon. Under certain conditions, this leads to a combination of non-coordinated control actions which can lead to sub-optimal traffic control.

An example of sub-optimal control due to two non-coordinated MPC-based on-ramp controllers is presented in Figure 5.15. In this example, two non-coordinated MPC-based ramp metering controllers one on the fourth and one on the fifth on-ramp of the case study motorway are simulated. The objective function parameter α_{var} is chosen equal to 40 as presented before. As we can see in Figure 5.15, both controllers start metering the on-ramps, which results in a queue at both on-ramps. Due to the presence of both non-coordinated controllers, more vehicles are stored in the on-ramp queues than in the previous simulation examples, which results in an increased time spent in the queues. This increased time spent in the queues is not compensated by an increased throughput on the motorway and thus the performance of the non-coordinated MPC-based controllers ($P_{\mathrm{TTS}} = 1927\,\mathrm{veh\,h}$) in this simulation example is lower than the performance of the single MPC-based controllers in the previous sim-

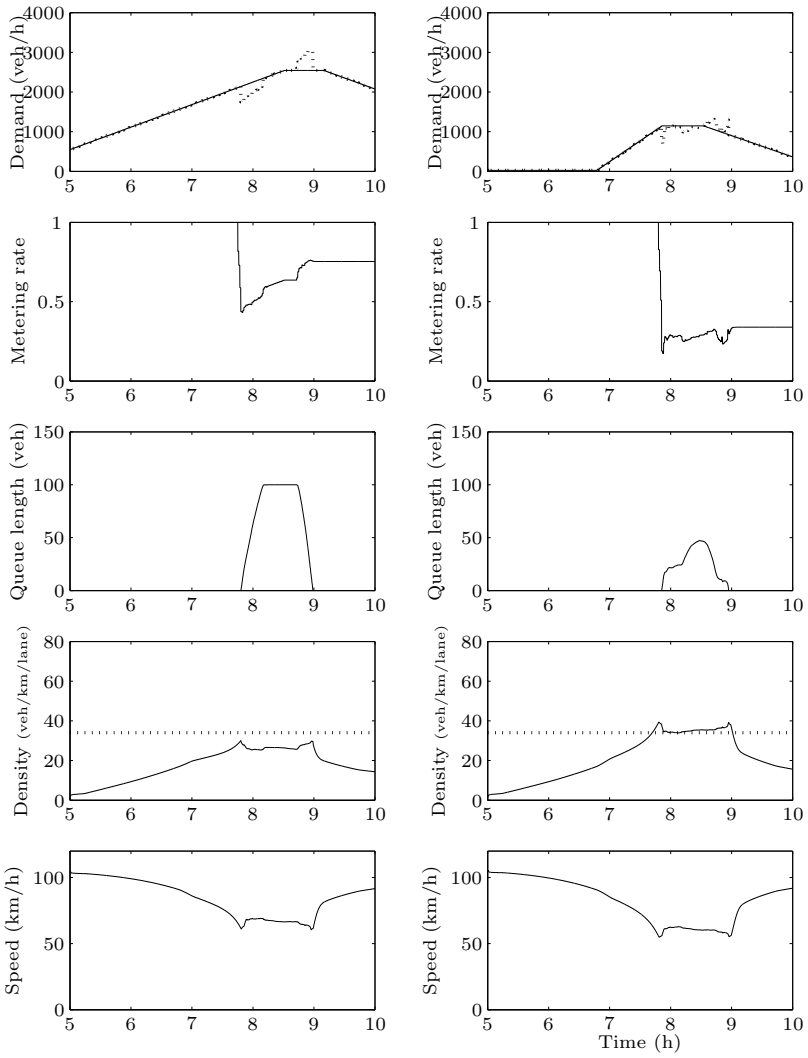


Figure 5.15: Simulation results of simultaneous, non-coordinated MPC-based ramp metering control on the fourth on-ramp and on the fifth on-ramp of the E17 Ghent–Antwerp case study motorway stretch. The dotted lines in the upper plots represent the real on-ramp traffic flows while the dotted lines in the traffic density plots represent the critical density ρ_{crit}. The traffic states corresponding to the fourth on-ramp are presented in the left column and the traffic states corresponding to the fifth on-ramp are presented in the right column. The four lower plots represent the traffic densities and the average speeds in the segments fed by the fourth (left) and the fifth (right) on-ramps.

ulation examples ($P_{\mathrm{TTS}} = 1911\,\mathrm{veh\,h}$) and ($P_{\mathrm{TTS}} = 1915\,\mathrm{veh\,h}$).

Coordinated MPC-based ramp metering on the fourth and on the fifth on-ramps

In general, coordination of control contributes to an additional increase in the performance of the system. The MPC-based controllers for the fourth and the fifth on-ramps presented in Figure 5.15 were not coordinated. Coordination of the MPC-based controllers on these on-ramps was investigated. An overview of the coordinated MPC-based ramp metering control experiments for varying values of α_{var} is presented in Table 5.4. It can be observed in Table 5.4 that the performance P_{TTS} of the coordinated controller improves with increasing value of α_{var}. For small values of α_{var}, the optimization easily gets stuck in local minima of the objective function (5.2). However, if the value of α_{var} becomes too large, the cost of variations of the control signal is so high that the performance of the controller decreases as was already discussed above.

The value of the objective function parameter α_{var} was chosen equal to 40 and the simulation results are presented in Figure 5.16. It was found that the MPC-based coordination of the ramp metering set-ups on the fourth and on the fifth on-ramps of the case study motorway did not enhance the performance (P_{TTS} = 1911 veh h) of the system for the simulated morning rush hour, although the computational complexity did increase. This can be understood when looking at the metering rates in Figure 5.16. The traffic demand during the morning rush hour is such that ramp metering on the fourth on-ramp can optimally control the traffic operation on the motorway. Although the maximal queue length is reached and the service rate of the controller on the fourth on-ramp needs to be equal to the traffic demand for about half an hour, the ramp metering set-up on the fifth on-ramp does not become active. The metering rate at the fifth on-ramp does drop, but the reduced metering rate at the fifth on-ramp does not result in a restriction of the on-ramp traffic at the fifth on-ramp since no queue is formed. Obviously, if during coordinated MPC-based ramp metering control the fifth ramp metering set-up does not become active, the same performance can be achieved by MPC-based ramp metering on the fourth on-ramp only. Note that for a higher demand, or for a more stringent constraint on the queue length, both coordinated ramp metering set-ups become active in order to keep the traffic operation on the motorway optimal.

Summary of the simulation results for MPC-based control

We applied MPC-based ramp metering control to the fourth on-ramp, to the fifth on-ramp, and to both the fourth and the fifth on-ramps of the case study motorway simultaneously. Both non-coordinated and coordinated MPC-based control of the fourth and the fifth on-ramps was investigated. First, the values of the MPC tuning parameters N_{p}, N_{c}, α_{queue} and α_{var} that result in a good

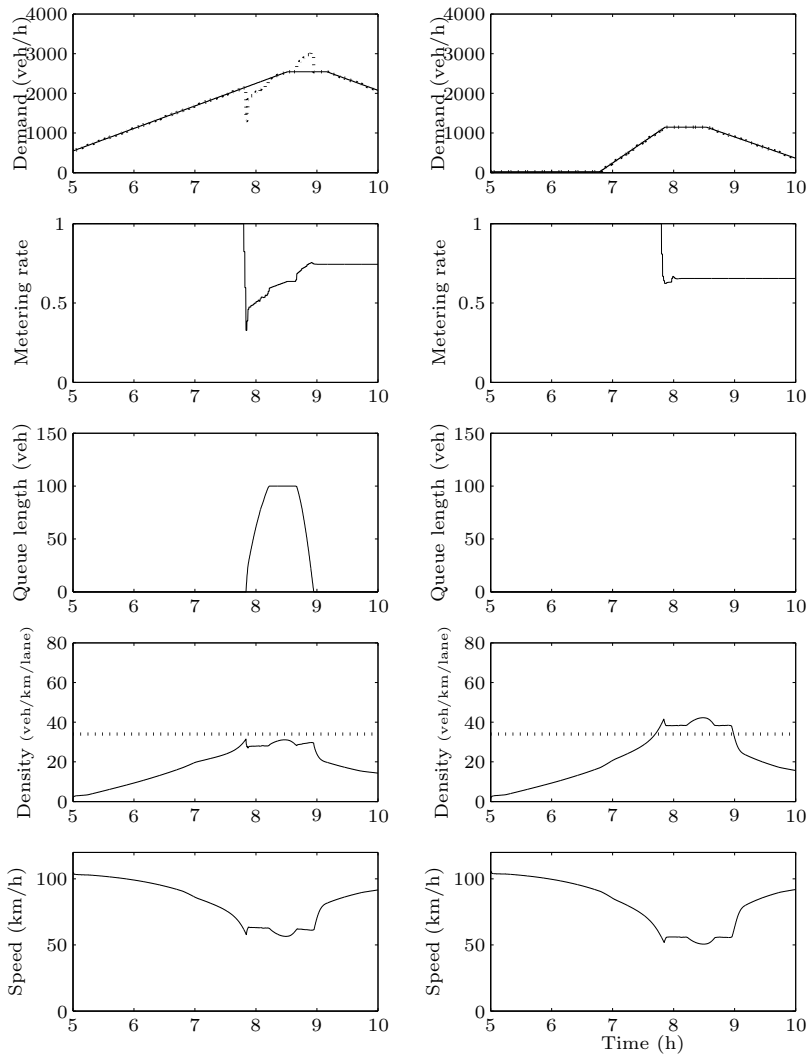


Figure 5.16: Simulation results of coordinated MPC-based ramp metering control on the fourth and the fifth on-ramps of the E17 Ghent–Antwerp case study motorway stretch. The dotted lines in the upper plots represent the real on-ramp traffic flows while the dotted lines in the traffic density plots represent the critical density ρ_{crit}. The traffic states corresponding to the fourth on-ramp are presented in the left column and the traffic states corresponding to the fifth on-ramp are presented in the right column.

controller performance for the case study traffic network were determined.

We observed that the performance of MPC-based ramp metering on the fourth on-ramp results in a higher performance (lower P_{TTS}) than MPC-based ramp metering on the fifth on-ramp. This was also observed for ALINEA-based control presented in the previous section. This difference is due to the different locations of the on-ramps and due to the the difference in the traffic demands at the on-ramps (see Figure 5.6).

The performance of non-coordinated MPC-based ramp metering control on the fourth and the fifth on-ramps was found to be worse than the performance of MPC-based ramp metering control on either the fourth or the fifth on-ramp. Due to the lack of communication between both MPC-based controllers, each controller limits the access to the motorway more than is required for an optimal traffic operation of the case study motorway system. This results in increased waiting times in the on-ramp queues, which are not compensated by an increased throughput on the motorway.

The coordinated MPC-based control of the fourth and the fifth on-ramps results in the same performance as MPC-based ramp metering control on the fourth on-ramp alone. This is due to the fact that the ramp metering controller at the fourth on-ramp alone is perfectly capable of optimally controlling the traffic for the simulated demands. The ramp metering set-up at the fifth on-ramp does not become active during coordinated operation. Based on this result we conclude that the most economical strategy to control the traffic on the case study motorway using ramp metering is to implement MPC-based ramp metering only on the fourth on-ramp since this results in the best performance at a minimal implementation cost. Note however that this recommendation is only valid if the maximal allowed queue lengths are 100 vehicles or more and if the traffic demands are smaller than or equal to the traffic demands used in the simulations above. However, since the traffic demands used in the simulations above were based on traffic measurements of four working days one can assume that the traffic demand will not consistently and substantially increase in the short term. For more stringent constraints on the maximal queue lengths and/or for higher traffic demands, the coordination of the fourth and the fifth on-ramps may be necessary to obtain optimal traffic operation.

Comparison of ALINEA-based and MPC-based ramp metering control for the case study motorway

To conclude this section, we compare the most important properties and the simulation results for ALINEA-based and MPC-based ramp metering control on the E17 motorway case study. An overview of the simulation results is provided in Table 5.5.

Controller	On-ramp	P_{TTS} (veh h)	TTS reduction (veh h)
No control	-	2038	0 (0 %)
ALINEA	4	1942	96 (4.7 %)
	5	1977	61 (3.0 %)
	4 and 5	1956	82 (4.0 %)
MPC	4	1911	127 (6.2 %)
	5	1915	123 (6.0 %)
	4 and 5 (nc)	1927	111 (5.5 %)
	4 and 5 (c)	1911	127 (6.2 %)

Table 5.5: Overview of the TTS on the E17 motorway case study for different ramp metering controllers. The last column presents the realized reduction in TTS compared to the no control case. The reduction relative to the no control case is presented between brackets. ALINEA-based and MPC-based ramp metering control is considered on the fourth on-ramp, on the fifth on-ramp, and on both the fourth and the fifth on-ramps of the case study motorway. Both non-coordinated (nc) and coordinated (c) MPC-based ramp metering control is considered.

- The MPC-based controller outperforms the ALINEA-based controller if the performance criterion P_{TTS} is considered. In Table 5.5 we see that the ALINEA-based controller for the fourth on-ramp of the case study motorway that we developed and tuned in Section 5.6 results in a reduction of 96 veh h (4.7 %) during the simulated rush hour compared to the no-control case. The reduction of the TTS that is realized by an MPC-based ramp metering controller at the fourth on-ramp of the case study motorway is 127 veh h (6.2 %) compared to the no-control case. An ALINEA-based ramp metering controller on the fifth on-ramp of the case study motorway is able to reduce the TTS by 61 veh h (3.0 %) compared to the no-control case. The MPC-based ramp metering controller on the fifth on-ramp results in a reduction of the TTS with 123 veh h (6.0 %).
- Since an MPC-based controller can easily be extended to implement coordinated control of ramp metering set-ups, we compared both non-coordinated and coordinated MPC-based ramp metering control for the fourth and the fifth on-ramps of the case study motorway. We also implemented two non-coordinated ALINEA-based ramp metering controllers on the fourth and the fifth on-ramps. From Table 5.5 we learn that non-coordinated MPC-based control performs worse than coordinated MPC-based control but better than ALINEA-based control on both on-ramps.
- We observed that while the ALINEA-based controller was not able to strictly respect the hard constraints on the queue lengths at the on-ramps (despite the overriding of the controller), the MPC-based ramp metering controller did never violate the queue length constraints. If the

ALINEA-based controller must be guaranteed to respect the queue length constraints, the queue length at which the controller is overridden must be set conservatively which reduces the controller performance.

- We compared the control signals and the evolution of the traffic states on the motorway for both ALINEA-based and MPC-based ramp metering control. We found that the evolution of the control signals (metering rates) and the resulting traffic states on the motorway over time are much smoother for the MPC-based controllers than for the ALINEA-based controllers.

- The computational complexity of an ALINEA-based controller is lower than the computational complexity of an MPC-based ramp metering controller. The optimization of the objective function over the prediction horizon is the most computationally intensive step in the MPC-based controller. In the receding horizon framework of the MPC-based controller, this optimization must be repeated for every controller step. However, by appropriately choosing the controller parameters ΔT_{ctrl}, N_{p} and N_{c}, the computational complexity of the MPC-based controller can be reduced.

Remark 5.3 The relative reduction of P_{TTS} that is realized by the ramp metering controllers in the simulations above seems rather small (see Table 5.5). As we can see in (5.1), the performance measure P_{TTS} computes the TTS in the whole case study network during the total simulation period. However, if we look e.g. at Figure 5.16, we observe that the ramp metering controller is only active for approximately one hour. The TTS spent by the vehicles in the case study network during the four hours of the simulation period that the ramp metering set-up was not active is also taken into account in P_{TTS}. This reduces the relative reduction of P_{TTS} due to ramp metering although the positive impact of ramp metering during the period of high traffic demand is significant. □

5.8 Conclusions

The motorway E17 Ghent–Antwerp in Belgium was modeled and ALINEA-based ramp metering control was compared to MPC-based ramp metering control using simulations.

First, the controller parameters of the ALINEA and the MPC-based controllers were determined after which the performance of a number of controllers was compared.

It was observed that the ALINEA-based controllers resulted in a gain in TTS during the simulated rush hour. The controllers were able to limit the queues

approximately to the maximal queue length. The metering rates oscillated due to the constraints on the maximal queue lengths overriding the controller. This results in oscillations in the traffic density and the average speed in the segment fed by the on-ramp. These oscillations need to be suppressed as much as possible. A big advantage of ALINEA is its limited computational complexity.

The MPC-based ramp metering controllers were observed to realize a higher performance (smaller P_{TTS}) than the ALINEA-based controllers. Moreover, the control signals of the MPC-based controllers are very smooth. The traffic density and the average speed in the segment fed by the on-ramps are behaving very smoothly, even during the peak traffic demands during the rush hour. No oscillations occurred. The constraints on the queue lengths at the on-ramps are strictly respected by the MPC-based controllers. The computational complexity of the MPC-based controllers is larger than the computational complexity of ALINEA due to the optimization problem that needs to be solved at every control step. By appropriately choosing the prediction and the control horizons, a trade-off was made between the performance and the computational complexity of MPC-based ramp metering on the E17 motorway in Belgium.

The first contribution of this chapter to the state of the art is the development of a METANET simulation model for a case study consisting of the E17 motorway Ghent–Antwerp using real traffic data. The second contribution of this chapter to the state of the art is the development of MPC-based ramp metering control and its application to the simulation model of the E17 case study. Both MPC-based control of a single on-ramp and MPC-based coordinated control of two on-ramps were investigated. Using simulations, we illustrated that the MPC-based controller was able to improve the performance of the case study motorway compared to the no control case and compared to the ALINEA-based control case. The MPC-based controller was also able to take hard constraints into account.

Chapter 6

Extensions to the MPC ramp metering framework

This chapter presents two new extensions to the MPC ramp metering framework. First, we present the combined MPC-identification approach to ramp metering. This new approach to MPC-based ramp metering control is able to eliminate the influences of non-modeled external disturbances on the performance of the MPC controller by performing a regular re-identification of the prediction model. Next, we develop anticipative MPC-based traffic control. Both extensions are combined in an anticipative MPC-identification approach to ramp metering. Note that although the focus in this thesis is on ramp metering, the extensions presented in this chapter can also be used for other traffic control measures.

6.1 Traffic model re-fitting

The MPC-based approach to ramp metering control uses a traffic simulation model to predict the future behavior of the traffic flows. As the traffic situation on motorways is influenced by various factors that are not all incorporated into the model, a regular re-fitting of the motorway model can significantly improve the accuracy of the control. Therefore, we discuss in this section how the MPC framework for ramp metering can be combined with an iterative identification procedure so that changes in the behavior of the traffic system (e.g., due to changes in the weather conditions (rain, fog, snow, ...), due to incidents, ...) can be taken into account.

First, a simulation example is presented where the impact of a misfit of the

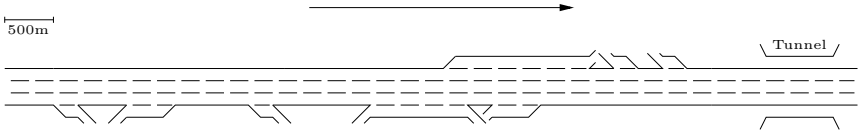


Figure 6.1: Schematic representation of the E17 motorway Ghent–Antwerp in Belgium. The arrow denotes the direction of the traffic flow.

model on the controller performance is investigated for the real-life situation on the E17 motorway Ghent–Antwerp already discussed in Chapter 5. Next, a combined MPC-identification approach to ramp metering control is presented.

6.1.1 Simulation example

In this section, a computer simulation of traffic is used to illustrate the importance of a good fit of the prediction model used in the MPC controller.

The E17 motorway Ghent–Antwerp (see Figure 6.1) is simulated using the METANET model as discussed in Chapter 5. At the fourth on-ramp, MPC-based ramp metering is simulated. The impact of a misfit of the MPC prediction model on the performance of the MPC controller is investigated.

In this simulation example, we use two traffic models: the MPC prediction model that is used in the MPC controller and a simulation model that serves as a reference model. This leads to the situation presented in Figure 6.2. The reference model represents the 'real-world' and is characterized by a parameter set θ_{ref}. The second model is the prediction model that is used in the MPC controller and is characterized by θ_{mpc}.

An MPC-based ramp metering controller is developed and tested for the fourth on-ramp. The parameters of the MPC controller in this simulation are chosen as follows. The length of the prediction horizon N_{p} of the MPC controller is 7 min and the length of the control horizon N_{c} is 5 min. The metering rate provided by the controller is updated every minute ($\Delta T_{\text{ctrl}} = 1$ min). The MPC controller uses a METANET prediction model, provided with a parameter set θ_{mpc}. Note that the MPC framework is generic and that another type of traffic model could be chosen. The MPC objective function is given by (5.2) and the parameters α_{queue} and α_{var} are chosen to be 1 and 10 respectively.

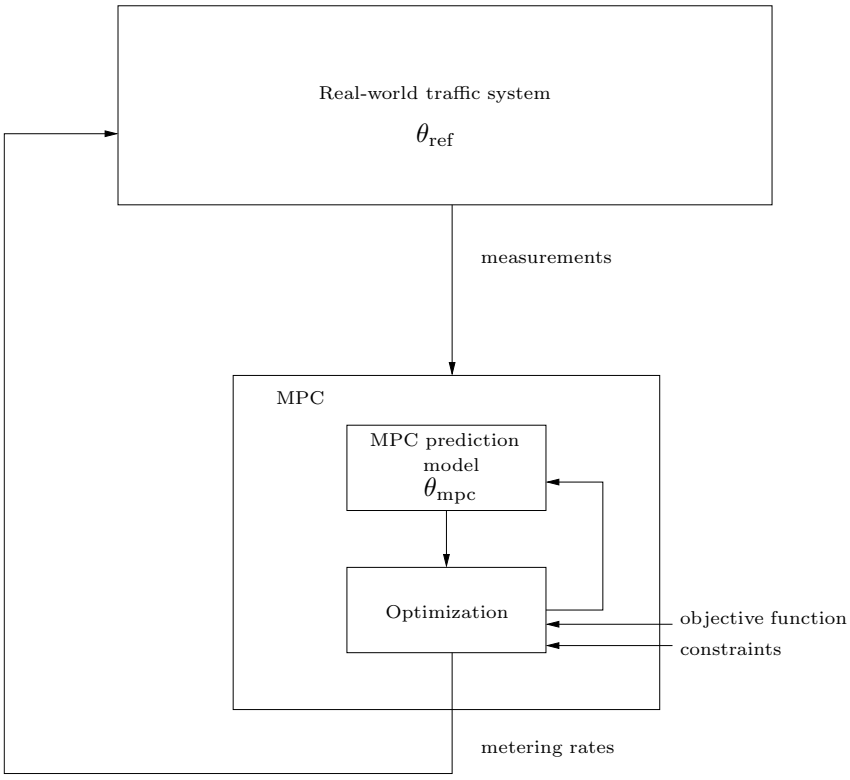


Figure 6.2: Schematic representation of the experiment setup. Two models are used. The first model represents the 'real-world' traffic system. It is a METANET model characterized by the parameter set θ_{ref}. The second model is the MPC prediction model used in the MPC-based ramp metering controller and is characterized by a parameter set θ_{mpc}.

Two situations are investigated:

- $\theta_{\text{mpc}} = \theta_{\text{ref}}$
  In the reference situation, the model parameters θ_{mpc} used in the MPC prediction model are equal to the model parameters θ_{ref} of the 'real-life' simulation model. The reference situation where the MPC controller uses a traffic flow prediction model with parameter set θ_{ref} serves as a reference in the investigation of the sensitivity of the MPC controller performance to perturbations in the prediction model parameters.

- $\theta_{\text{mpc}} = \theta_{\text{perturb}}$:
  In the second situation that is investigated, the controller uses a perturbed parameter set θ_{perturb} in the prediction model, which makes predictions over the prediction horizon N_{p}. The free flow speeds and the critical densities in parameter set θ_{perturb} were chosen about 10 % larger than those in the reference set θ_{ref} of the simulation model. This results in a prediction model that overestimates the capacity of the motorway.

The simulation experiment spans 2.5 hours. The traffic demands on the motorway and at the fourth on-ramp are presented in Figure 6.3. The motorway traffic demand is considered to be constant and equal to 5750 vehicles per hour, while the capacity of the motorway is 6269 vehicles per hour. At the fourth on-ramp a demand peak with a maximal demand of 750 vehicles per hour is simulated.

Since the total traffic demand on the mainline during the peak period is larger than the capacity of the motorway, congestion or a queue (or both) will occur. Figure 6.4 shows the evolution of the traffic density in the segment fed by the metered on-ramp. The solid line in the upper plot shows the evolution of the traffic density realized by the MPC controller in the reference situation ($\theta_{\text{mpc}} = \theta_{\text{ref}}$). The density increases gradually with increasing traffic demand, but once the density becomes too high, the metering rate drops to a lower value as presented in Figure 6.5. The controller starts metering the on-ramp, resulting in the build-up of a queue. The lower plot in Figure 6.4 shows the evolution of the queue length at the fourth on-ramp for both situations. During the operation of the MPC controller in the second situation ($\theta_{\text{mpc}} = \theta_{\text{perturb}}$, dashed line), no queue is formed at the fourth on-ramp (Figure 6.4). However, the rush hour traffic density in the segment fed by the fourth on-ramp is higher for the second situation, resulting in a lower traffic flow[1]. This is due to the fact that the prediction model used in the second situation overestimates the motorway capacity, which results in a metering rate that is too high for the capacity of the motorway (Figure 6.5).

[1] According to the fundamental diagram in Figure 4.1, a higher traffic density results in a lower traffic flow for traffic densities above the critical density ρ_{crit}.

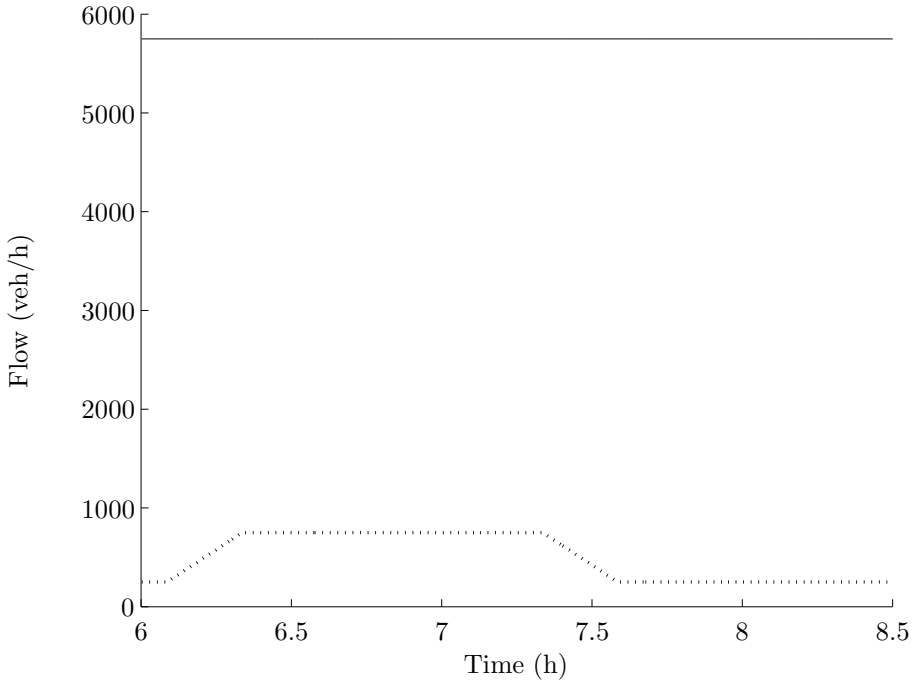


Figure 6.3: Plot of the traffic demand on the motorway (solid line) and of the traffic demand on the fourth on-ramp (dotted line) for the motorway of Figure 6.1.

The performance of both controllers can be assessed by computing the TTS (P_{TTS}) over the 2.5 hour simulation interval according to (5.1). The controller in the reference situation ($P_{\mathrm{TTS,ref}} = 1555$) outperforms the controller in the perturbed situation ($P_{\mathrm{TTS,perturb}} = 1593$). From Figures 6.4 and 6.5 it can be seen that the mode of traffic operation realized by the MPC controllers in both situations is totally different. In the reference situation, the MPC controller keeps the traffic on the motorway flowing smoothly at the cost of a queue at the fourth on-ramp while in the second situation, the MPC controller does not cause a queue at the fourth on-ramp but at the cost of a higher traffic density and delays on the motorway. The difference between both situations originates from the perturbations on the model parameters of the MPC prediction model.

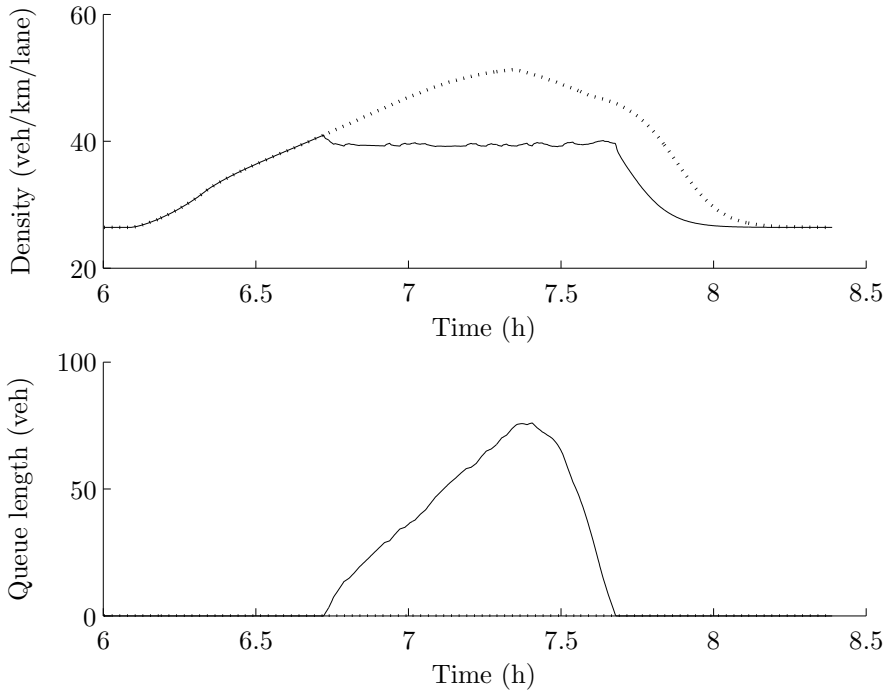


Figure 6.4: Plot of the evolution of the traffic density in the segment fed by the fourth on-ramp (upper plot) and of the queue length at the fourth on-ramp (lower plot) for two situations with MPC-based ramp metering controllers. The solid line represents the reference situation ($\theta_{\text{mpc}} = \theta_{\text{ref}}$). The dotted line represents the situation with the perturbed prediction model parameter set ($\theta_{\text{mpc}} = \theta_{\text{perturb}}$). Note that in the second situation no queue is formed at the on-ramp.

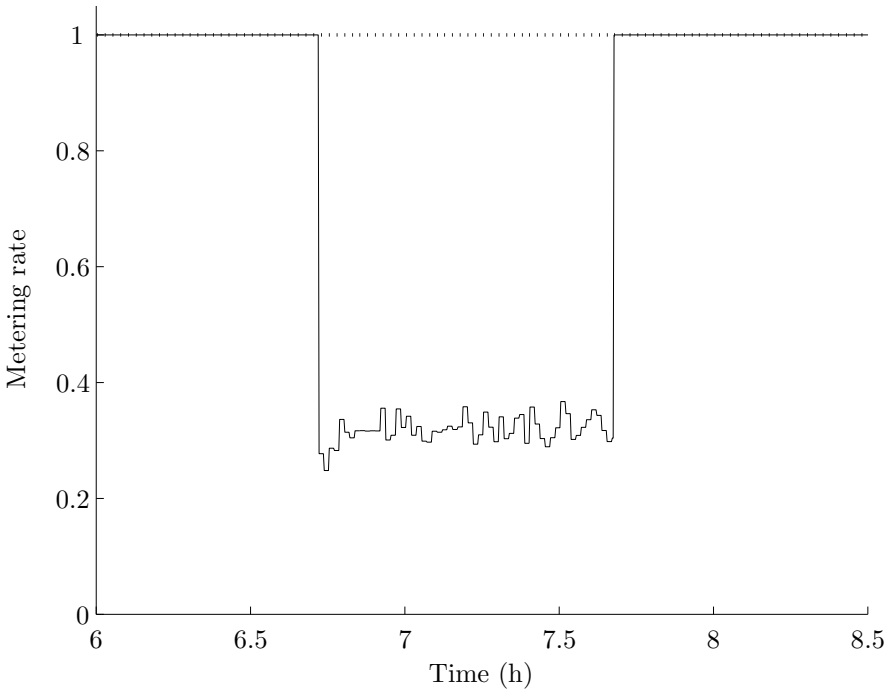


Figure 6.5: Plot of the evolution of the metering rate provided by the MPC-based controller for both experiments. The solid line represents the reference experiment using the exact model parameters ($\theta_{\text{mpc}} = \theta_{\text{ref}}$). The dotted line represents the experiment with the perturbed parameter set ($\theta_{\text{mpc}} = \theta_{\text{perturb}}$). Note that the MPC controller using the perturbed parameter set in the prediction model never becomes active (metering rate remains 1) due to the overestimation of the motorway capacity.

6.1.2 The combined MPC-identification approach to ramp metering

In the previous section it was illustrated that a misfit of the traffic model that is used in the MPC controller to predict the traffic behavior over the prediction horizon can lead to a decreased performance of the MPC controller. This misfit between the behavior of the real traffic system and the model can occur due to external disturbances such as e.g., influences of the weather, incidents, ... In order to deal with this misfit and the corresponding decrease in performance, the model parameters can be updated at regular time intervals using the traffic state measurements that are available from the MPC controller. This leads to a combined MPC-identification approach to the ramp metering problem as shown in Figure 6.6.

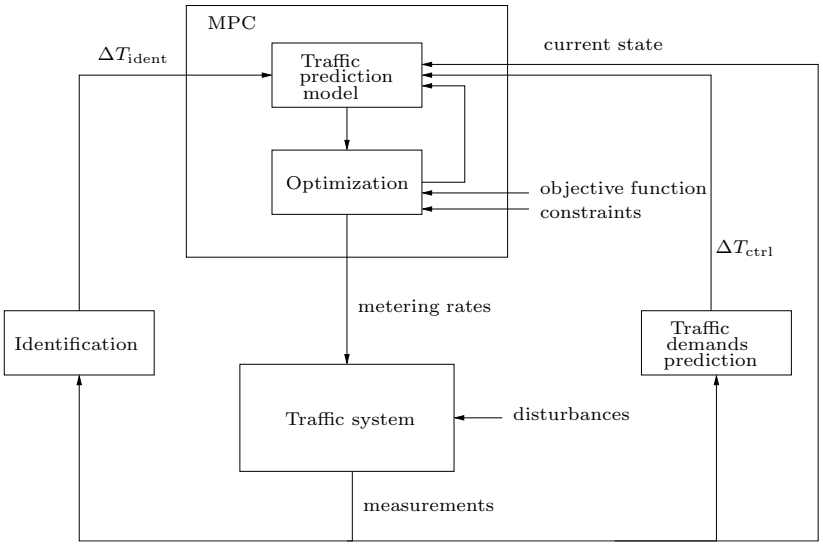


Figure 6.6: Schematic representation of the MPC-identification approach to ramp metering. The right-hand side of the figure represents the MPC-based ramp metering controller that was also presented in Figure 4.8. This MPC-based controller is extended with an identification block that regularly updates the traffic prediction model parameters. The model parameters are updated at a lower frequency than the metering rates (e.g., $\Delta T_{\text{ident}} = 30$ min and $\Delta T_{\text{ctrl}} = 1$ min).

Figure 6.6 shows the MPC controller, which receives measurements of the traffic state on the motorway on a regular basis. Based on these measurements, a prediction of the future traffic demand during the prediction horizon is made. These predicted demands are fed to the traffic prediction model that is used during the optimization of the metering rates. The first metering rate is applied to the system and the whole process starts all over again (receding horizon approach). Figure 6.6 differs from a regular MPC implementation in that an identification part is added to the system. The identification of a traffic model is a computationally intensive task as was discussed in Section 3.4. Since the dynamics of the non-modeled influences can be assumed to be limited, the frequency of the traffic prediction model updates can be lower than the frequency of the updates of the metering rate ($\Delta T_{\text{ident}} > \Delta T_{\text{ctrl}}$). E.g., in the MPC-identification implementation presented in Figure 6.6, the period between two model updates could be chosen to be $\Delta T_{\text{ident}} = 30\,\text{min}$ while the control step is chosen to be $\Delta T_{\text{ctrl}} = 1\,\text{min}$. This provides more time to run the identification process. The computational complexity of the identification process is also reduced by the availability of a good starting point for the optimization. Since a parameter set is available for the model and since the external factors can in general be expected to be small enough not to alter the model parameters too much in a short time period, the model parameters estimated in the previous identification cycle can be used as the starting point for the identification process. In order to limit the computation time required by the parameter identification, the identification process can be implemented as an adaptive procedure where at regular time intervals a fixed number of iterations (e.g., steepest descent steps) of the identification are executed. The traffic data provided to the identification process consist of a collection of traffic state measurements that are collected during the operation of the MPC controller.

6.2 Anticipative ramp metering control

If ramp metering is implemented in a motorway network with alternative routes, traffic can spontaneously re-route due to the response of the drivers to the applied control actions. Although spontaneous re-routing can have a significant influence on the resulting traffic situation in the traffic network and on the performance of the traffic network as a whole, re-routing is usually not automatically included in current motorway traffic control frameworks. In this section, we develop a new method to efficiently calculate and incorporate re-routing effects into the MPC framework. In this way, anticipative model predictive control for ramp metering in motorway networks is realized.

In a motorway network with multiple routes from the origins to the destinations, drivers have to choose their route in the network. Given the traffic demands for each origin-destination pair and the topology of the network, the

traffic needs to be assigned to the routes before a simulation can be run. In this section we present anticipative ramp metering control, which takes the re-routing effects due to ramp metering into account. Note that although the focus in this thesis is on ramp metering, the presented technique can also be applied to other traffic control measures. First, a static equilibrium traffic assignment algorithm that assigns the traffic to the different routes based on the collective behavior of the drivers is presented. Next, a new dynamic traffic assignment method based on the static traffic assignment method is presented. This dynamic traffic assignment method is incorporated in the MPC framework to realize anticipative MPC-based control that takes the spontaneous re-routing of drivers due to control measures into account.

6.2.1 Static equilibrium traffic assignment

When traveling in a motorway network, drivers try to find the route that is optimal for themselves. In fact, it appears as if every driver assigns a cost to every alternative route leading to his destination and chooses the one with the smallest cost. First, we present an expression for the cost assigned by an average driver to a route. Next, we discuss static equilibrium traffic assignment where the traffic demands between origins and destinations are assigned to the available routes taking the behavior of the drivers with respect to the route costs into account.

Two important factors in the cost assigned to a route are [38, 126]: the travel time along the route and the length of the route. Hence, the cost of a route can be computed as follows:

$$\text{route cost} = \alpha_{\text{t}} \,\text{travel time} + \alpha_{\text{l}} \,\text{route length}, \tag{6.1}$$

where the parameters α_{t} and α_{l} are the relative importance of the travel time and the route length in the total route cost respectively. The importance an individual driver assigns to each of these factors varies from driver to driver. The values of α_{t} and α_{l} in (6.1) are the values for an average driver. Since a route consists of consecutive links connecting the origin with the destination, the cost assigned to a route according to (6.1) can be calculated by adding the link costs.

In order to illustrate the implementation of the static equilibrium traffic assignment method, the reasoning in the remainder of this section will be conducted based on the METANET model equations presented in Chapter 3[2]. Note however that this choice is free and that another model can be selected if desired.

In the context of the METANET model, each link is subdivided in segments. Hence, the link cost from (6.1) can be computed as the sum of the segment

[2]Note that since the routing of the traffic through the network is important, the destination oriented version of the METANET model is used.

costs:

$$c_m(q_m(l)) = \alpha_t \sum_{i \in S_m} \frac{l_{m,i}}{v_{m,i}(l)} + \alpha_l \sum_{i \in S_m} l_{m,i}, \tag{6.2}$$

where S_m is the set of segments in the link m. According to (6.2), the link cost depends on the traffic demand $q_m(l)$ through the speed in the sections. Indeed, if more drivers use a link, the densities in the sections of that link increase and the desired average speeds in the corresponding sections decrease, increasing the travel time (cost) of the link. The travel time computed by (6.2) is an instantaneous travel time and can differ from the experienced travel time [145]. However, since static equilibrium assignment is considered, the traffic states in the network are considered invariant in time, and the instantaneous and the experienced travel times are equal. Given the fact that the computation of the experienced travel time is computationally more involved than the computation of the instantaneous travel time, the instantaneous travel time is used in (6.2). In order to be able to compute the link costs for a whole motorway network using (6.2), the average speeds in the segments given a set of link flows need to be computed. Since static equilibrium traffic assignment implies that the traffic flows in the network are invariant in time, it is assumed that the traffic flows in the links are in equilibrium. By consequence, the equilibrium average segment speeds can be computed using the METANET model equations with $\rho_{m,i,j}(k+1) = \rho_{m,i,j}(k)$ and $v_{m,i}(k+1) = v_{m,i}(k)$. In what follows, we present how a static equilibrium traffic assignment can be computed using the expression of the route cost presented in (6.2). In the context of the METANET model, the splitting rates at the network nodes can be determined based on the traffic assignment.

Wardrop stated in 1952 [147] that the traffic in a network distributes over the links in such a way that an equilibrium occurs where no individual driver can lower his travel time by changing routes. In equilibrium all used routes between an origin-destination pair have the same travel cost and non-used routes have a higher travel cost. The resulting equilibrium is called the user optimal equilibrium since it occurs if every driver individually optimizes his route.

There exist several methods to compute the user optimum as defined by Wardrop's principle such as e.g. the Frank-Wolfe algorithm [144], the method of the successive averages [119], feedback strategies [133], iterative strategies [153], predictive feedback strategies [146], ... For a more detailed discussion of the methods to compute the user optimum the interested reader is referred to [126]. Since it is guaranteed to converge to a solution, we discuss the method of the successive averages (MSA) here. The MSA is an iterative static equilibrium traffic assignment method that takes the impact of vehicle flows on the link costs into account through the cost function (6.2). The algorithm uses link flow vectors $\mathbf{q}^i$, which contain the link flows for all the links in the network at iteration i. The algorithm contains the following steps:

initialization: $i = 1$, $\mathbf{q}^i = 0$, $\phi = 1$
repeat
Step 1 calculate costs $\mathbf{c}^i(\mathbf{q}^i)$ according to (6.2)
Step 2 determine $\mathbf{q}^i_{\text{aon}}$ by all-or-nothing assignment
Step 3 $\mathbf{q}^{i+1} = (1-\phi)\mathbf{q}^i + \phi\mathbf{q}^i_{\text{aon}}$
Step 4 $i = i+1$, $\phi = \frac{1}{i}$
until stopping criterion reached

In Step 1, the vector $\mathbf{c}^i(\mathbf{q}^i)$ containing the link costs associated with the traffic assignment in iteration i is calculated. The link costs are computed based on the vector with the link flows $\mathbf{q}^i$ and the equations (3.15), (3.3) and (3.18) of the METANET model as described before. In Step 2, all the traffic between an origin-destination pair is assigned to the cheapest route. This is called an all-or-nothing assignment. During the i-th iteration for every origin-destination pair the cheapest route is searched based on the link cost vector $\mathbf{c}^i(\mathbf{q}^i)$. For smaller networks, the cheapest route can be searched exhaustively but for larger networks a more advanced method like Dijkstra's shortest path algorithm [42, 152] is needed. The flows in the links caused by the all-or-nothing assignments for all origin-destination pairs lead to a link flow vector $\mathbf{q}^i_{\text{aon}}$. In Step 3 a new link flow vector $\mathbf{q}^{i+1}$ is computed as a convex combination of the previous link flow vector and the link flow vector $\mathbf{q}^i_{\text{aon}}$ resulting from the all-or-nothing assignments in Step 2. The meaning of ϕ is the following: In iteration i, the value of ϕ is such that the new link flow vector that is found is the average of all i link flow vectors. Hence the name: the method of successive averages. The stopping criterion for the MSA can be e.g., a maximal number of iterations, a distance criterion between the flows $\mathbf{q^i}$ and $\mathbf{q^{i+1}}$ resulting from two successive iterations of the MSA algorithm becoming lower than a threshold value, or a distance criterion between the costs $\mathbf{c}^i(\mathbf{q}^i)$ and $\mathbf{c}^{i+1}(\mathbf{q}^{i+1})$ resulting from two successive iterations of the MSA algorithm becoming lower than a threshold value.

In the next section we will discuss an anticipative traffic assignment strategy that is based on the principles of the static equilibrium traffic assignment presented in this section.

6.2.2 Anticipative traffic assignment

The previous section dealt with static equilibrium traffic assignment where the traffic demands were invariant in time. In this section anticipative traffic assignment is presented. Anticipative traffic assignment takes the response of the drivers to the changes of the traffic state over time into account. The presented method only uses the experienced traffic state and does not require the traffic demands to be constant. The anticipative traffic assignment is incorporated in the MPC framework for ramp metering in order to allow for the computation of control signals that anticipate the response of the drivers to the control actions.

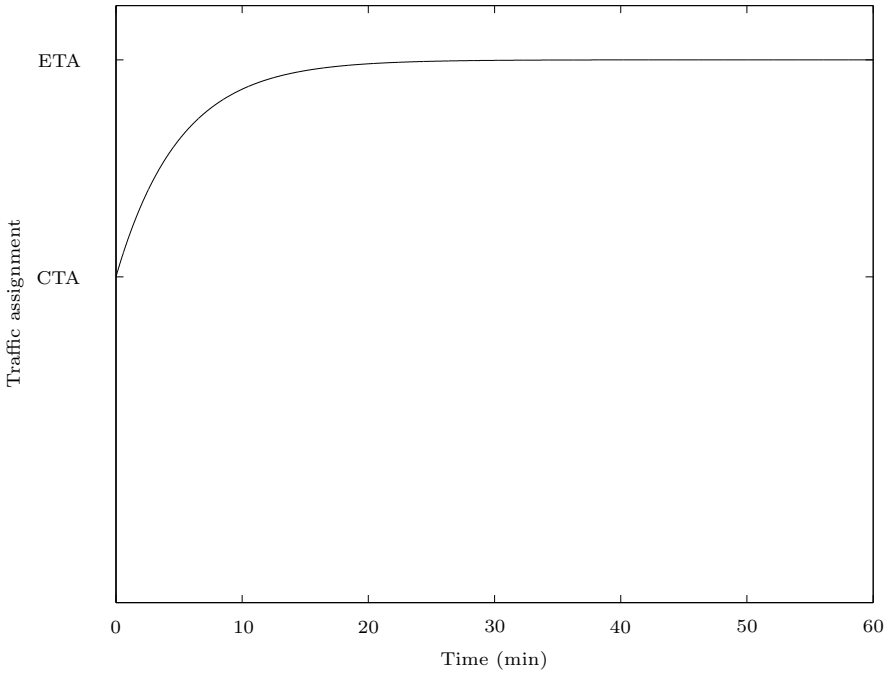


Figure 6.7: Exponential evolution from the current traffic assignment (CTA) to the equilibrium traffic assignment (ETA) for $\tau_{\text{evol}} = 5$ min.

During their trip in the motorway network, drivers experience the traffic state of the network. Based on the information they gather on the global state of the traffic network, drivers will determine their optimal route. The current traffic state assessment, which is used to determine the optimal route, is typically based on information of the traffic state in the near past. This process of gathering traffic state information can be modeled as follows: The traffic state in the near past results from the traffic demands and the metering rates in the near past. The fact that the gathering of traffic state information takes some time is modeled by averaging the traffic demands and the metering rates over a time horizon τ_{info}. By averaging the traffic demands and the metering rates over the time interval $[t - \tau_{\text{info}}, t]$, we obtain the information about the traffic state and past/expected traffic conditions as they are perceived by the drivers and as they are used at the current time to determine their optimal route. The parameter τ_{info} needs to be tuned and is influenced by the network dimensions and topology, the availability of information (e.g., radio bulletins) and so on. E.g., a value of 30 minutes could be chosen for τ_{info}.

Drivers use the gathered traffic state information to optimize their route. Using the static equilibrium assignment method presented in Section 6.2.1, combined

with the average values of the traffic demands and the metering rates, the equilibrium traffic assignment (ETA) associated with the traffic situation experienced by the drivers can be computed. Since the average traffic demand is used, the static equilibrium traffic assignment method yields acceptable results even in the case of time varying traffic demands. Based on the assignment of the flows to the routes, the splitting rates used in the METANET simulation model are computed.

It takes some time before the traffic flows in the motorway network will reorganize according to the ETA. This is due to the fact that not all drivers will decide to use the new route immediately. The transition from the current traffic assignment (CTA) to the ETA is modeled as an exponential evolution of the splitting rates over time as shown in Figure 6.7. The time constant τ_{evol} (63%) needs to be tuned such that the evolution from CTA to ETA occurs in a realistic time frame.

6.2.3 The anticipative MPC-based traffic control strategy

The anticipative MPC-based traffic control strategy combines the MPC-based control strategy and the anticipative traffic assignment technique. A schematic representation of the overall control strategy is presented in Figure 6.8.

The MPC-based control module produces control signals in the form of metering rates, which are applied to the traffic system. The state of the traffic situation is measured (e.g., every minute) and fed back to the MPC module. These measurements are used to update the state and to subsequently start a new optimization over the shifted prediction horizon (recall that MPC uses a moving horizon approach).

During the traffic assignment, a prediction of the ETA is made using information of the traffic situation experienced by the drivers during the interval $[t - \tau_{\text{info}}, t]$, where t is the time at which the ETA is computed. This experienced traffic situation is computed based on the traffic demands and the metering rates as described in Section 6.2.2. The ETA and the CTA are combined to compute a description of the evolution of the splitting rates in time. This evolution is fed to the MPC control module in order to be used for the MPC prediction during the prediction horizon. By supplying the MPC with the evolution of the splitting rates from CTA to ETA, the controller is able to take the re-routing behavior of the drivers into account. Since the dynamics of the re-routing are slower than the dynamics of the traffic system near the on-ramps, it suffices to update the traffic assignment at a slower pace than the traffic states (e.g., every $\Delta T_{\text{anticip}} = 15$ min).

The assumption of the dynamics of traffic re-routing being slower than the dynamics of the traffic behavior on the motorway near the on-ramps is an im-

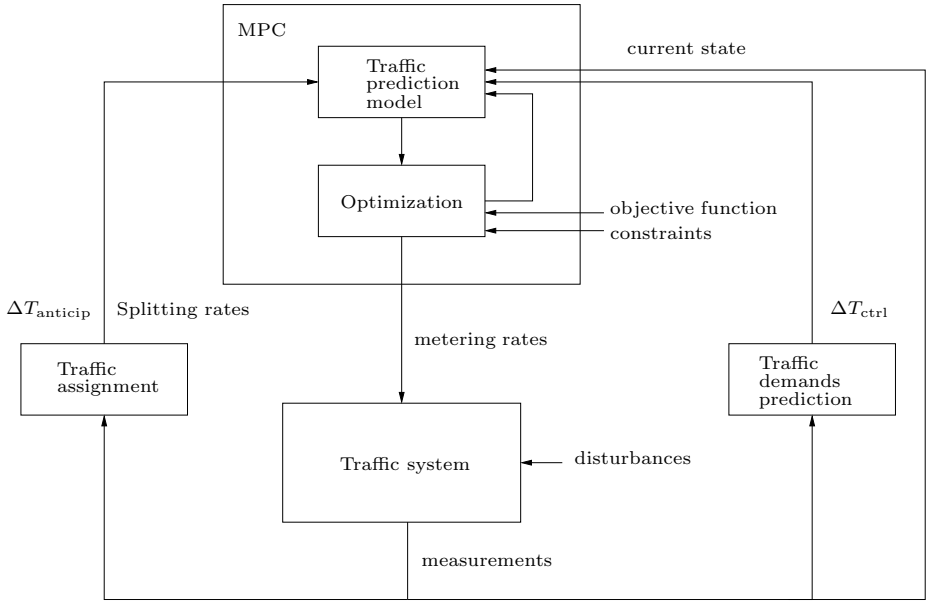


Figure 6.8: Schematic representation of the anticipative MPC-based control strategy for ramp metering. The right-hand side of the figure represents the MPC-based ramp metering controller that was also presented in Figure 4.8. On the left-hand side of the figure, a traffic assignment module is added to provide the traffic prediction model in the MPC controller with estimations of the routing of the traffic in the network (splitting rates). This routing information is obtained using the anticipative traffic assignment from Section 6.2.2. The traffic routing is updated at a slower pace ($\Delta T_{\text{anticip}} = 15$ min) than the metering rates ($\Delta T_{\text{ctrl}} = 1$ min).

portant assumption. The slower the dynamics of the re-routing process are, the larger the parameter $\Delta T_{\text{anticip}}$ can be chosen. This is important since the computation of the ETA is rather computationally intensive. Also from the point of view of stability of the anticipative MPC-based ramp metering controller it is desirable that the dynamics of the re-routing are sufficiently small compared to the dynamics of the traffic operation [132]. The necessary condition for stability of the anticipative MPC-based traffic control methodology as well as the impact of the objective function of the MPC controller on stability are subject to further research.

Remark 6.1 In the METANET traffic simulation software package [133], several dynamic traffic assignment algorithms are incorporated ranging from feedback to iterative strategies. However, since the dynamic traffic assignment is updated at every simulation step, as is also described in [145], this results in a high computational complexity compared to the simulation of the METANET model without dynamic traffic assignment. The anticipative MPC-based control strategy that we presented in this section reduces the computational complexity by updating the dynamic traffic assignment at a pace $\Delta T_{\text{anticip}}$ corresponding to the re-routing dynamics. □

6.2.4 Simulation example

In this section, we illustrate anticipative MPC-based ramp metering control using a simulation example. First, we present the layout of the network and the simulated traffic scenario. Based on a simulation of a reduction of the number of lanes due to e.g., an incident, maintenance works on the motorway, ..., we illustrate the behavior of the anticipative traffic assignment method discussed in Section 6.2.2. First, we consider a situation without ramp metering. Next, we repeat the simulation with MPC-based ramp metering enabled and compare the performance of the traffic network with and without MPC-based ramp metering control.

Traffic network layout

The traffic network we consider in this simulation example is presented in Figure 6.9. The traffic network consists of a motorway with four lanes which bifurcates into two branches of two lanes each. Downstream both branches join in a four lane motorway. Both four lane motorway links are 3 kilometers long. The lower two lane branch is the primary branch. The primary or main branch is shorter than the upper branch and consists of two links. The primary branch is 6 kilometers long and an on-ramp is present in the middle of the branch. The secondary branch is longer than the primary branch and is 8 kilometers long. During simulation, the traffic originating from the mainstream origin distributes over the two branches. The distribution of the traffic over

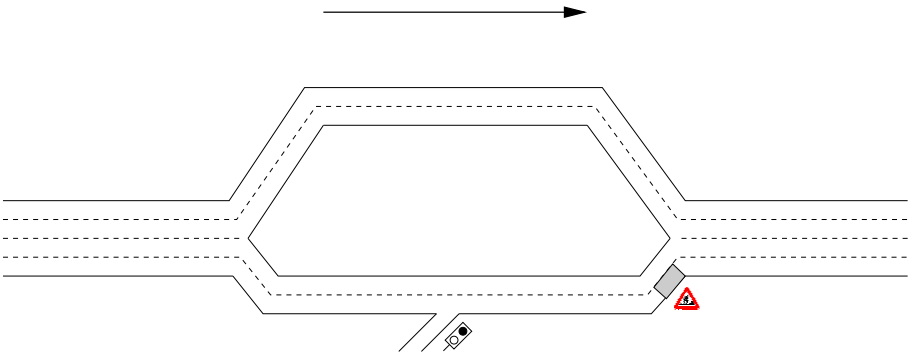


Figure 6.9: Schematic representation of the motorway network studied in the simulation example. The traffic network consists of a four lane motorway link, which bifurcates into two 2 lane branches. The four lane motorway link is 3 kilometers long. In the middle of the 6 kilometer long primary branch an on-ramp, which can be metered, is present. The secondary branch is 8 kilometers long. Both branches join into a four lane motorway link which is 3 kilometers long. The arrow denotes the direction of traffic flow. We can simulate road maintenance works that block one lane of the last segment of the primary branch.

both branches depends on the state of both branches and is modeled using the anticipative traffic assignment discussed in Section 6.2.2.

Traffic model parameters

The destination oriented METANET traffic flow model is used to implement a simulation model for the motorway network presented in Figure 6.9. The links are subdivided in segments of 500 m length each. The splitting rate $\beta(l)$ determines the fraction of the traffic that uses the main branch at simulation step l[3]. According to (3.12), the fraction of the traffic originating from the mainstream origin that uses the secondary branch at simulation step l is given by $1-\beta(l)$. The value of $\beta(l)$ is computed using the anticipative traffic assignment presented in Section 6.2.2.

The METANET model parameters are chosen equal for all segments: $v_{\text{free},m}$ = 105 km/h, $\rho_{\text{crit},m}$ = 34 veh/km/lane, τ_m = 18 s, ν_m = 60 km^2/h, κ_m = 40 veh/lane/km and $a_m = 1.867$. The merging term parameter δ_{m_o} from (3.5) is chosen equal to 0.0122. The parameters ρ_{jam,m_o} and $Q_{\text{cap},o}$, which are used

[3]Since there is only one node with more than one leaving link in the motorway network shown in Figure 6.9, we denote the splitting rate as $\beta(l)$, omitting the node index m that was used in the general description of the METANET model in Chapter 3. Accordingly the link exponent m is also omitted since there are only two links leaving the node.

in (3.9) or in (3.10) to determine the service rate of the on-ramp are chosen equal to 180 veh/km/lane and 1000 veh/h respectively. The minimal metering rate of the on-ramp $r_{\min}$ is 0.1.

Traffic scenario

We simulate a traffic scenario on the motorway network with road maintenance works on the primary branch as shown in Figure 6.9. The maintenance works result in a reduction of the number of lanes from 2 to 1 in the last segment (i.e., the last 500 m) of the primary branch. The maintenance works start at 5.30 a.m. and persist during the remainder of the simulation. We will denote the start of the maintenance works in the figures with a vertical dashed line as illustrated in Figure 6.10.

The traffic demands at the on-ramp and at the mainstream origin are shown in Figure 6.10. The traffic demand on the mainstream is considered constant and equal to 4500 veh/h in this simulation for illustrative purposes. The traffic demand on the on-ramp is equal to 200 veh/h with a peak traffic demand of 500 veh/h around 6 a.m.

Anticipative traffic assignment parameters

The parameter τ_{info} of the anticipative traffic assignment determines the size of the horizon over which the drivers gather information to assess the global state of the network. The state of the network results from the traffic demands and the metering rates from the past. Based on this information the equilibrium traffic assignment is computed using the method described in Section 6.2.1. The larger the horizon τ_{info}, the slower the response of the route choice behavior of the drivers on varying traffic demands and metering rates will be. We choose $\tau_{\text{info}} = 30\,\text{min}$ in this simulation example.

The parameter τ_{evol} models the swiftness of the response of the drivers to a difference between the current traffic assignment and the equilibrium traffic assignment. Since the incident or the maintenance works are located near the downstream end of the 6 kilometer main two lane branch, it takes some time for the disturbance caused by this lane reduction to propagate upstream and reach the drivers that still need to make a route choice. The time it takes for the drivers to adapt to the ETA depends e.g. on the situation and the familiarity of the drivers with the motorway network. We choose the value of the parameter $\tau_{\text{evol}} = 45\,\text{min}$.

The rate at which the traffic assignment needs to be updated depends on the dynamics of the re-routing in the network. The re-routing dynamics are typically much slower than the dynamics of the traffic system near the on-ramps and depend on the topology of the network. A trade-off can be made between

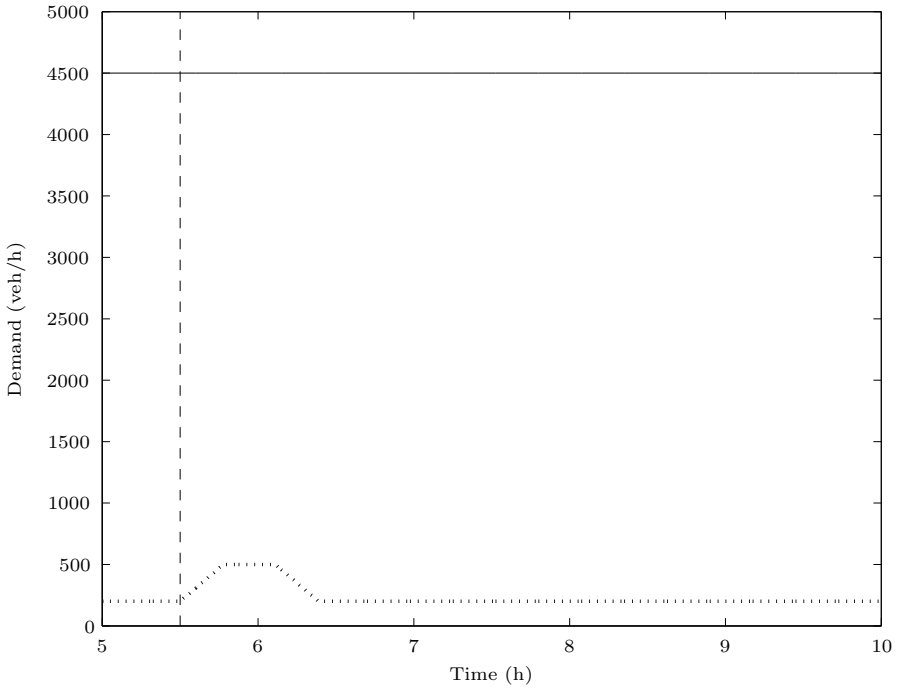


Figure 6.10: Plot of the traffic demand at the mainstream origin (solid line) and at the on-ramp (dotted line) for the motorway network of Figure 6.9.

the accuracy of the traffic assignment and the computational complexity of the anticipative traffic assignment by tuning the time $\Delta T_{\text{anticip}}$ between two ETA updates. In this simulation example we choose $\Delta T_{\text{anticip}} = 15\,\text{min}$.

The cost we assign to a link is given by (6.2) with α_{t} =1 and $\alpha_{\text{l}} = 0$, i.e., we choose the average instantaneous travel time through a link to be the cost of that link.

No ramp metering control

Before discussing anticipative MPC-based ramp metering for the traffic network from Figure 6.9, we discuss a simulation without ramp metering but with anticipative traffic assignment. The simulation results are summarized in Figure 6.11.

At the start of the simulation, the traffic network is in steady state and the instantaneous travel times along both alternative routes are equal. The demands

at both the mainstream and the on-ramp are constant and there is no incident on the main branch. The split rate $\beta(l)$ is equal to 0.86 (see middle plot in Figure 6.11). The main branch is the shortest route but due to the traffic flow through this branch, the density in this branch increases and thus the speed decreases. This decrease in speed results in an increased travel time. As a result, a fraction of the drivers chooses the longer secondary branch since it has a lower density with corresponding higher speed resulting in a travel time that is equal to the travel time of the main branch. The vehicles entering the main branch through the on-ramp also contribute to the increased travel time of the main branch. In this way, a change in the on-ramp traffic flow results in a reorganization of the mainstream traffic flows.

At 5.30 a.m., the traffic demand on the on-ramp increases while at the same time the maintenance works start at the downstream end of the primary branch (see Figure 6.9). Due to the increasing traffic demand on the on-ramp (see Figure 6.10) the traffic density in the first segment downstream of the on-ramp starts to increase immediately, which results in a decrease of the average speed in the segment (see Figure 6.11).

Due to the maintenance works, the number of lanes of the primary branch is reduced from two to one. Since the traffic flow carried by the primary branch is too high for one single lane, congestion sets in. In the upper plot of Figure 6.11 the traffic density of the first segment after the on-ramp is shown. Since the reduction of the number of lanes is located downstream of this segment, it takes some time for the congestion to propagate to this segment. This is observed as the delay between the start of the construction works at 5.30 a.m. and the peak in the traffic density in the segment. The traffic density in the segment exceeds the critical density of the segment and congestion sets in in the segment.

The average speed in a segment decreases with increasing traffic density as was shown in Figure 3.2. Hence, the increased traffic density in the primary branch results in a decrease of the average speed in the branch (see Figure 6.11) and thus in an increased travel time. This phenomenon can be observed in the lower plot of Figure 6.11, which presents the travel time from the mainstream origin to the destination for both alternative routes. The travel time of the route including the primary branch (solid line) increases immediately with the start of the maintenance works in the primary branch and with the increase in traffic demand at the on-ramp.

As a result of the increase of the travel time on the primary route, the secondary route becomes the fastest route. Hence, the drivers will be inclined to start using the secondary route. This can be observed in the third plot of Figure 6.11 where the evolution of the split rate $\beta(l)$ over time is plotted. The split rate starts to decrease from the moment the travel time on the primary route becomes larger than the travel time on the secondary route. In the lower plot

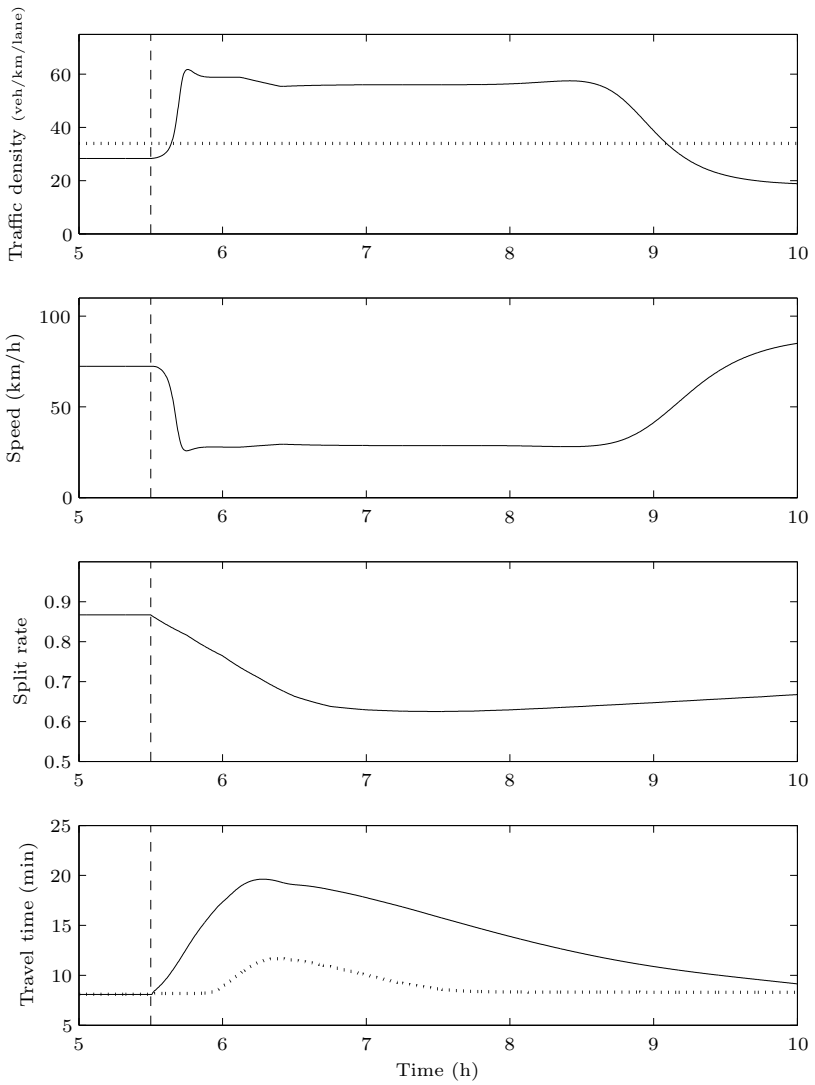


Figure 6.11: Simulation results of anticipative traffic assignment without control for the network presented in Figure 6.9. Plot of the evolution of the traffic density (upper plot) and the average speed (second plot) in the first segment downstream of the on-ramp in the traffic network. The dotted line in the upper plot represents the critical traffic density ρ_{crit} of the segment. The third plot represents the evolution of the splitting rate $\beta(l)$ over time. In the lower plot, the instantaneous travel time is shown for the primary (solid line) and the secondary (dotted line) route from the mainstream origin to the destination in Figure 6.9. The vertical dashed line represents the start of the construction works on the primary branch.

of Figure 6.11 we observe an increase in the travel time on the secondary route, which results from the increased traffic volume on the secondary route but also from the spill-back of congestion of the primary branch into the upstream four lane motorway link. The congestion in the four lane motorway link starts resolving as soon as a sufficiently large number of drivers starts using the longer secondary route. As the four lane mainstream link is common to both the primary and the secondary route, this can be observed as the decrease in travel time for both routes around 7 a.m.

Once the congestion on the mainstream link is resolved, the congestion in the primary branch starts to resolve since more drivers are now using the secondary route. Since more and more segments of the primary branch are becoming uncongested, the travel time of the primary route decreases. Around 9 a.m. the traffic density in the segment right after the on-ramp drops below the critical density and the congestion in this segment is resolved.

Eventually, the system reaches a new equilibrium state with equal travel times for both routes. In this new equilibrium state there still is congestion in the segments directly upstream of the maintenance works. However, in the upstream segments of the primary branch, the traffic density during the new equilibrium is lower than the traffic density in the equilibrium state at the beginning of the simulation due to the lower traffic volume using the primary branch. As an illustration of this we refer to the traffic density in the segment behind the on-ramp at the start and at the end of the simulation period (see Figure 6.11).

The split rate associated with this new equilibrium state is lower than the split rate at the start of the simulation since more drivers take the secondary route to avoid the congestion near the maintenance works on the primary route. The travel time associated with this new equilibrium state is higher than the travel time associated with the equilibrium at the start of the simulation experiment.

Similarly as in Chapter 5, we define the total time spent by all the vehicles in the traffic network as the performance measure P_{TTS} of the traffic operation. The performance P_{TTS} can be computed using (5.1). For the scenario presented above the performance P_{TTS} of the traffic state in the network was found to be 4886 veh/h. All this time is spent in the traffic network by the vehicles since there is no queue present during the simulation.

Anticipative MPC-based ramp metering control

After the extensive discussion of the simulation with anticipative traffic assignment but without ramp metering control, we now present the simulation results of anticipative MPC-based ramp metering control for the motorway network presented in Figure 6.9 and the traffic scenario discussed above.

The MPC parameters N_{p} and N_{c} are chosen as described in Section 4.4.3. We

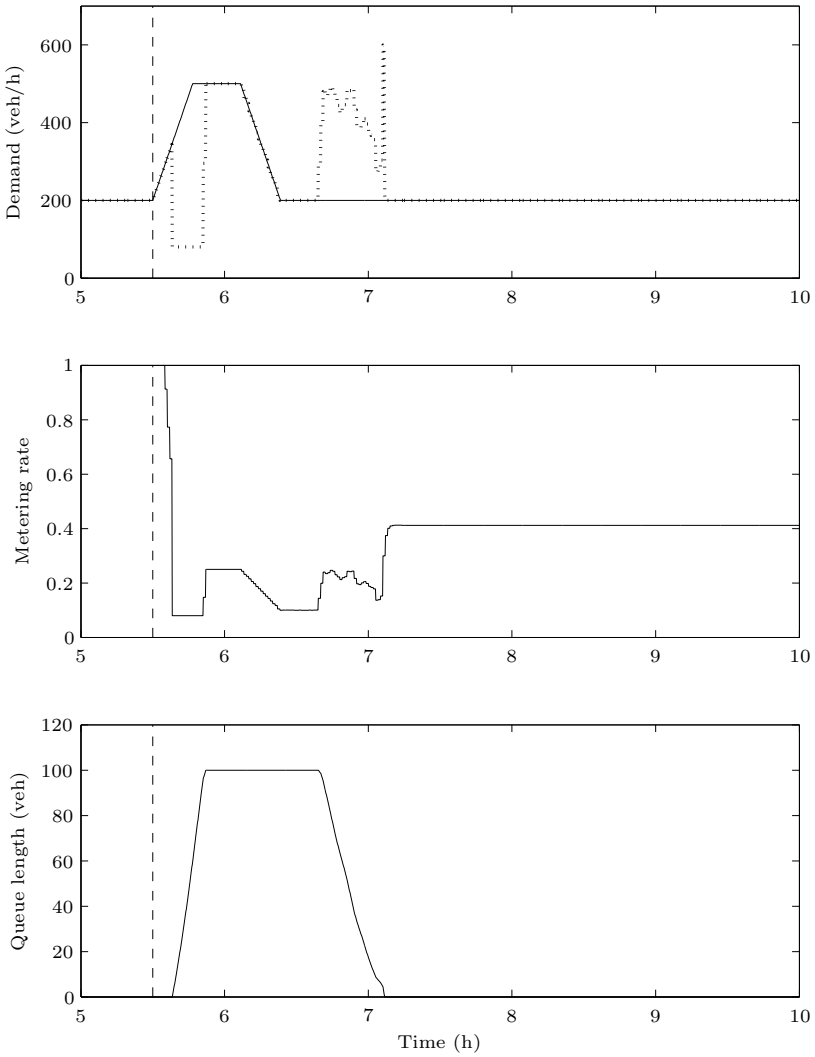


Figure 6.12: Simulation results of anticipative MPC-based ramp metering applied to the network presented in Figure 6.9. Plot of the traffic demand (solid line) and the traffic that is allowed to enter the primary motorway branch (dotted line) of the traffic network. The middle plot presents the evolution of the metering rate of the on-ramp over time. The evolution of the queue length at the on-ramp over time is shown in the lower plot. The vertical dashed line represents the start of the construction works on the primary branch.

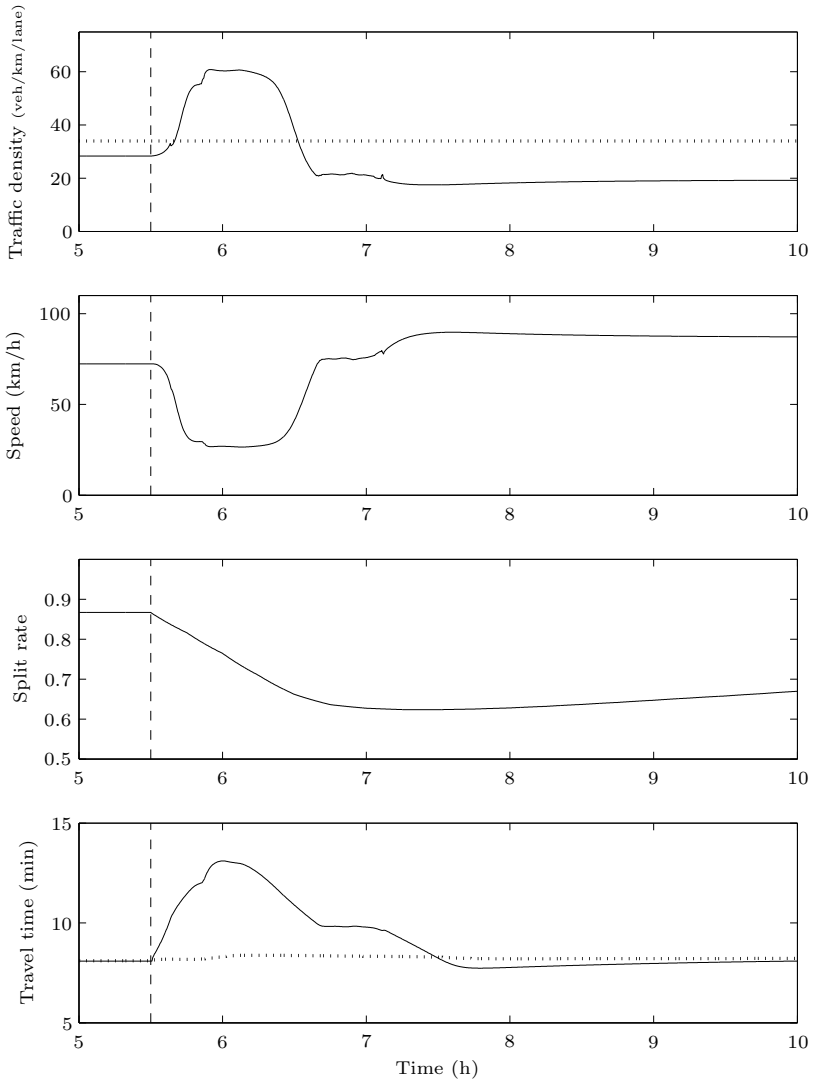


Figure 6.13: Simulation results of anticipative MPC-based ramp metering applied to the network presented in Figure 6.9. Plot of the evolution of the traffic density (upper plot) and the average speed (second plot) in the first segment downstream of the on-ramp. The dotted line in the upper plot represents the critical traffic density ρ_{crit} of the segment. The third plot represents the evolution of the splitting rate $\beta(l)$ over time. In the lower plot the instantaneous travel time is shown for the primary (solid line) and the secondary (dotted line) route from the mainstream origin to the destination in Figure 6.9. The vertical dashed line represents the start of the construction works on the primary branch.

choose the prediction horizon N_p to be 15 minutes and the control horizon N_c is chosen to be 8 minutes. We choose (4.3) as the MPC objective function with $\alpha_{queue} = 1$ and $\alpha_{var} = 30$, which yields good results as will become clear in the discussion of the simulation results below.

The simulation example with anticipative MPC-based ramp metering control starts from the same equilibrium state as the simulation without ramp metering control discussed in the previous section. At 5.30 a.m., the traffic demand at the on-ramp starts to increase as presented in Figure 6.10 and the maintenance works start, which results in a reduction of the number of lanes from 2 to 1 near the downstream end of the primary branch.

As a result of the increased traffic demand at the on-ramp and the reduced number of lanes in the last segment of the primary branch, the traffic density in the first segment downstream of the on-ramp starts to increase immediately as can be observed in the upper plot of Figure 6.13. In order to avoid the traffic density in the primary branch from becoming larger than the critical density ρ_{crit}, ramp metering becomes active as can be seen in the middle plot of Figure 6.12 where the evolution of the metering rate at the on-ramp over time is presented. In the upper plot of Figure 6.12 we see the traffic demand at the on-ramp plotted in a solid line and the traffic flow that is allowed to enter the motorway according to (3.10) plotted as a dotted line. In the lower plot, the evolution of the queue length at the on-ramp over time is shown.

Shortly after the activation of the ramp metering, the metering rate drops to its minimal value ($r_{min} = 0.1$) and after some time the queue length becomes equal to the maximal number of 100 vehicles that is allowed at the on-ramp. The MPC controller is able to take the constraint on the queue length into account by increasing the metering rate. The metering rate must be large enough to prevent the queue from growing and the metering rate must be as small as possible to keep the traffic density on the motorway below the critical density ρ_{crit}. This trade-off results in a metering rate which is such that the number of vehicles that is allowed to enter the motorway is equal to the traffic demand (see Figure 6.12).

In the upper plot of Figure 6.13 we observe that despite ramp metering the traffic density in the first segment downstream of the on-ramp grows larger than the critical density ρ_{crit}. This is due to the congestion from the maintenance works spilling back in the upstream direction. However, if we compare the traffic density in the control case (Figure 6.13) with the traffic density in the no control case (Figure 6.11), we observe that the traffic density in the ramp metering control case remains lower than in the no control case. This results in a lower instantaneous travel time for the primary route in the controlled case.

Minutes before 6 a.m., the MPC-based ramp metering controller is forced by the constraint on the queue length to allow more traffic to enter the motorway

in order to avoid the queue becoming too long. This results in an immediate increase of the traffic density in the segment fed by the on-ramp as can be seen in Figure 6.13. The increased traffic density leads to a decrease of the average speed in the segment and thus the travel time of the primary route increases due to the increased metering rate (see Figure 6.13).

The third plot in Figure 6.13 shows the evolution of the split rate over time. Due to the larger travel time on the primary branch, a fraction of the traffic shifts from the primary to the secondary route. The impact of the shift of traffic from the primary to the secondary motorway branch on the instantaneous travel time of the secondary route is rather small.

As the traffic demand at the on-ramp decreases again, the traffic density on the motorway starts to decrease as well. At a certain point, the traffic density on the primary motorway branch is low enough for the controller to start dissolving the queue. After 7 a.m., all the traffic that wants to enter the primary motorway branch through the on-ramp is allowed to do so, despite the metering rate that differs from 1. Indeed, we observe in Figure 6.12 that the queue length is 0 and that the traffic flow entering the motorway through the on-ramp is equal to the traffic demand at the on-ramp after 7 a.m. The metering rate used in (3.10) differs from 1 but does not restrict the on-ramp traffic.

As was the case in the no control case, the traffic evolves to a new equilibrium state with equal travel times for both alternative routes. In this new equilibrium state, there is some congestion in the segments directly upstream of the maintenance works. Due to the delays and the increased travel time caused by this congestion in the primary branch, there are more drivers using the longer secondary route.

We computed the performance measure P_{TTS} for the simulation of the traffic scenario in combination with anticipative MPC-based ramp metering control. The total time spent by all the vehicles in the network was 3781 veh/h. A part of this time is spent by vehicles in the waiting queue at the on-ramp.

Simulation summary

In this section, we discussed the simulation of the traffic network from Figure 6.9 for a scenario with and without ramp metering traffic control. Since there are two alternative routes from the mainstream origin to the destination, re-routing will occur if there occurs a difference in the costs of both alternative routes. In this example we considered the instantaneous travel time through a link as the cost associated with that link.

In the simulated traffic scenario, maintenance works were started on the primary branch of the network, thus reducing the number of lanes from 2 to 1 (see Figure 6.9). We also simulated a peak in the traffic demand at the on-ramp as

shown in Figure 6.10. Both the maintenance works and the peak in the traffic demand resulted in a shift of some of the traffic from the primary route to the longer secondary route.

In the no control case we observed congestion spilling all the way back from the location of the maintenance works to the upstream four lane link. As can be observed from Figure 6.11, the traffic density in the first segment downstream of the on-ramp remains higher than the critical density ρ_{crit} for a long time.

In the control case, we observe that the MPC-based ramp metering controller starts metering immediately to avoid congestion in the primary branch. However, the congestion that sets in at the maintenance works spills back past the first segment after the on-ramp in the primary branch, which becomes congested despite the ramp metering. Moreover, the MPC-based controller needs to increase the metering rate due to the constraint on the queue length. However, by comparing Figures 6.11 and 6.13, we see that the duration of the congestion in the segment fed by the on-ramp is much shorter in the control case (approximately 1 hour) than in the no control case (approximately 3.5 hours).

The total time spent was computed for both the no control and the anticipative MPC-based ramp metering control case. In the no control case, the drivers try to limit their travel time by choosing the fastest route. This resulted in P_{TTS} = 4886 veh h. In the control case, the MPC-based ramp metering controller takes the re-routing behavior of the drivers into account and is as such able to further reduce the total time spent in the network to $P_{\text{TTS}} = 3781$ veh h. This is a gain of approximately 23% despite the presence of a queue at the on-ramp.

6.2.5 A combined anticipative MPC-identification approach to ramp metering

The MPC-identification approach to ramp metering discussed in Section 6.1.2 can be combined with the anticipative MPC-based ramp metering strategy from Section 6.2.3. This combined framework is able to deal with slowly varying non-modeled disturbances as well as with the re-routing behavior of drivers in response to control actions. A schematic representation of the combined anticipative MPC-identification approach is given in Figure 6.14.

6.3 Conclusions

This chapter presented two new extensions to the MPC-based ramp metering framework: the combined MPC-identification approach and the anticipative MPC-based traffic control strategy.

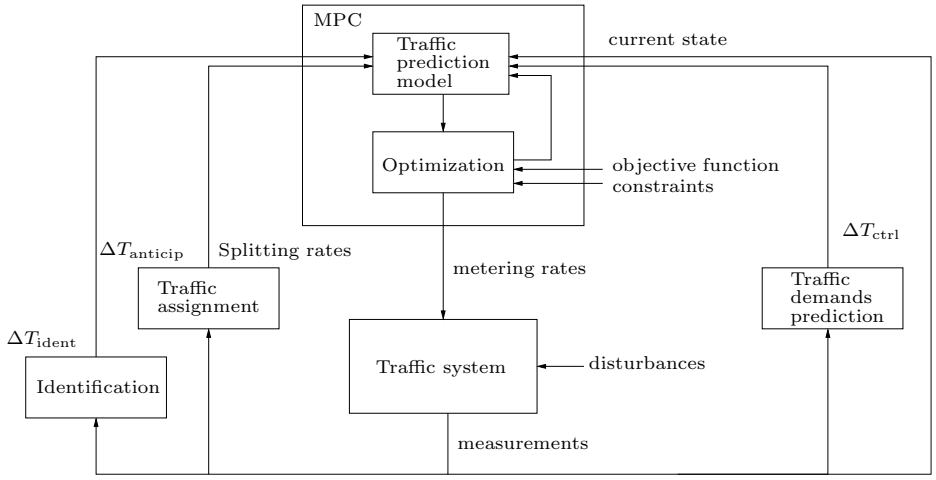


Figure 6.14: Schematic representation of the combination of the MPC-identification approach to ramp metering presented in Figure 6.6 and the anticipative MPC-based control strategy for ramp metering presented in Figure 6.8.

The MPC-identification approach to ramp metering we discussed in this chapter performs a re-identification of the MPC prediction model on a regular basis as presented in Figure 6.6. In this way, the misfit between the prediction model and the real-life situation, which can reduce the performance of the MPC controller, is reduced. In a simulation example the impact on the controller performance of a misfit of the MPC prediction model was illustrated. A misfit of the prediction model can result from unpredictable external factors that cannot be modeled. Since the dynamics of the unpredictable external factors, such as e.g., the weather, can be assumed to be limited, the update frequency of the prediction model parameters can be relatively low. The prediction model parameters from the previous iteration constitute in general good starting values for the identification process. The low pace of the model re-identification and the availability of good starting values substantially reduce the computational complexity of the model identification.

Since drivers tend to use the cheapest route available, a control action can cause a re-routing of traffic. The anticipative MPC-based ramp metering control strategy that we developed in this chapter, takes into account that due to the control actions in the motorway network, the drivers may choose an alternative route. The anticipative MPC-based ramp metering strategy optimizes the metering rates in such a way that the desired (lowest cost) traffic state is obtained despite the re-routing behavior of the drivers. It was assumed that the re-routing takes place at a slower time-scale since drivers do not re-route

instantaneously. The re-routing of the drivers was modeled as an exponential transition from the current traffic assignment (CTA) to the equilibrium traffic assignment (ETA). The ETA is determined based on the information that is available to the drivers i.e., the traffic situation in the near past. In case there is some kind of information provided to the drivers, this can be modeled by shortening the transition time from CTA to ETA.

We illustrated this new anticipative MPC-based ramp metering control strategy with a simulation example. In this simulation example, re-routing due to changing traffic demands and due to a sudden reduction of the number of lanes of a motorway branch (e.g. due to an incident, maintenance works, ...) was considered. A simulation with and without anticipative MPC-based ramp metering was performed and it was found that the anticipative MPC-based ramp metering controller was able to reduce the total time spent by the vehicles in the network by 23 %.

This chapter concluded with the suggestion to combine the two presented methods: the combined MPC-identification approach and the anticipative MPC-based approach to ramp metering. This is a topic for further research as we describe in more detail in Chapter 7. Other topics for further research resulting from this chapter include the investigation of the stability of the anticipative MPC-based ramp metering controller and a more elaborate investigation of the MPC-identification approach to ramp metering regarding performance and computational complexity.

The main contributions of this chapter to the state of the art are the development of the combined MPC-identification approach to ramp metering and the development of anticipative MPC-based ramp metering control. In the scope of the anticipative MPC-based ramp metering control framework, we developed an anticipative traffic assignment strategy, which models the re-routing behavior of the drivers based on their perception of the traffic state of the motorway network and the traffic demands. We also suggested the combination of both methods presented above as a topic for further research.

Chapter 7

Conclusions and Further Research

7.1 Conclusions

In this thesis we have studied advanced traffic management on motorways. We have applied model predictive control (MPC) concepts known from process industry to ramp metering control for motorway systems.

An important advantage of the application of MPC-based control in advanced traffic management systems is that the desired controller behavior is defined as an objective function. This allows the road administrator to define the desired mode of traffic operation as an objective function that is directly provided to the MPC-based controller. Besides the convenient definition of the controller behavior, MPC-based control is able to easily take constraints into account. These constraints can result from government policies but also from practical considerations. In this thesis, we have illustrated the inclusion of constraints in the control problem by applying a maximal queue length constraint to the queue lengths at the on-ramps. The maximal queue length constraint avoids spill-back of the queue into the urban traffic network where it could disrupt traffic, and it limits the waiting time for the vehicles in the queue. In an MPC framework, we can extend the control of one single traffic control measure to the coordinated control of several traffic control measures. An MPC-based controller is a non-local controller, which is able to take the traffic state in a large area into account. In this way, optimally controlling traffic on one location at the expense of creating congestion elsewhere in the traffic network is avoided. In the simulation example we have presented in Chapter 5, we have applied coordinated ramp metering to the E17 motorway case study. Due to

the receding horizon approach, MPC-based control is robust against measurement errors and it can adapt to slow changes in the system conditions and the environment.

We have chosen ramp metering as the motorway traffic control measure in this thesis since it has the potential to improve the traffic operation on the motorway. The implementation of ramp metering can be achieved in the short term and at minimal costs. The acceptance of ramp metering traffic control by the public is not hindered by the need for the drivers to invest in in-vehicle technology. From the control point of view it is important to note that ramp metering can easily be enforced, e.g., by a red-light running camera. Since the control can be enforced, we can assume that a control signal that is applied to the system is also followed up exactly. This simplifies the development of the controller.

Let us now briefly summarize the topics treated in this thesis. In Chapter 2, we have presented a general overview of three important components in an MPC-based advanced traffic management system: traffic models, sensing, and control measures. The general overview of the traffic model classification criteria allows for a classification of the large number of existing traffic models based on their most important properties. This classification has allowed us to choose the appropriate model for the case study presented in Chapter 5. The traffic sensors provide the advanced traffic management system with measurements based on which the control signals for the control measures are computed. In Chapter 2, we have given an overview of both the most popular traffic sensor technologies today and of some of the traffic control measures that can be applied in the MPC framework. We have studied the data provided by the traffic sensors, which has yielded a computationally inexpensive estimation scheme for missing data based on the concept of a reference day.

Based on the traffic model classification from Chapter 2, we have chosen the deterministic, macroscopic METANET traffic flow model as the prediction model in the MPC-based ramp metering controller. We have presented the METANET traffic flow model in Chapter 3 in both its non-destination oriented and in its destination oriented form. The non-destination oriented traffic flow model does not explicitly take traffic routing through the network into account. The non-destination oriented METANET traffic flow model has been implemented to simulate the E17 motorway case study in Chapter 5. We have chosen to implement the non-destination oriented METANET model for this case study since the case study consists of a single traffic corridor and thus it follows that we do not need to consider routing effects. We have also discussed the destination oriented METANET model in Chapter 3 since we needed a traffic model that takes the traffic routing information explicitly into account for the discussion of traffic re-routing effects in Chapter 6. In order to conclude Chapter 3, we have briefly discussed the advantages and the disadvantages of the METANET traffic flow model, and the identification of the model param-

eters.

In Chapter 4 we have discussed ramp metering control in general and we have focused in more detail on MPC-based ramp metering control. In order to illustrate the ramp metering concept, we have presented the fundamental diagram, which shows that beyond a critical density, traffic operates in a suboptimal regime (i.e., the flow is lower than the capacity flow). The ramp metering concept consists of limiting the motorway traffic density to a value lower than or equal to the critical density in order to realize an optimal traffic flow on the motorway. Note that during ramp metering, a queue can form at the on-ramp. We have discussed two ramp metering algorithms in detail in Chapter 4: ALINEA-based and MPC-based control.

- The ALINEA-based ramp metering controller is a PID-like controller that tries to achieve a predetermined setpoint. This setpoint is typically a traffic density on the motorway. If the traffic density on the motorway becomes too high, the ramp metering set-up limits the on-ramp flow. Note that the ALINEA-based controller is a local controller that takes only the traffic density on the motorway near the on-ramp into account. Since the storage capacity of the queue at the on-ramps is often limited, a constraint needs to be imposed on the queue length. The ALINEA-based controller is not able to take a hard constraint on the queue length into account. In order to take a hard constraint into account, a mechanism that overrides the ALINEA-based controller needs to be implemented. The threshold on the queue length at which the ALINEA controller is overridden needs to be set conservatively in order to avoid a violation of the hard constraint.

- We have presented MPC-based ramp metering control in detail in Chapter 4. First, we have presented the general MPC concept. In contrast to an ALINEA-based controller, MPC-based ramp metering control does not try to achieve a preset setpoint but it optimizes an objective function. This objective function contains the desired traffic control policy. In order to find the desired control signals, the MPC-based controller minimizes the objective function over a prediction horizon. We have presented an objective function that consists of the total time spent by all the vehicles in the motorway and in the on-ramp queues plus a term to penalize variations in the control signal. The last term allows us to obtain a smooth control signal. The prediction model in the MPC controller is used to predict the traffic behavior during the prediction horizon. The MPC-based controller does not only take the state of the motorway near the on-ramp into account, but it optimizes the objective function over the whole traffic network implemented in the prediction model. The constraints are directly taken into account by providing them to the optimization algorithm that minimizes the objective function over the prediction horizon.

We want to emphasize that the MPC-based approach to traffic control is a modular approach. Based on a trade-off between model accuracy and computational complexity, we have chosen the METANET traffic flow model as the prediction model for the MPC-based controller developed in Chapter 5. However, another model can also be chosen. We can also replace the objective function by another objective function in case we would like to implement a new traffic policy. MPC-based control is not restrained to ramp metering control. We can use an MPC-based controller to control the traffic control measures presented in Chapter 2 and we can easily extend the MPC-based control from one single on-ramp to the coordinated control of several on-ramps as we have done in the case study in Chapter 5. The MPC control concept can even be used to implement the integrated control of a combination of traffic control measures.

We have put both the ALINEA and the MPC-based ramp metering control algorithms to the test in a case study in Chapter 5. We have modeled the E17 motorway Ghent-Antwerp using the non-destination oriented METANET traffic flow model. Since we have observed from measurements that congestion occurs near the fourth and the fifth on-ramps in the case study area, ramp metering has been simulated for these on-ramps. We have defined the total time spent by all vehicles is the traffic network, consisting of the motorway and the on-ramps, during a simulated morning rush hour ranging from 5 a.m. to 10 a.m., as a performance criterion of the traffic operation on the motorway. We have investigated three major cases: the no-control case, ALINEA-based ramp metering control and MPC-based ramp metering control. In order to obtain a reference value of the total time spent performance criterion, we have simulated the morning rush hour without control. We have also simulated the implementation of the ALINEA ramp metering algorithm, including a constraint on the queue length, for the fourth, for the fifth and for both the fourth and the fifth on-ramps in order to be able to compare the MPC-based ramp metering controller we have presented in this thesis with a widely adopted technology. The same ramp metering simulations have been conducted with an MPC-based controller. We have added one extra experiment for the MPC-based controller, the coordination of ramp metering on the fourth and the fifth on-ramps.

In Chapter 6 we have presented two extensions of the MPC-framework: traffic model re-fitting and anticipative ramp metering control. The properties of a traffic system are always subject to external, non-modeled disturbances such as e.g., weather influences, incidents, ... These external influences result in a misfit between the real traffic system and the MPC prediction model. We have illustrated the impact that a misfit of the prediction model can have on the performance of the resulting MPC-based controller with a simulation example in Chapter 6. Based on this experiment we have concluded that it is useful to regularly re-fit the traffic prediction model of the MPC-based controller.

Since the MPC-based controller needs to run in a real-time environment, the computational complexity of the model re-fitting needs to be as low as possible. The computational complexity of the identification process is reduced by the fact that not all parameters need to be re-fitted due to the limited sensitivity of some parameters and due to the fact that the re-fitting of the model need not be performed at every controller step. Moreover, we can assume to have good starting values for the optimization of the model parameters available, namely the model parameters of the previous iteration.

In Chapter 6 we have also discussed the concept of anticipative ramp metering control. Wardrop's principle states that the traffic in a network organizes in such a way that the cost assigned by the drivers to all routes between an origin and a destination is equal for all used routes. Using the equilibrium principle, we have assigned traffic between origins and destinations to alternative routes in the traffic network. However, if we apply a control measure to the traffic system, we alter the cost of the routes and the drivers will re-route in order to obtain the equilibrium assignment. We have presented an anticipative MPC-based ramp metering control strategy that takes this re-routing of the drivers into account when computing the metering rates for the prediction horizon. In this way, we avoid the implementation of a metering rate which has adverse effects on the motorway performance due to re-routing effects. The anticipative control strategy that we have presented in Chapter 6 was obtained by computing the static equilibrium assignment and by assuming that the traffic gradually evolves to this static equilibrium assignment. The adoption of the static equilibrium assignment by the drivers is not instantaneous since it takes some time for the drivers to learn about the changed situation and to effectively re-route. The static equilibrium assignment is updated at every controller step in order to track changes in the traffic situation and in the traffic demand. In order to assess the performance increase due to the anticipative MPC-based ramp metering control, we implemented a simple case study with alternative routes in Chapter 6. We used the destination oriented METANET model from Chapter 3 since we wanted to take re-routing effects due to control actions into account. It was found that, compared to the no control case, a significant improvement in the total time spent could be realized by anticipative MPC-based ramp metering control.

In the appendix, we have reconsidered the first-order macroscopic traffic flow model of Lighthill, Whitham and Richards and we have proposed some modifications that solve some of the deficiencies related to this model, while at the same time preserving one of the main advantages of this model compared to the higher order models, namely the low computational complexity. A low computational complexity is important if the model is used in on-line real-time traffic control applications.

The contributions of this thesis to the state of the art can be summarized as follows:

- we presented an efficient interpolation scheme for the estimation of missing traffic measurement data using a reference day,
- we made a distinction between the simulation time step of the simulation model and the control time step of the controller,
- we implemented a METANET simulation model for the E17 motorway case study,
- we applied the model predictive control framework to ramp metering control,
- we used the simulation model of the case study to assess the performance of MPC-based single on-ramp control and coordinated MPC-based ramp metering control,
- we developed the combined MPC-identification approach to ramp metering, and
- we developed anticipative MPC-based ramp metering control.

7.2 Further research

Some open questions and possible directions for further research are:

- In this thesis, we focused mainly on MPC-based ramp metering for motorways. Although we did not deal with the other traffic control measures presented in Section 2.3, all of these traffic control measures can be used in an MPC-framework, e.g., variable speed limits as was shown in [22, 129]. However, some of these traffic control measures, such as e.g., lane closures, lead to integer optimization problems, which are computationally very intensive. The application of these control measures in the MPC-framework should be further investigated, especially with regard to the computational complexity as these systems need to be used in real-time on-line applications. For some traffic control measures compliance rates need to be considered since they are indicative rather than obliged (e.g., route guidance).
- After the investigation of other traffic control measures than ramp metering, it is a logical step to extend coordinated control to integrated control, or the combination of different traffic control measures. We expect that the performance of a motorway system with ramp metering can

be enhanced by adding other control measures such as e.g., route guidance, variable speed limits, ... due to the additional means to actuate the traffic flows towards efficient traffic operation. Also the combination of MPC-based advanced traffic management systems with new technologically advanced traffic control measures such as e.g., intelligent speed adaptation or in the future even with automated intelligent vehicle and highway systems looks very promising.

- As we have shown in Chapter 4, an MPC controller uses a prediction model to simulate the traffic system over the control horizon. In this thesis we have assumed the future traffic demands to be known. However, in a real-life application, an estimation of the future traffic demands over the prediction horizon needs to be made. The estimation of the traffic demands could be made based on the current upstream traffic measurements, which are an indication for the future traffic demands downstream, and based on historical traffic data. The performance of this method, as well as the sensitivity of the MPC-based controller on the prediction of the traffic demand still needs to be investigated.

- It would be interesting to implement and verify the performance of the combined MPC-identification approach to ramp metering presented in Chapter 6 in a real-life situation. Since it is hardly possible to run experiments on a real-life situation, the performance of the combined MPC-identification approach could be simulated using e.g., a microsimulation model with varying motorway properties. Simulating the real-life situation using a detailed traffic microsimulation model in order to assess the performance of traffic control strategies was recently done by Ben-Akiva *et al.* in [16] and by Logghe and Immers in [90–92]. Topics that need to be further investigated in this respect are the choice of the update interval of the prediction model parameters and the computational complexity of the combined MPC-identification approach to ramp metering. Note that in order to set up a realistic simulation experiment, the dynamics of the motorway properties in real-life situations need to be investigated. As a starting point, the impact of the weather conditions on the traffic measurements from Chapter 2 and on the identified METANET traffic flow model parameters could be investigated.

- The case study we have investigated in Chapter 5 consisted of a single motorway corridor. Since there are five on-ramps present in the case study, we were able to investigate the effect of coordinated ramp metering. However, due to the combination of the motorway configuration and the traffic demands in the case study, coordinated ramp metering was not able to yield much performance improvement for the case study. Coordinated ramp metering should be investigated for other combinations of traffic demands and motorway topology in order to fully assess the improvements that can be realized. In particular, we suggest to investigate

the performance of coordinated and integrated control in long motorway corridors with many on-ramps.

- Some publications in the literature describe the occurrence of instabilities if traffic assignment and traffic control are combined [132]. Since the anticipative MPC-based ramp metering control takes the traffic state history into account and since the delay in the route choice of the drivers is taken into account, we expect to be able to avoid the instabilities mentioned in [132]. If necessary, we can add an extra term to the MPC objective function to impose stability. The choices of the length of the traffic state history that is taken into account τ_{info} and the time constant τ_{evol} for the evolution from the current traffic assignment to the equilibrium assignment need further investigation, especially the relation of these parameters with the occurrence of instabilities.

- Since ramp metering can cause delays for the traffic that wants to enter the motorway (on-ramp queue), there will be re-routing effects on the urban network due to ramp metering (i.e., drivers taking the next or the previous on-ramp), especially in areas with short distances between the on-ramps (as is the case in Belgium). If we want to include these effects in the motorway corridor model, we need to extend the motorway model with the alternative urban routes and use the anticipative MPC-based ramp metering strategy.

- In Chapter 4 the total time spent in the network plus an extra term added to smooth the control signals, was chosen as the objective function. Currently we are involved in a project sponsored by the Belgian government to implement MPC-based control using a sustainable cost function which consists of the total time spent by all the vehicles in the network extended with terms to account for traffic safety, pollution, travel time predictability, ... Since the MPC-based ramp metering controller makes a trade-off between several conflicting objectives, the impact of the different terms in the objective function as well as the resulting traffic operation are being investigated.

- Another topic for further research is the coordinated and the integrated control of very large networks. If the motorway traffic network to be controlled grows to the dimensions of large areas or even countries, the MPC-based control problem becomes intractable. Indeed, the increased number of control signals that we need to determine as more traffic control set-ups are coordinated or integrated, drastically increases the computational complexity of the optimization problem to be solved. In order to tackle this problem we would need to investigate hierarchical control or multi-agent control [93, 125, 139].

- An important topic to be investigated is the robustness and the reliability of the controlled traffic network. In this thesis, the focus has been pri-

marily on optimal control of traffic. However, as a controller deals with larger networks, the reliability of the traffic operation becomes more important. E.g., a route that becomes blocked due to an incident may not lead to traffic breakdown in a whole area. The robustness of an MPC-based controller to unforeseeable events needs to be investigated and this concept of robustness needs to be incorporated in the controller.

- A motorway network does not operate on its own but it interacts with the urban road network. As such, the control of motorway traffic could be integrated with the control of urban traffic in order to obtain a smooth operating road transportation system [70]. This yields a large control problem that can be tackled using hierarchical control or multi-agent control. Currently, the traffic data needed for such an advanced system are generally not available, however we expect that due to decreasing traffic detector prices and due to the emergence of new technologies to obtain traffic data such as e.g., floating car data from cellular phones, sufficient traffic data will become available in the future.

Appendix A

An improved first-order model

We reconsider the first-order macroscopic traffic flow model of Lighthill, Whitham and Richards [87,123] and we propose some modifications that solve some of the deficiencies related to this model, while at the same time preserving one of the main advantages of this model compared to higher-order models, namely its low computational complexity, which is important if the model is used in on-line real-time traffic control applications such as e.g. the MPC framework presented in Chapter 4. Although a first-order model is less detailed than higher order models, first-order models play an important role for the simulation of very large networks or in applications where simulation speed is more important than accuracy.

A.1 Introduction

In this appendix we reconsider the first-order macroscopic traffic flow model of Lighthill, Whitham and Richards [87, 123]. Although this model has since long been complemented by second-order models [31,110,117] in order to avoid several deficiencies and to improve the modeling accuracy [107], it is still used in practice especially in on-line control applications [30] for large networks where the computational complexity of second-order models is too high.

In this appendix we will briefly revisit the model of Lighthill, Whitham and Richards and then illustrate some of the problems that can occur when this model is used for simulation. Next, we present some modifications for the model that take care of these problems without increasing the order of the model such

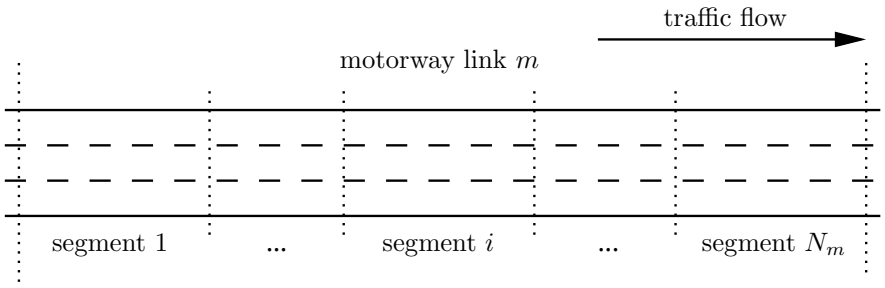


Figure A.1: Schematic representation of a motorway stretch. In a similar way as for the METANET model presented in Chapter 3, the motorway stretch is modeled as a link, which is divided into segments. A segment is characterized by a couple (m,i) where m is the link index and i is the segment index in link m.

that the computational requirements of the improved model remain low. As we point out in Remark A.1, a similar result as the modifications presented in this appendix was also presented by Daganzo.

A.2 The traffic flow model of Lighthill and Whitham

We present a discretized version [30, 95, 109] of the macroscopic traffic flow model that was introduced by Lighthill and Whitham in 1955 [86, 87] and independently by Richards in 1956 [123]. For the sake of simplicity of the exposition, we do not consider on-ramps and off-ramps. Note however that the model can easily be adapted to include on-ramps and off-ramps. Also for the simplicity of the exposition, the numerical scheme resulting from the discretization of the continuous model will be referred to as the LWR model.

In a similar way as described in Chapter 3 for the METANET model, the motorway can be discretized in links and segments (see Figure A.1) [32]. A typical length of the segments is 500 m. Recall from Chapter 3 that the structural properties of the segments in a link are assumed to be constant. Whenever there is a change in the properties of the motorway, a new link is defined. The Lighthill, Whitham and Richards model is also discretized in time. The simulation step is ΔT_{sim}. A typical value for ΔT_{sim} is 15 s [32].

The LWR model is a first-order macroscopic model of which the traffic state in a motorway segment is uniquely described by only one state variable, the traffic density. The traffic flow and the average speed in a segment can then

be derived from the traffic density. These state variables are updated at each simulation step.

Next, we give an overview of the equations that provide the relations between the traffic density, the average speed and the flow in a segment.

Let $q_{\text{in},m,i}(l)$ and $q_{\text{out},m,i}(l)$ be respectively the inflow and the outflow for segment i of link m in the period $[l\Delta T_{\text{sim}}, (l+1)\Delta T_{\text{sim}})$. The conservation of flow then results in

$$\rho_{m,i}(l+1) = \rho_{m,i}(l) + \frac{\Delta T_{\text{sim}}}{n_m l_{m,i}} \big[q_{\text{in},m,i}(l) - q_{\text{out},m,i}(l)\big], \tag{A.1}$$

where $l_{m,i}$ is the length of the i-th segment of link m and n_m is the number of lanes in link m. This equation is identical to (3.2) from the METANET model discussed in Chapter 3.

In Chapter 2 it was illustrated that the average speed in a segment varies with varying traffic density. Although this relation is highly nonlinear, it is often approximated by a two-segment linear function [27, 104, 159]. However, the nonlinear behavior of the speed-density relation cannot completely be captured by this piece-wise linear approximation [30]. Hence, the following nonlinear empirical formula is often used in the LWR model instead of the approximation [95]:

$$v_{m,i}(l) = V[\rho_{m,i}(l)] \triangleq v_{\text{free},m}\left[1 - \left(\frac{\rho_{m,i}(l)}{\rho_{\text{jam},m}}\right)^{\alpha}\right]^{\beta} \tag{A.2}$$

where $v_{\text{free},m}$ is the free flow speed in the link and $\rho_{\text{jam},m}$ is the jam density as described in Chapter 2. The relation between $v_{m,i}(l)$ and $\rho_{m,i}(l)$ presented in (A.2) is plotted in Figure A.2 for typical values of the parameters $v_{\text{free},m}$, $\rho_{\text{jam},m}$, α and β. Note that in general $v_{\text{free},m}$ and $\rho_{\text{jam},m}$ are different for each link since they are the structural parameters that describe the properties of the motorway. Features that influence these parameters are e.g., the curvature of the link, the width of the lanes, ... The parameters $v_{\text{free},m}$ and $\rho_{\text{jam},m}$ are generally fitted on traffic data using regression techniques [32] (see also Section 3.4) while the parameters α and β are usually taken from literature [95] since they have a smaller impact on the quality of the fit.

Given the traffic density and the average speed in a segment, the traffic flow in the segment is given by:

$$q_{m,i}(l) = \rho_{m,i}(l) v_{m,i}(l) n_m \ . \tag{A.3}$$

If a motorway link or a motorway network is simulated, consecutive segments have to be coupled together. This can be done in two ways:

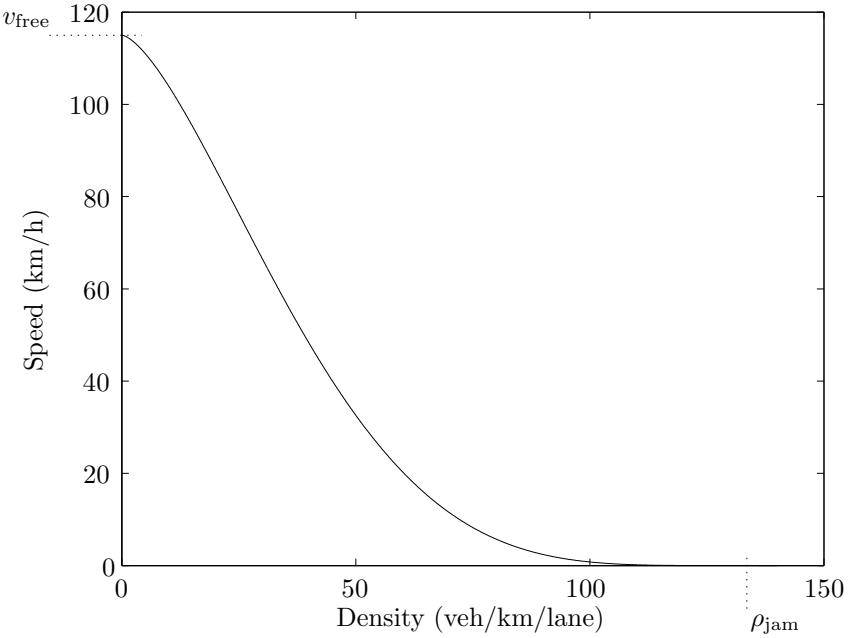


Figure A.2: The nonlinear speed-density relation according to (A.2). For this plot we used $v_{\text{free},m} = 115\,\text{km/h}$, $\rho_{\text{jam},m} = 136\,\text{vehicles/km/lane}$, $\alpha = 1.5$ and β=5 as was found by fitting (A.2) to traffic data of the E17 motorway Ghent–Antwerp in Section 2.2.4.

- the inflow of segment i can be set equal to the flow in the i-th segment:

$$q_{\text{in},i}(l) = q_i(l) \ . \tag{A.4}$$

  This will be called *upstream coupling* as will be explained later on.

- the outflow of segment i can be set equal to the flow of the i-th segment:

$$q_{\text{out},i}(l) = q_i(l) \ . \tag{A.5}$$

  This will be called *downstream coupling*.

Furthermore, we have $q_{\text{in},i}(l) = q_{\text{out},i-1}(l)$.
We shall see later on that expressions (A.4) and (A.5) both result in a discrete first-order model for the motorway or motorway network. Note that it is possible to define more complex coupling relations. However, this leads to more complex or to higher-order models.

Equations (A.4) and (A.5) both result in a different numerical scheme for the first-order LWR model.

The average speed and the flow in a segment can be determined based on the density $\rho_{m,i}(l)$ in the segment via (A.2) and (A.3). The difference between the coupling equations (A.4) and (A.5) becomes clearer if the nonlinear state space model for both coupling equations is rewritten with the expressions for the average speed and the flow substituted.

If we combine (A.1), (A.2), (A.3) and (A.4) we obtain the following nonlinear difference equation for the traffic density in segment i of link m:

$$\rho_{m,i}(l+1) = \rho_{m,i}(l) + \frac{\Delta T_{\text{sim}}}{l_{m,i}} \Big(\rho_{m,i}(l) V[\rho_{m,i}(l)] - \rho_{m,i+1}(l) V[\rho_{m,i+1}(l)] \Big) \quad . \tag{A.6}$$

So $\rho_{m,i+1}(l)$ and $\rho_{m,i}(l)$ determine $\rho_{m,i}(l+1)$. Therefore, (A.4) is called *upstream coupling*. Figure A.3 clarifies this naming. Three successive motorway segments are represented at two consecutive simulation steps. The thick arrows in Figure A.3 represent the information flows between the segments due to the upstream coupling described by (A.4). Clearly, the information flows in the upstream direction of the motorway. Note that (A.6) is a first-order model in which only one variable is needed to completely characterize the traffic situation in a given segment, i.e., there is only one independent variable per segment.

When combining (A.1), (A.2), (A.3) and (A.4) we obtain the following nonlinear difference equation for the traffic density in segment i of link m:

$$\rho_{m,i}(l+1) = \rho_{m,i}(l) + \frac{\Delta T_{\text{sim}}}{l_{m,i}} \Big(\rho_{m,i-1}(l) V[\rho_{m,i-1}(l)] - \rho_{m,i}(l) V[\rho_{m,i}(l)] \Big) \quad . \tag{A.7}$$

So $\rho_{m,i-1}(l)$ and $\rho_{m,i}(l)$ determine $\rho_{m,i}(l+1)$. The coupling described by (A.5) is called *downstream coupling*. In Figure A.4 a graphical representation of the data flows in time and space for the downstream case is given.

A.3 Model behavior and deficiencies

In this section we discuss some specific cases to illustrate some properties and deficiencies of the discrete models obtained by (A.6) and (A.7). In Section A.4 we will then present some improvements of the model that will address the problems encountered in this section.

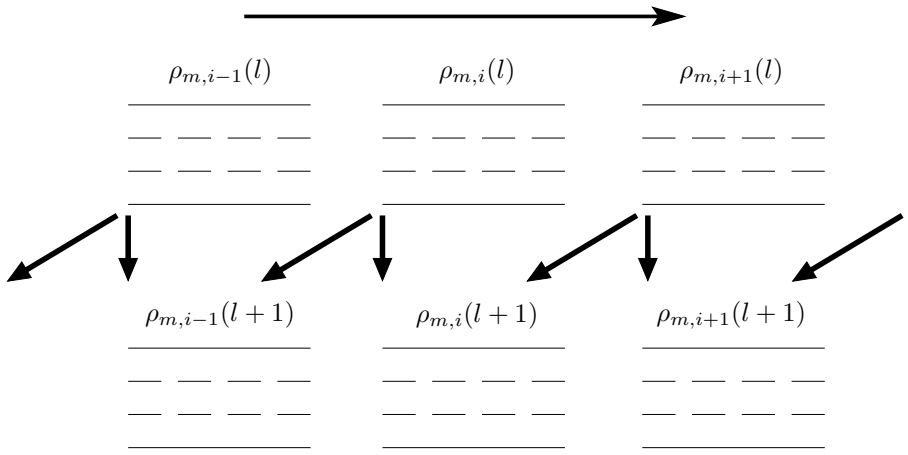


Figure A.3: Schematic presentation of upstream coupling. The figure shows the same motorway at two consecutive simulation steps l and $l+1$. The thick arrows between the segments indicate the dependencies of the states of the segments through time as described by (A.6).

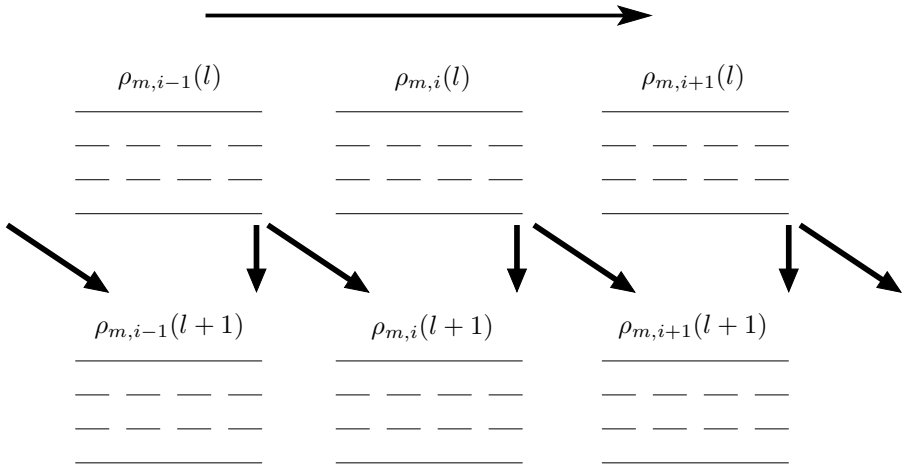


Figure A.4: Schematic representation of downstream coupling. The figure shows three successive motorway segments at two consecutive simulation steps l and $l+1$. The thick arrows between the segments indicate the dependencies of the states of the segments through time as described by (A.5).

A.3.1 Upstream coupling

Consider the situation depicted in Figure A.3 where information about the state of a segment is provided to the upstream segment. Now suppose that segment i is empty at simulation step l. So $\rho_{m,i}(l) = 0$. The segments i and $i+1$ of link m are now observed at simulation step $l+1$.

- Segment i: Since $\rho_{m,i}(l) = 0$, it follows from (A.3) that $q_{m,i}(l) = 0$ and thus also $q_{\text{in},m,i}(l) = 0$ according to (A.4). This means that segment i remains empty independently of what happens upstream of i. This is an unrealistic situation which will not occur in practice unless a road is completely blocked upstream.
- Segment $i+1$: According to (A.4), the inflow of segment $i+1$ is given by $q_{\text{in},m,i+1}(l) = q_{m,i+1}(l)$. Hence,

  $$q_{\text{out},m,i}(l) = q_{\text{in},m,i+1}(l) = \rho_{m,i+1}(l) v_{m,i+1}(l) n_m,$$

  which is not necessarily zero. So in this case vehicles are assumed to leave segment i even though the segment is empty. This implies that cars are "created" out of the blue and that the flow conservation property is violated as well as the physical constraint that the density in i must be nonnegative.

It is clear from the two presented cases that for upstream coupling not only the state of the considered segment has to be taken into account but also the state of the preceding segment. Sometimes this is implemented by using a more complex coupling equation such as

$$q_{\text{in},m,i}(l) = \gamma q_{m,i-1}(l) + (1-\gamma) q_{m,i}(l) \ . \tag{A.8}$$

The estimation of γ and the lack of a physical motivation are two drawbacks of (A.8) [32, 114]. Moreover, in order to avoid under all circumstances that vehicles are created out of the blue in an empty upstream segment as described above, γ needs to be chosen equal to 1, which leads to downstream coupling (see (A.5)). In the next subsection we will show that downstream coupling does not exhibit the problems connected to empty segments that have been discussed above.

A.3.2 Downstream coupling

Consider the situation of Figure A.4 where the information of the state of a segment is provided to the downstream segment. The properties of the downstream coupled model are illustrated with some examples.

If it is assumed that segment i is empty at simulation step l, the following phenomena are observed:

- If $q_{m,i-1}(l) \neq 0$ then vehicles can enter the empty segment i from the upstream segment $i-1$ since $q_{\text{in},m,i}(l) = q_{\text{out},m,i-1}(l) = q_{m,i-1}(l)$ by (A.5). So in that case segment i starts filling up at simulation step $l+1$.
- If segment i is empty at simulation step l, $\rho_{m,i}(l) = 0$ and thus by (A.3) also $q_{m,i}(l) = 0$. Hence, $q_{\text{out},m,i}(l) = q_{m,i}(l) = 0$ by (A.5).

So in case $\rho_{m,i}(l) = 0$, the problems of blocking and vehicle generation out of the void that occurred with upstream coupling have disappeared by using downstream coupling. This does not mean that all problems have disappeared though. The following two cases illustrate that also with downstream coupling (dual) problems can occur. Assume that segment i of link m is congested at simulation step l, which means that $\rho_{m,i}(l) = \rho_{\text{jam},m}$. Note that this implies that $v_{m,i}(l) = 0$ and thus $q_{m,i}(l) = 0$ as can be seen from (A.2) and (A.3). In what follows, we investigate what happens in segments i and $i+1$ of link m at simulation step $l+1$:

- Segment i: By (A.5) the inflow in segment i is given by $q_{\text{in},m,i}(l) = q_{\text{out},m,i-1}(l) = q_{m,i-1}(l) = \rho_{m,i-1}(l)v_{m,i-1}(l)n_m$. So the inflow in segment i depends on the outflow of segment $i-1$ but not on the state of segment i. This means that in the model with downstream coupling as it was presented before, the flow of vehicles leaving a segment is always fed into the next segment irrespective the state of the next segment. This poses problems when the state of the segment receiving the flow is such that the segment is not able to accept the vehicles. E.g., in the example presented here, vehicles enter into segment i even though the density $\rho_{m,i}(l)$ is already maximal and $q_{\text{out},m,i}(l) = q_{m,i}(l) = 0$. This implies a violation of the flow conservation property since the entering vehicles cannot be accounted for by an increased density (the jam density ρ_{jam} is the upper limit of the density). It seems as if the vehicles disappear. The state of the downstream segment i should not only be taken into account when calculating the flow out of i, but it should also be used to check whether the inflow does not exceed the maximum allowable value. A discussion of how this can be done is provided in the Section A.4.
- Segment $i+1$: According to (A.5) the inflow in segment $i+1$ is given by $q_{\text{in},m,i+1}(l) = q_{\text{out},m,i}(l) = q_{m,i}(l) = 0$. This has two consequences.

First, the density in segment $i+1$ can only decrease since the inflow is always zero. Second, once a segment, say i, becomes congested it stays congested throughout the whole simulation since the outflow stays zero indefinitely.

It is clear that the behavior of the model in the two cases above (upstream and downstream coupling) does not mimic the behavior of real traffic flows in an analogous real-life situation very well. In real-life, traffic state information propagates downstream during free-flow operation and upstream during congested traffic operation [95]. Therefore, we modify the model such that its behavior better corresponds to real-life traffic flows. However, the complexity of the model is kept as low as possible e.g., for application of the model in on-line applications. Two modifications to the numerical scheme are presented in the next section.

A.4 An improved first-order traffic flow model

There exist different solutions to overcome the model deficiencies presented in the previous section.

A first solution consists in choosing a more advanced model. Examples of such more advanced models are the second-order models where the speed in the segment is taken as an extra state in the model and a difference equation replaces (A.2) such as e.g., the METANET model presented in Chapter 3. These models are able to better deal with the problems concerning the flow conservation law as well as with the problem of the outflow out of a congested segment [107,133]. Information about the traffic density and the average velocity in the neighboring segment is processed in the model. However, these more advanced models require more parameters to be tuned when calibrating the model. Moreover, they have a higher computational complexity, which is disadvantageous if the model is used in on-line applications where computation speed is important.

In the case of upstream coupling, the state of the upstream segment could be taken into account in order to prevent cars from being created and to allow an empty segment to fill. When applying coupling equation (A.8), a new parameter is introduced which has to be tuned [32] and which has no physical motivation. Moreover, in order to prevent cars from being created at an empty segment, we must choose $\gamma = 1$, which is equivalent to (A.5). Therefore, we start from downstream coupling and address the following two problems that occur with downstream coupling:

- overflow in a segment, and
- inability to simulate disappearing congestion.

To prevent overflow in segment i, the net inflow in the segment must be limited such that the density in segment i cannot grow larger than the maximal density $\rho_{\text{jam},m}$. This implies that the following condition should always be satisfied:

$$\rho_{m,i}(l) + \frac{\Delta T_{\text{sim}}}{l_{m,i} n_m} \big[q_{\text{in},m,i}(l) - q_{\text{out},m,i}(l) \big] \leq \rho_{\text{jam},m}$$

or equivalently

$$q_{\text{in},m,i}(l) \leq q_{\text{out},m,i}(l) + \frac{l_{m,i} n_m}{\Delta T_{\text{sim}}} \big[\rho_{\text{jam},m} - \rho_{m,i}(l) \big] \ . \tag{A.9}$$

Equation (A.9) can be interpreted as follows. The maximal inflow in a segment i is equal to the flow leaving the segment, $q_{\text{out},m,i}(l)$, plus a flow of cars which are absorbed by the segment and which bring the traffic density to its maximal value, $\rho_{\text{jam},m}$. The inflow of a segment can be decomposed in two parts: a flow that leaves the segment again (throughput) and a flow changing the density in the segment (storage capacity of the segment).

Equation (A.9) provides an upper bound for the inflow in segment i. By combining this upper bound (A.9) with (A.5), a coupling between the segments $i-1$ and i is realized, which takes the states of both segments into account. The resulting expression for the inflow in segment i is given by

$$\begin{aligned} q_{\text{in},m,i}(l) = \min \Bigg(& q_{m,i-1}(l), \\ & q_{\text{out},m,i}(l) + \frac{l_{m,i} n_m}{\Delta T_{\text{sim}}} \big[\rho_{\text{jam},m} - \rho_{m,i}(l) \big] \Bigg) \ . \end{aligned} \tag{A.10}$$

This equation prevents cars from disappearing (overflow of the density in segment i).

Finally the inconsistent behavior of the model when congestion occurs needs to be solved. Once $\rho_{\text{jam},m}$ is reached in a segment the outflow of this segment becomes zero according to (A.2) and remains zero indefinitely. This does not correspond to the behavior that can be observed in practice. In order to prevent the outflow from becoming zero, (A.2) is modified as follows:

$$v_{m,i}(l) = \begin{cases} v_{\text{free},m} \left[1 - \left(\dfrac{\rho_{m,i}(l)}{\rho_{\text{jam},m}} \right)^{\alpha} \right]^{\beta} & \text{if } \rho_{m,i}(l) \leq (1-\varepsilon)\rho_{\text{jam},m} \\ \zeta v_{\text{free},m} & \text{if } \rho_{m,i}(l) > (1-\varepsilon)\rho_{\text{jam},m} \end{cases} \tag{A.11}$$

where ζ and ε are small positive real numbers. A discontinuity of $v_{m,i}(l)$ for $\rho_{m,i}(l) = (1-\varepsilon)\rho_{\text{jam},m}$ can be avoided by by imposing the following relation between ε and ζ:

$$\zeta = \big(1 - (1-\varepsilon)^{\alpha} \big)^{\beta} \tag{A.12}$$

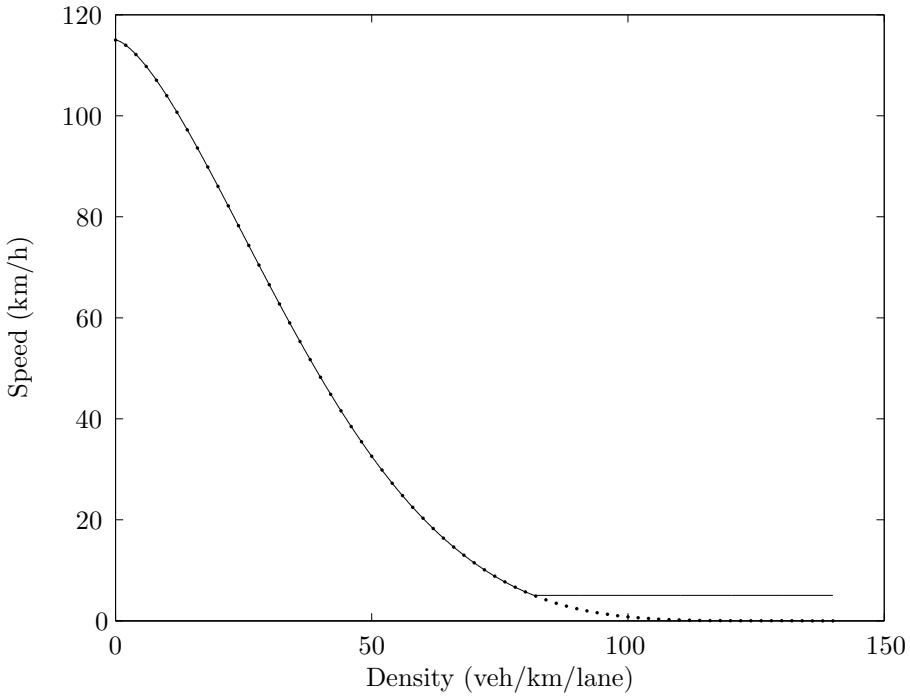


Figure A.5: The altered speed-density relation as given by (A.11). For traffic densities larger than $(1-\varepsilon)\rho_{\text{jam},m}$, the speed is equal to $\zeta v_{\text{free},m}$. The dashed line represents the original speed-density relation given by (A.2) and plotted in Figure A.2.

The corrected speed-density relation according to (A.11) is plotted in Figure A.5. In order to avoid a discontinuity in (A.11), the parameters ε and ζ are chosen such that they satisfy (A.12).

The two corrections presented in this section enable the simulation of the traffic flows in a motorway (network) using a first-order model. Since the corrected model uniquely describes the state of a segment using one traffic state it is still of first order although it does no longer exhibit the problems of the original model that were discussed in the previous section.

Remark A.1 Daganzo [34, 35, 37] described a similar result as the correction we presented in (A.10) by defining a traffic supply (send) and a traffic demand (receive) function in order to take the upstream and the downstream propagation of traffic state information between the segments into account. Lebacque showed in 1996 [83] that this approach also follows from the Godunov [51] discretization scheme applied to the continuous LWR model. □

A.5 Conclusions

In this appendix, we have considered the first-order macroscopic traffic flow model of Lighthill, Whitham and Richards. The main advantage of this model is the trade-off it provides between model accuracy and computational complexity. Indeed, for some applications, such as e.g. on-line traffic control applications, the computational complexity of the more detailed higher order models is too high while the level of detail provided by these models is not always necessary. In these circumstances, the first-order traffic flow model of Lighthill, Whitham and Richards provides a viable alternative.

First, we have shown that simple numerical schemes, such as e.g. downstream coupling and upstream coupling of motorway segments, exhibit some deficiencies. Next, we proposed two modifications to the downstream coupling of the segments, which resolve the two main shortcomings of downstream coupling: the overflow in congested segments and the poor capability to simulate resolving congestion in a segment. While resolving the two shortcomings mentioned above, these modifications preserve the low computational complexity of the model, which is the main advantage of the first-order model compared to higher-order models.

Bibliography

[1] Detection Technology for IVHS - Final Report Addendum. Technical report, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 1995. Contract DTFH61-91-C-00076.

[2] Detection Technology for IVHS - Task L Final Report. Technical report, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 1995. Contract DTFH61-91-C-00076.

[3] J. D. Anderson. *Modern compressible flow: With historical perspective*. Mc Graw Hill, 2002.

[4] K. Ashok and M. Ben-Akiva. *Dynamic Origin-Destination Matrix Estimation and Prediction for Real-Time Traffic Management Systems*. Elsevier Science, 1993.

[5] A. Aw and M. Rascle. Resurrection of second order models of traffic flow. *SIAM journal on applied mathematics*, 60(3):916–938, 2000.

[6] J. Barceló. Dynamic network simulation with Aimsun. In *Proceedings of the International Symposium on Transport Simulation*, Yokohama, August 2002. http://www.aimsun.com/documents.html.

[7] T. Bellemans, B. De Schutter, and B. De Moor. Data acquisition, interfacing and pre-processing of highway traffic data. In *Proceedings of Telematics Automotive 2000*, pages 4/1–4/7, Birmingham, UK, April 2000. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `00-43.ps`.

[8] T. Bellemans, B. De Schutter, and B. De Moor. On data acquisition, modeling and simulation of highway traffic. In *Proceedings of the 9th International Federation of Automatic Control (IFAC) Symposium on Transportation Systems*, pages 22–27, Braunschweig, Germany, June 2000. IFAC. Available by ftp from

`ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `00-42.ps`.

[9] T. Bellemans, B. De Schutter, and B. De Moor. An improved first-order macroscopic flow model for highway traffic simulation. In *Proceedings of the 2001 American Control Conference (ACC 2001)*, pages 2105–2109, Washington DC, Virginia, June 2001. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `00-45.ps`.

[10] T. Bellemans, B. De Schutter, and B. De Moor. Model predictive control with repeated model fitting for ramp metering. 2002. Submitted for publication in a journal. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `02-185.ps.`

[11] T. Bellemans, B. De Schutter, and B. De Moor. Model predictive control with repeated model fitting for ramp metering. In *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems (ITCS2002)*, pages 236–241, Singapore, September 2002. (CDROM), Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `02-35.pdf`.

[12] T. Bellemans, B. De Schutter, and B. De Moor. Models for traffic control. *Journal A*, 43(3–4):13–22, 2002. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `01-89.pdf`.

[13] T. Bellemans, B. De Schutter, and B. De Moor. Anticipative model predictive control for ramp metering in freeway networks. In *Proceedings of the 2003 American Control Conference (ACC 2003)*, 2003. Accepted for publication. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `02-153.ps`.

[14] T. Bellemans, B. De Schutter, and B. De Moor. Model predictive control for ramp metering of motorway traffic : A case study. 2003. Submitted for publication in a journal. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `02-207.ps.`

[15] T. Bellemans, S. Logghe, B. De Moor, and B. Immers. Het fileprobleem in België : Wiskundige modellen, analyse, regeling en acties. Technical report, ESAT-SCD (SISTA), K.U.Leuven, Leuven, Belgium, 2002. Final report for DWTC project MD/01/24. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/reports` as `02-15.pdf`. In Dutch.

[16] M. Ben-Akiva, D. Cuneo, M. Hasan, M. Jha, and Q. Yang. Evaluation of freeway control using a microscopic simulation laboratory. *Transportation Research C*, 11(1):29–50, 2003.

[17] E. Bernauer, L. Breheret, S. Algers, M. Boero, C. Di Taranto, M. Dougherty, K. Fox, and J.-F. Gabard. A review of micro-simulation models. Webextract, March 1998. http://www.its.leeds.ac.uk/projects/smartest/intro.html.

[18] P.T. Boggs and J.W. Tolle. Sequential quadratic programming. *Acta Numerica*, 4:1–51, 1995.

[19] H. Botma. State-of-the-art report "Traffic Flow Models". Technical Report R-78-40, SWOV, 1978.

[20] G.E. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, USA, 1976.

[21] D. Branston. Models of single lane time headway distributions. *Transportation Science*, 10:125–148, 1976.

[22] P. Breton, A. Hegyi, B. De Schutter, and H. Hellendoorn. Shock wave elimination/reduction by optimal coordination of variable speed limits. In *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems*, pages 225–230, Singapore, September 2002.

[23] K. A. Brookhuis and D. De Waard. Limiting speed, towards an Intelligent Speed Adapter (ISA). *Transportation Research F*, pages 81–90, 1999.

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

[25] E.F. Camacho and C. Bordons. *Model Predictive Control in the Process Industry*. Springer-Verlag, Berlin, Germany, 1995.

[26] Cambridge Systematics, Inc., Oakland, California. *Twin Cities Ramp Meter Evaluation – Final Report*, February 2001. Prepared for the Minnesota Department of Transportation.

[27] G.L. Chang, J. Wu, and S.L. Cohen. Integrated real-time ramp metering model for nonrecurrent congestion: Framework and preliminary results. *Transportation Research Record*, (1446):56–65, 1994.

[28] N. A. Chaudhary and C. J. Messer. Ramp metering technology and practice. Technical Report 2121-1, Texas Transportation Institute, The Texas A&M University System, College Station, Texas 77843-3135, May 2000.

[29] C. Chen, Z. Jia, and P. Varaiya. Causes and cures of highway congestion. *IEEE Control Systems Magazine*, pages 26–32, December 2001.

[30] O.J. Chen, A.F. Hotz, and M.E. Ben-Akiva. Development and evaluation of a dynamic ramp metering control model. In *Proceedings of the 8th IFAC/IFIP/IFORS Symposium on Transportation Systems*, pages 1162–1168, Chania, Greece, June 1997.

[31] M. Cremer and A.D. May. An extended traffic flow model for inner urban freeways. In *Proceedings of the 5th IFAC/IFIP/IFORS International Conference on Control in Transportation Systems*, pages 383–388, Vienna, Austria, 1986.

[32] M. Cremer and M. Papageorgiou. Parameter identification for a traffic flow model. *Automatica*, 17(6):837–843, 1981.

[33] C.R. Cutler and B.L. Ramaker. Dynamic matrix control – a computer control algorithm. In *Proceedings of the 86th AIChE National Meeting*, Houston, Texas, April 1979.

[34] C. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research B*, 28B(4):269–287, 1994.

[35] C. Daganzo. The cell transmission model, Part II: Network traffic. *Transportation Research B*, 2(2):79–93, 1995.

[36] C. Daganzo. Requim for second-order fluid approximations of traffic flow. *Transportation Research B*, 29B(4):277–286, 1995.

[37] C. F. Daganzo. A finite difference approximation of the kinematic wave model. *Transportation Research B*, 29B(4):261–276, 1995.

[38] C. F. Daganzo and Y. Sheffi. On stochastic models of traffic assignment. *Transportation Science*, 11:253–274, 1977.

[39] C.F. Daganzo. *Fundamentals of Transportation and Traffic Operations*. Pergamon, 1997.

[40] B. De Moor and T. Bellemans. Niet opgelost, wel uitgelegd : file ! file ! file ! *TEEK! magazine*. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `99-90.pdf`. In Dutch, Nov 1999.

[41] B. De Schutter, T. Bellemans, S. Logghe, J. Stada ad B. De Moor, and B. Immers. Advanced traffic control on highways. *Journal A*, 40(4):42–51, 1999. Available by ftp from `ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/bellemans/papers` as `99-59.ps`.

[42] E. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.

[43] R. C. Dorf and R. H. Bishop. *Modern control systems*. Addison Wesley, 1995.

[44] J. S. Drake, J. L. Schofer, and A. D. May. A statistical analysis of speed density hypothesis. In *Proceedings of the Third Symposium on the Theory of Traffic Flow*, New York, 1967. Elsevier North Holland, Inc.

[45] D. R. Drew. *Traffic Flow Theory and Control*, chapter 12. McGraw Hill Book Company, New York, 1968.

[46] L. Elefteriadou. Freeway merging operations: A probabilistic approach. In *Proceedings of the 8th International Federation of Automatic Control (IFAC) Symposium on Transportation Systems*, pages 1351–1356, Chania, Greece, 1997.

[47] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company, San Francisco, 1979.

[48] N. H. Gartner, C. J. Messer, and A. Rathi. Traffic flow theory. Technical report, Transportation Research Board, 1997. Available on the web at `http://www.tfhrc.gov/its/tft/tft.htm`.

[49] S. Geerts and E. Kenis. Verkeerscentrum Vlaanderen leidt verkeer in goede banen. *Het Ingenieursblad*, pages 28–37, October 2001. In Dutch.

[50] D. L. Gerlough and M. J. Huber. Traffic flow theory: a monograph. Technical report, Transportation Research Board, Washington D.C., 1975. Special Report 165.

[51] S. Godunov. A difference scheme for numerical computation of discontinuous solutions of equations of fluid dynamics. *Mathematics of the USSR - Sbornik*, 47(89):271–306, 1959.

[52] H. Greenberg. An analysis of traffic flow. *Operations Research*, 7:78–85, 1959.

[53] B. D. Greenshields. A study in highway capacity. In *Proceedings of the Highway Research Board*, volume 14, page 458, 1935.

[54] M. Hasan, M. Jha, and M. Ben-Akiva. Evaluation of ramp control algorithms using microscopic traffic simulation. *Transportation Research C*, 10:229–256, 2002.

[55] Masroor Hasan. Evaluation of ramp control algorithms using a microscopic traffic simulation laboratory, MITSIM. Master's thesis, Massachusetts Institute of Technology, February 1999.

[56] A. Hegyi, B. De Schutter, and J. Hellendoorn. Optimal coordination of variable speed limits to suppress shock waves. In P.H.L. Bovy, editor, *Proceedings of the 7th TRAIL Congress 2002*, pages 197–220, Rotterdam, The Netherlands, November 2002.

[57] A. Hegyi, B. De Schutter, and H. Hellendoorn. Model predictive control for optimal coordination of ramp metering and variable speed control. In *Proceedings of the 1st International ICSC Conference on Neuro-Fuzzy Technologies (NF 2002)*, Havana, Cuba, January 2002.

[58] A. Hegyi, B. De Schutter, H. Hellendoorn, and T. van den Boom. Optimal coordination of ramp metering and variable speed control — An MPC approach. In *Proceedings of the 2002 American Control Conference*, pages 3600–3605, Anchorage, Alaska, May 2002.

[59] D. Heidemann. Some critical remarks on a class of traffic flow models. *Transportation Research B*, 33:153–155, 1999.

[60] D. Helbing. Gas-kinetic derivation of Navier-Stokes-like traffic equations. *Physical Review*, E53(3):2266–2381, 1996.

[61] D. Helbing. *Verkehrsdynamik - neue Physikalische Modellieringskonzepte*. Springer-Verlag, 1997.

[62] D. Helbing, A. Hennecke, V. Shvetsov, and M. Treiber. MASTER: Macroscopic traffic simulation based on a gas-kinetic, non-local traffic model. *Transportation Research B*, 35(2):183–211, February 2001.

[63] D. Helbing, A. Hennecke, V. Shvetsov, and M. Treiber. Micro- and macrosimulation of freeway traffic. *Mathematical and Computer Modelling*, 35(5–6):517–547, March 2002.

[64] B. Hellinga and M. Van Aerde. Examining the potential of using ramp metering as a component of an ATMS. *Transportation Research Record*, 1494:75–83, 1995.

[65] F.-S. Ho and P. Ioannou. Traffic flow modeling and control using artificial neural networks. *IEEE Control Systems*, pages 16–26, 1996.

[66] S. P. Hoogendoorn and P.H.L. Bovy. Modelling multiple user-class traffic flow. *Transportation Research B*, 34(2):123–146, 2000.

[67] S.P. Hoogendoorn. *Multiclass Continuum Modelling of Multiclass Traffic Flow*. PhD thesis, TRAIL Thesis Series, December 1999. Delft.

[68] S.P. Hoogendoorn and P.H.L. Bovy. State-of-the-art of vehicular traffic flow modelling. Technical report, Transportation and Traffic Engineering Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands. Accepted for publication in the

Special Issue on Road Traffic Modelling and Control of the Journal of Systems and Control Engineering – Proceedings of the Institution of Mechanical Engineers, Part I, 2000.

[69] S.P. Hoogendoorn and P.H.L. Bovy. Generic gas-kinetic traffic systems modeling with applications to vehicular traffic flow. *Transportation Research B*, 35(4):317–336, May 2001.

[70] Ben Immers and Bart Egeter. OWN als redmiddel tegen verkeersinfarct. *Verkeerskunde*, 10(2):18–25, 2002. In Dutch.

[71] L. Jacobson, K. Henry, and O. Mehyar. Real-time metering algorithm for centralized control. *Transportation Research Record*, 1232:17–26, 1989.

[72] B. Johnson. Keeping the world flowing. *Traffic Technology International*, pages 38–40, June-July 1999.

[73] P. Kachroo and K. Özbay. *Feedback Control Theory for Dynamic Traffic Assignment*. Advances in Industrial Control. Springer, 1999.

[74] K. G. Keen, M. J. Schofield, and G. C. Hay. Ramp metering access control on M6 motorway. In *Proceedings of the IEE 2nd International Conference on Road Traffic Control*, London, 1986.

[75] J. H. Kell, I. J. Fullerton, and M. K. Mills. *Traffic Detector Handbook, Second Edition*. Number FHWA-IP- 90-002. July 1990.

[76] E. Kenis and R. Tegenbos. Ramp metering synthesis. Technical report, CENTRICO, DG TREN European Commission, Dec 2001. http://www.itsproj.com/centrico/publications.html.

[77] L.A. Klein, M.R. Kelley, and M.K. Mills. Evaluation of overhead and in-ground vehicle detector technologies for traffic flow measurement,. *Journal of Testing and Evaluation (JTEVA)*, 25(2):215–224, March 1997.

[78] A. Kotsialos, M. Papageorgiou, H. Haj-Salem, S. Manfredi, J. van Schuppen, J. Taylor, and M. Westerman. Coordinated control strategies, 1997. Deliverable D06.1 of the DACCORD project (TR1017), European Comission, Brussels, Belgium.

[79] A. Kotsialos, M. Papageorgiou, M. Mangeas, and H. Haj-Salem. Coordinated and integrated control of motorway networks via non-linear optimal control. *Transportation Research C*, 10(1):65–84, February 2002.

[80] A. Kotsialos, M. Papageorgiou, and A. Meßmer. Integrated optimal control of motorway traffic networks. In *Proceedings of the 1999 American Control Conference (ACC'99)*, pages 2183–2187, San Diego, California, June 1999.

[81] A. Kotsialos, M. Papageorgiou, and A. Messmer. Optimal co-ordinated and integrated control motorway traffic control. In A. Ceder, editor, *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, 1999.

[82] B. Kouvaritakis, M. Cannon, and J.A. Rossiter. Non-linear model based predictive control. *International Journal of Control*, 72(10):919–928, 1999.

[83] J.P. Lebacque. The Godunov scheme and what it means for first order traffic flow models. In J.B. Lesort, editor, *Proceedings of the 13th International Symposium of Transportation and Traffic Theory*, pages 647–677, Lyon, France, 1996.

[84] B. Li and B. De Moor. Recursive optimization for intersection equality-constrained optimization for intersection origin-destination matrices. *Transportation Research B*, 33B(3):203–214, April 1999.

[85] B. Li and B. De Moor. Dynamic identification of origin-destination matrices in the presence of incomplete observations. *Transportation Research B*, 36B(1):37–57, 2001.

[86] M. Lighthill and G. Whitham. On kinematic waves I. Flood movement in long rivers. In *Proceedings of the Royal Society*, volume 229A, pages 281–316, London, May 1955.

[87] M. Lighthill and G. Whitham. On kinematic waves II. A theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society*, volume 229A, pages 317–345, London, May 1955.

[88] G. Liu, A.S. Lyrintzis, and P.G. Michalopoulos. Improved high-order model for freeway traffic flow. *Transportation Research Record*, 1644:37–46, 1998.

[89] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River, New Jersey, 2nd edition, 1999.

[90] S. Logghe and B. Immers. Representation of traffic shock waves in micro-simulation models. In *Proceedings of the 7th World congress on ITS*, Turin, November 2000.

[91] S. Logghe and B. Immers. Evaluatie effecten DVM-maatregelen met behulp van microsimulatie. In *Proceedings van het 4de DVM symposium*, Rotterdam, April 2001. In Dutch.

[92] S. Logghe and B. Immers. Micro-simulation of congested traffic flows. In *Proceedings of the Seminar on Advances and Applications in Dynamic Traffic Flow Modeling*, Delft, The Netherlands, February 2001.

[93] F. Logi and S. G. Ritchie. A multi-agent architecture for cooperative inter-jurisdictional traffic congestion management. *Transportation Research C*, 10(5):507–527, 2002.

[94] J.M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, Harlow, England, 2002.

[95] A. D. May. *Traffic flow fundamentals*. Prentice Hall, Englewood Cliffs, NJ,, 1990.

[96] A. Messmer and M. Papageorgiou. METANET: a macroscopic simulation program for motorway networks. *Traffic Engineering and Control*, 31(9):466–470, 1990.

[97] A. Messmer and M. Papageorgiou. Automatic control methods applied to freeway network traffic. *Automatica*, 30(4):691–702, 1994.

[98] F. Middelham. A synthetic study of the network effects of ramp metering. In *Proceedings of the Seminar Advanced Traffic Control*, Delft, The Netherlands, March 2001. TRAIL Research School.

[99] F. Middelham and H. Taale. The application of FLEXSYT in traffic management systems. In *Proceedings of the Managing Traffic and Transportation Conference*, pages 158–167, Amsterdam, The Netherlands, April 1992.

[100] D. Middleton, D. Gopalakrishna, and M. Raman. Advances in traffic data collection and management. Technical Report BAT-02-006, Texas Transportation Institute, Cambridge Systematics, Inc., Washington, DC, December 2002. http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/13766.html.

[101] K. Nagel. Particle hopping models and traffic flow theory. *Physical review*, E(53):4655–4672, 1996.

[102] K. Nagel. From particle hopping models to traffic flow theory. *Transportation Research Record*, 1644:1–9, 1998.

[103] G. Newell. A simplified theory of kinematic waves in highway traffic, Part I: General theory. *Transportation Research B*, 27B(4):281–287, 1993.

[104] G. Newell. A simplified theory of kinematic waves in highway traffic, Part II: Queueing at freeway bottlenecks. *Transportation Research B*, 27B(4):289–303, 1993.

[105] J. Nocedal and S. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, New York, 1999.

[106] M. Papageorgiou. *Applications of Automatic Control Concepts to Traffic Flow Modeling and Control*. Lecture Notes in Control and Information Sciences. Springer Verlag, Berlin, Germany, 1983.

[107] M. Papageorgiou. Some remarks on macroscopic traffic flow modelling. *Transportation Research A*, 32(5):323–329, 1998.

[108] M. Papageorgiou. *Handbook of Transportation Science*, chapter 8. Traffic Control, pages 233–267. Kluwer Academic Pulishers, Boston, 1999.

[109] M. Papageorgiou, J. M. Blosseville, and H. Hadj-Salem. Macroscopic modelling of traffic flow on the boulevard périphérique in Paris. *Transportation Research B*, 23(1):29–47, 1989.

[110] M. Papageorgiou, J.-M. Blosseville, and H. Hadj-Salem. Modelling and real-time control of traffic flow on the southern part of boulevard périphérique in Paris: Part I: Modelling. *Transportation Research A*, 24A(5):345–359, 1990.

[111] M. Papageorgiou, J.M. Blosseville, and H. Hadj-Salem. Modelling and real-time control of traffic flow on the southern part of boulevard périphérique in Paris: Part II: Coordinated on-ramp metering. *Transportation Research A*, 24A(5):361–370, 1990.

[112] M. Papageorgiou, H. Hadj-Salem, and J.-M. Blosseville. ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record*, 1320:58–64, 1991.

[113] M. Papageorgiou, H. Hadj-Salem, and F. Middelham. ALINEA local ramp metering: Summary of field results. *Transportation Research Record*, 1603:90–98, 1997.

[114] M. Papageorgiou, B. Posh, and G. Schmidt. Comparison of macroscopic models for control of freeway traffic. *Transportation Research B*, 17B(2):107–116, 1983.

[115] P.M. Pardalos and M.G.C. Resende, editors. *Handbook of Applied Optimization*. Oxford University Press, Oxford, UK, 2002.

[116] S.L. Paveri-Fontana. On Boltzmann-like treatments for traffic flow: a critical review of the basic model and an alternative proposal for dilute traffic analysis. *Transportation Research B*, 9:225–235, 1975.

[117] H. J. Payne. Models of freeway traffic and control. In G. A. Bekey, editor, *Mathematical Models of Public Systems*, volume 1 of *Simulation Council Proceedings Series*, pages 51–61, La Jolia, California, 1971.

[118] L. A. Pipes. Car-following models and the fundamental diagram of road traffic. *Transportation Resarch*, 1(1):21–29, 1967.

[119] W. B. Powell and Y. Sheffi. The convergence of equilibrium algorithms with predetermined step sizes. *Transportation science*, 16(1):45–55, January 1982.

[120] I. Prigogine and R. Herman. *Kinetic Theory of Vehicular Traffic*. American Elsevier, New-York, 1971.

[121] S. J. Qin and T. A. Badgwell. An overview of industrial model predictive control technology. In J. C. Kantor, C. E. Garcia, and B. Carnahan, editors, *Fifth International Conference on Chemical Process Control*, AIChE Symposium Series, pages 232–256, 1997.

[122] J. Richalet, A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14(5):413–428, September 1978.

[123] P. I. Richards. Shockwaves on the highway. *Operations Research*, 4:42–51, 1956.

[124] J. Robinson and M. Doctor. Ramp metering status in North America - Final Report. Technical Report DOT-T-90-01, Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., Sep. 1989.

[125] R. Schleiffer. Intelligent agents in traffic and transportation. *Transportation Research C*, 10(5):325–329, 2002.

[126] Y. Sheffi. *Urban transportation networks*. Prentice-Hall, Englewood Cliffs, N.J., 1985.

[127] B. L. Smith. Forecasting freeway traffic flow for intelligent transportation system applications. Master's thesis, Department of Civil Engineering, University of Virginia, Charlottesville, USA, 1995.

[128] B. L. Smith, B. M. Williams, and R. K. Oswald. Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research C*, 10(4):303–321, August 2002.

[129] S. Smulders. Control of freeway traffic flow by variable speed signs. *Transportation Research B*, 24(2):111–132, 1990.

[130] S. A. Smulders. *Control of Freeway Traffic Flow*. PhD thesis, University of Twente, Enschede, 1989.

[131] A.R.M. Soeterboek. *Predictive Control – A Unified Approach*. Prentice Hall, Englewood Cliffs, New Jersey, 1992.

[132] H. Taale and H.J. van Zuylen. Traffic control and route choice: Occurrence of instabilities. In *TRAIL 5th Annual Congress 1999: Five Years Crossroads of Theory and Practice*, volume 2, pages 1–19. Trail, 1999.

[133] Technical University of Crete, Dynamic Systems and Simulation Laboratory and A. Messmer. *METANET – A simulation program for motorway networks*, July 2000.

[134] The Mathworks, Natick, Massachusetts. *Optimization toolbox user's guide v. 2.2*, 2002.

[135] M. Treiber and D. Helbing. Microsimulations of freeway traffic including control measures. *Automatisierungstechnik*, 49:478–484, 2001.

[136] Twin cities ramp meter evaluation — Executive summary. Cambridge Systematics, Inc., February 2001.

[137] R. T. Underwood. Speed, volume, and density relationships, quality and theory of traffic flow. Technical report, Yale Bureau of Highway Traffic, New Haven, Connecticut, 1961.

[138] N. J. van der Zijpp. *Estimation of origin-destination tables*. PhD thesis, Delft University of Technology, Delft, 1995.

[139] R. van Katwijk and P. van Koningsbruggen. Coordination of traffic management instruments using agent technology. *Transportation Research C*, 10(5):455–471, 2002.

[140] A. van Loon and L. Duynstee. Intelligent Speed Adaptation (ISA): A succesfull test in the Netherlands. In *Proceedings of the Canadian Multidisciplinary Road Safety Conference XII*, Ontario, Canada, June 2001. Ministry of Transport, Transport Research Centre (AVV).

[141] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer Academic Publishers, 1996.

[142] J. H. van Schuppen and A. G. Steenbeek. Variable route directives for motorway networks. In *DACCORD workshop on Advanced Motorway Traffic Control*, pages 1–9, UK, December 1998. Lancaster University.

[143] H.J. Van Zuylen and L.G. Willumson. The most likely trip matrix estimated from traffic counts. *Transportation Research B*, 14B(3):281–293, 1980.

[144] D. Van Vliet. The Frank-Wolfe algorithm for equilibrium traffic asignment viewed as a variational inequality. *Transportation Research B*, 21(1):87–89, 1987.

[145] Y. Wang and M. Papageorgiou. A predictive feedback routing control strategy for freeway network traffic. In *Proceedings of the American Control Conference*, pages 3606–3611, Anchorage, May 2002. IEEE.

[146] Y. Wang, M. Papageorgiou, and A. Messmer. A predictive feedback routing control strategy for freeway network traffic. In *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*, Washington, D.C, January 2003.

[147] J. Wardrop. Some theoretical aspects of road traffic research. pages 326–378. Road engineering division, Jan 1952.

[148] A. Waters. Theory and measurement of private and social cost of highway congestion. *Econometrica*, 29:676–699, 1961.

[149] J. A. Wattleworth. Peak-period analysis and control of a freeway system. *Highway Research Record*, 157:1–21, 1965.

[150] G. B. Whitham. Linear and nonlinear waves. In *Pure and Applied Math.* Wiley Interscience Monographs and Tracts, New York, 1974.

[151] B. M. Williams. *Modeling and Forecasting Vehicular Traffic Flow as a Seasonal Stochastic Time Series Process.* PhD thesis, Department of Civil Engineering, University of Virginia, Charlottesville, USA, 1999.

[152] W. L. Winston. *Operations Research: Applications and Algorithms.* Duxbury, 1993.

[153] M. B. Wisten and M. J. Smith. Distributed computation of dynamic traffic equilibria. *Transportation Research C*, 5:77–93, 1997.

[154] N. Wu and W. Brilon. Cellular automata for highway traffic flow simulation. In Ceder A., editor, *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, pages 1–18, 1999.

[155] Q. Yang. *A Simulation Laboratory for Evaluation of Dynamic Traffic Management Systems.* PhD thesis, Massachusetts Institute of Technology, 1997.

[156] J. T. Yuan. A new interpolation lattice structure. In *Proceedings of the 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 448–451, Minneapolis, MN, USA, Apr. 1993.

[157] H. Zhang and S. G. Ritchie. Freeway ramp metering using artificial neural networks. *Transportation Research C*, 5(5):273–286, 1997.

[158] H. Zhang, S. G. Ritchie, and Z. Lo. Macroscopic modeling of freeway traffic using an artificial neural network. *Transportation Research Record*, 1588:110–119, 1997.

[159] H. Zhang, S.G. Ritchie, and W.W. Recker. Some general results on the optimal ramp control problem. *Transportation Research C*, 4(2):51–69, 1996.

[160] H. M. Zhang. A mathematical theory of traffic hysteresis. *Transportation Research B*, 33:1–23, 1999.

[161] H. M. Zhang. Structural properties of solutions arising from a nonequilibrium traffic flow theory. *Transportation Research B*, 34:583–603, 2000.

[162] H. M. Zhang. On the consistency of a class of traffic flow models. *Transportation Research B*, 37:101–105, 2003.

Tom Bellemans was born on March 28, 1975 in Leuven, Belgium. He received the degree of electrotechnical engineering in control theory from the Katholieke Universiteit Leuven, Belgium in July 1998. Since then he has been working as a researcher at the Electrical Engineering Department of the Katholieke Universiteit Leuven on a DWTC (Federal Office for Scientific, Technical and Cultural Affairs) grant.