

Betrouwbaarheid van MoCo-data voor het bepalen van intensiteiten voor het verkeersmodel

Door: Daniel Lie

Mei 2012



Begeleiders VU
dr. Sandjai Bhulai
prof. dr. Maarten van Steen

Gemeente Amsterdam



Begeleider DIVV
ir. Jeroen Schrijver

Betrouwbaarheid van MoCo-data voor het bepalen van intensiteiten voor het verkeersmodel

Daniel Lie

Stageverslag

Vrije Universiteit Amsterdam
Faculteit der Exacte Wetenschappen
Studierichting Bedrijfswiskunde en Informatica
De Boelelaan 1081a
1081 HV Amsterdam

Dienst Infrastructuur Verkeer en Vervoer
Nieuwevaart 5 - 9
1018 AA Amsterdam

April 2012

Voorwoord

De masteropleiding Business Mathematics and Informatics aan de Vrije Universiteit (VU) wordt afgesloten met een afstudeerstage. In deze afstudeerstage past de student zijn kennis die hij tijdens zijn opleiding heeft opgedaan in de praktijk toe bij een organisatie.

Ik heb mijn stage gelopen bij de Dienst Infrastructuur Verkeer en Vervoer (DIVV) van de gemeente Amsterdam. De stageperiode was voor mij zeer leerzaam en ik heb een goed beeld gekregen van wat er komt kijken bij het werken in een organisatie. In een organisatie sta je niet alleen. Ik heb geleerd hoe je de juiste mensen kan vinden die de informatie hebben die je nodig hebt. Door de behulpzame sfeer die heerst binnen Dienst Infrastructuur Verkeer en Vervoer was dit goed mogelijk.

Er zijn een aantal mensen die ik wil bedanken. Ten eerste wil ik Jeroen Schrijver van DIVV bedanken voor zijn uitstekende begeleiding. Elke twee weken maakte hij een uur voor mij vrij om mij te voorzien van adviezen en nieuwe inzichten.

Dank gaat ook uit naar Sandjai Bhulai van de VU voor zijn scherpe feedback tijdens onze besprekingen. Altijd wist hij iets toe te voegen of beter te maken aan het werk dat ik op dat moment presenteerde.

Dan wil ik Nelly Kalfs van DIVV bedanken voor het beschikbaar stellen van deze stageplek. Ook wil ik Maarten van Steen van de VU bedanken voor zijn enthousiaste bijdrage aan dit project. Zijn nieuwe ideeën bleken altijd zeer waardevol.

Deze stageplek was niet mogelijk zonder het project ‘Vernieuwingstraject afdeling Verkeersonderzoek van Fred Harms’. Daarom gaat mijn dank ook uit naar hem. Tenslotte wil ik al mijn collega’s van DIVV bedanken voor de gezellige afstudeerperiode.

Amsterdam, April 2012

Management Samenvatting

MoCo is het Amsterdamse systeem om reistijden te meten en is opgezet door Vialis. Verspreid over de stad staan er camera's die kentekens kunnen herkennen van voertuigen. Wanneer een voertuig langs een camera rijdt, wordt het kenteken met een versleuteling geregistreerd. Tevens wordt ook het tijdstip waarop dit kenteken is herkend geregistreerd. Wanneer een voertuig langs meerdere camera's rijdt is de tijd die het voertuig erover heeft gedaan om van de ene camera naar de andere camera te komen bekend. Zo is dus de reistijd van een voertuig te meten.

Momenteel worden de MoCo-camera's alleen gebruikt om reistijden te meten. Op basis van de data uit de MoCo-camera's zou het ook mogelijk moeten zijn om de intensiteiten van voertuigen te bepalen op de plaatsen waar de MoCo-camera's staan. Een intensiteit is het aantal voertuigen dat in een bepaalde periode een locatie passeert. Deze intensiteiten zijn voor Verkeersonderzoek interessante gegevens. Verkeersonderzoek werkt met het verkeersmodel GenMod. GenMod kan prognoses maken van intensiteiten voor de wegen in Amsterdam. Om te zorgen dat GenMod betrouwbare prognoses blijft geven, wordt GenMod elke vier jaar gekalibreerd op de meest recente verkeerstellingen. Het laten uitvoeren van deze verkeerstellingen kost geld. Wanneer blijkt dat de intensiteiten op basis van de MoCo-data betrouwbaar zijn, kan worden onderzocht of bepaalde verkeerstellingen die worden uitgevoerd kunnen worden vervangen door intensiteiten op basis van MoCo-data. Zo kunnen kosten worden bespaard.

Het bepalen van de intensiteit op basis van MoCo-data is eenvoudig. Door het aantal herkende kentekens in een bepaalde periode bij elkaar op te tellen ontstaat er een intensiteit. De vraag is echter hoe betrouwbaar deze intensiteiten zijn. Het doel van het onderzoek is om te bepalen of deze intensiteiten gebaseerd op MoCo-data betrouwbaar zijn om gebruikt te kunnen worden voor kalibratie van het verkeersmodel.

Er zijn twee punten die de betrouwbaarheid van de intensiteiten op basis van data uit MoCo-camera's beïnvloeden. Ten eerste is dit de betrouwbaarheid van de data. In hoeverre worden alle voertuigen die langs een MoCo-camera rijden geregistreerd? Een ander punt is dat op sommige wegen niet op alle rijstroken een camera staat. Een weg kan uit meerdere rijstroken bestaan. Om de totale intensiteit van een weg te bepalen heb je de intensiteit nodig van alle rijstroken op deze weg. Deze twee punten worden in dit onderzoek behandeld.

We hebben een steekproef genomen van vijf weken. Voor elk van deze vijf weken hebben we op de dinsdag, woensdag en donderdag de gemeten avondspitsintensiteit bekeken voor alle camera's. Bij het onderzoeken van de gemeten intensiteiten was er een opvallend verschil tussen de camera's. Bij een aantal camera's verschilden de gemeten intensiteiten veel van elkaar, en bij andere camera's verschilden de gemeten intensiteiten minder van elkaar. Bij sommige camera's waren er ook redelijk lage intensiteiten gemeten.

Op basis van alleen de gemeten intensiteiten kunnen deze verschillen en lage intensiteiten niet verklaard worden. Enerzijds is het mogelijk dat bij sommige camera's het aantal voertuigen dat werkelijk langs de camera reed ongeveer

zo laag was als gemeten. Anderzijds is het mogelijk dat er aanzienlijk meer voertuigen reden dan gemeten en dat de lage gemeten intensiteiten veroorzaakt zijn doordat de camera's veel van deze voertuigen niet hebben geregistreerd.

Bij Vialis is er echter storingsdata aanwezig. Van elke locatie is bekend van hoeveel procent van de voertuigen het kenteken in een maand is herkend. Ook is bekend wat het percentage van de tijd is dat de camera werkte en data verstuurde naar Vialis. Wanneer deze herkenningpercentages en beschikbaarheidspercentages per uur per camera bekend zouden zijn, dan kan met een grote betrouwbaarheid bepaald worden wat de werkelijke intensiteit was in dat uur. Echter in dit onderzoek waren alleen de gemiddelde herkennings- en beschikbaarheidspercentages per maand aanwezig. Een negatieve correlatie was gevonden tussen de verschillen van gemeten intensiteiten van een bepaalde camera en de herkennings- en beschikbaarheidspercentages. Hoe hoger de herkennings- en beschikbaarheidspercentages waren, hoe lager de verschillen waren tussen de gemeten intensiteiten van een camera.

Dit betekent dat de verschillen in gemeten intensiteiten grotendeels verklaard kunnen worden door onbetrouwbaarheden van de camera. Een betrouwbare intensiteit voor het verkeersmodel mag niet op deze onbetrouwbaarheden gebaseerd zijn. Daarom moeten de gemeten intensiteiten waarbij veel onbetrouwbaarheden van de camera optraden, voor zo veel mogelijk niet meegenomen worden in de intensiteit voor het verkeersmodel. Deze onbetrouwbaarheden treden voornamelijk op bij de laagste gemeten intensiteiten.

Op basis daarvan is bepaald dat voor een betrouwbare intensiteit voor het verkeersmodel een gemiddelde kan worden genomen van een aantal van de hoogste gemeten intensiteiten. Hoe hoog dit exacte aantal moet zijn kon op basis van de aanwezige data niet worden bepaald. Wel kan gesteld worden dat dit aantal niet te hoog mag zijn.

Met behulp van een regressiemodel is geprobeerd om de intensiteit te schatten op locaties waar niet alle rijstroken zijn bemeten. Op locaties met twee rijstroken, waarvan er slechts één bemeten is, staat de camera meestal op de rechter rijstrook. Het doel van het regressiemodel is om de intensiteit op de linker rijstrook te schatten. Dit regressiemodel is alleen van toepassing voor locaties met twee rijstroken, waarvan een rijstrook niet was bemeten. Voor de locaties met meer dan twee rijstroken was er niet genoeg data om een regressiemodel op te stellen.

Het regressiemodel heeft als beperking dat het maar gebaseerd is op de data van camera's op zes locaties. Er zijn twee variabelen gevonden die kunnen voorspellen hoe groot de totale intensiteit op een weg zou zijn als alleen de intensiteit op de rechter rijstrook bekend is. Dit zijn de intensiteit op de rechter rijstrook en de afstand van het meetpunt van de camera tot het vorig kruispunt. Dit regressiemodel verklaarde 80% van de variantie in de dataset.

Vervolgens hebben we bekeken hoe betrouwbaar de intensiteiten worden als we een gemiddelde nemen van een aantal van de hoogste intensiteiten. Ook hebben we bekeken hoe betrouwbaar het regressiemodel is. Dit hebben we kunnen toetsen door de intensiteiten voor het verkeersmodel die uit onze methodes volgen, te vergelijken met andere tellingen op exacte plaatsen van MoCo-camera's.

De resultaten hebben we opgedeeld in locaties waar op alle rijstroken een camera staat en locaties waar niet op alle locaties een camera staat.

Op de wegen waar op alle rijstroken een camera staat lijken de intensiteiten voor het verkeersmodel met de gebruikte methodes erg betrouwbaar. Een gemiddeld verschil van 10% met andere tellingen was gemeten. Er moet rekening mee worden gehouden dat de tellingen waarmee we onze intensiteiten voor het verkeersmodel vergelijken ook slechts steekproeven zijn. We kunnen dus alleen een indicatie geven van de betrouwbaarheid van de intensiteiten voor het verkeersmodel. Voor de locaties waar niet alle rijstroken waren bemeten en dus een voorspelling is gedaan van de linker rijstrook om de intensiteit voor het verkeersmodel te bepalen, bleken de resultaten minder betrouwbaar. Een gemiddeld verschil van 35% met de andere tellingen was nu waargenomen. Hier zaten zelfs verschillen bij van rond de 80%. Op basis hiervan kunnen we concluderen dat het regressiemodel geen betrouwbare voorspellingen geeft. Vermoedelijk was de dataset waarop we het regressiemodel hebben gebaseerd te klein om een betrouwbaar regressiemodel op te stellen.

Toch kunnen we stellen dat de locaties waarbij alle rijstroken bemeten zijn betrouwbaar zijn om intensiteiten te bepalen voor het verkeersmodel. Hiervoor kunnen de intensiteiten bepaald worden door een gemiddelde te nemen van een aantal van de hoogste gemeten intensiteiten. In totaal zijn er 95 locaties waarbij alle rijstroken zijn bemeten die gebruikt kunnen worden voor het verkeersmodel.

Inhoudsopgave

1	Inleiding	11
1.1	Over Dienst Infrastructuur Verkeer en Vervoer	11
1.2	Probleemstelling	11
1.2.1	Intensiteiten	11
1.2.2	Intensiteiten en MoCo-camera's	11
1.2.3	Onderzoeksvraag	12
1.3	Motivatie en doelstelling	12
1.4	Aanpak	13
1.5	Indeling van het verslag	13
2	Wat is MoCo?	15
2.1	Doel	15
2.2	Werking	15
2.3	Architectuur	15
2.4	Toevoeging milieuzonecamera's	17
2.5	Meetlocaties	18
3	Dataset	20
3.1	Dataverzameling	20
3.1.1	Kentekenregistraties	20
3.1.2	Betrouwbaarheidsdata	21
3.2	Beschrijving metadata	22
3.2.1	Kentekenregistraties	22
3.2.2	Betrouwbaarheidsdata	23
3.3	Verkennde Data-analyse	23
3.3.1	Kentekenregistraties	23
3.3.2	Storingsdata	36
4	Filtermethode	38
4.1	Notatie	38
4.2	Schema van kentekenherkenning	39
4.3	Analyse van betrouwbaarheid camera's	40
4.3.1	Onderzoek Rotterdam	40
4.3.2	Betrouwbaarheidsdata	44
4.4	Methode	50
4.4.1	Definitie $q_m(l)$	50
4.4.2	Bepalen $q_m(l)$	50
5	Voorspelling meerdere rijstroken	54
5.1	$q_{m,rechts}(l)$ en p	56
5.2	Afstand vorige kruispunt en verhouding	60
5.3	Meer dan twee rijstroken	64

6 Resultaten	65
6.1 Alle rijstroken bemeten	66
6.2 Twee rijstroken, waarvan een bemeten	71
7 Conclusie en aanbevelingen	77
7.1 Deelvraag 1	77
7.2 Deelvraag 2	78
7.3 Aanbevelingen	79
7.4 Algemene aanbevelingen	79
A Alternatieve filtermethoden	82
A.1 Volgtijden	82
A.2 Relaties tussen camera's	82
B Uitdagingen tijdens het onderzoek	84
B.1 betrouwbaarheid van coördinaten van meetlocaties	84
B.2 bepaling van het aantal rijstroken	84
B.3 het verwerken van kentekenregistraties tot gemeten intensiteiten	85
B.4 interpretatie van betrouwbaarheid van gemeten intensiteiten . . .	85
C Meetlocaties	86

1 Inleiding

1.1 Over Dienst Infrastructuur Verkeer en Vervoer

Dienst Infrastructuur Verkeer en Vervoer (DIVV) bedenkt, ontwikkelt, onderhoudt en verbetert de verkeersinfrastructuur van Amsterdam. DIVV is de organisatie die verantwoordelijk is voor [1]:

- een veilige doorstroom van alle verkeer in en naar de stad te bewerkstelligen
- de beschikbaarheid van de infrastructuur te waarborgen
- het toedelen en beprijzen van de schaarse ruimte en de stimulering van het gebruik van fiets en openbaar vervoer

De afdeling Verkeersonderzoek (VO) binnen DIVV doet evaluatiestudies naar de mobiliteit in Amsterdam. Daarnaast is VO verantwoordelijk voor het leveren van verkeerscijfers.

1.2 Probleemstelling

1.2.1 Intensiteiten

Een van de verkeerscijfers die geleverd worden zijn prognoses van intensiteiten. De intensiteit is het aantal gepasseerde voertuigen op een bepaalde plaats in een bepaalde periode. Voor het maken van intensiteitsprognoses wordt gebruik gemaakt van het verkeersmodel GenMod. Dit model maakt intensiteitsprognoses voor wegen in Amsterdam.

Met behulp van intensiteitsprognoses kunnen bijvoorbeeld de effecten van verandering in beleid of infrastructuur worden doorgerekend. De intensiteitsprognoses ondersteunen derden in het maken van beslissingen.

Om te zorgen dat de intensiteitsprognoses accuraat zijn, wordt GenMod elke twee jaar bijgewerkt met de meest recente invoer. Dit wordt het actualiseren van het model genoemd. Daarnaast wordt GenMod elke vier jaar tijdens de modelactualisatie gekalibreerd op de meest recente verkeerstellingen [2]. VO laat elke twee jaar verkeerstellingen uitvoeren.

1.2.2 Intensiteiten en MoCo-camera's

DIVV beschikt over meetinstrumenten in de stad die gebruikt worden voor dynamisch verkeersmanagement. Onder deze meetinstrumenten bevinden zich MoCo-camera's. Op basis van data die MoCo-camera's genereren kunnen mogelijk intensiteiten worden geschat.

MoCo-cameras registreren namelijk kentekens van voertuigen. Het aantal geregistreerde kentekens op een plaats over een bepaalde periode geeft een schatting van hoeveel voertuigen over die periode werkelijk die plaats zijn gepasseerd. Het is een schatting omdat het aantal geregistreerde kentekens niet overeen hoeft te komen met het aantal werkelijk gepasseerde voertuigen.

Een weg kan uit meerdere rijstroken bestaan. Er zijn wegen met meerdere rijstroken waar niet een MoCo-camera staat op elke rijstrook. Voor het bepalen van de intensiteit op deze wegen op basis van data uit MoCo-camera's betekent dit dat een deel van de informatie op deze wegen mist.

Verkeersonderzoek wil weten in hoeverre geschatte intensiteiten op basis van data uit MoCo-camera's bruikbaar zijn voor kalibratie van het verkeersmodel.

Er zijn twee punten die de kwaliteit van intensiteiten op basis van data uit MoCo-camera's beïnvloeden:

1. Betrouwbaarheid van de data
2. Ontbreken van MoCo-camera's op sommige rijstroken

1.2.3 Onderzoeksvraag

In dit onderzoek wordt de volgende onderzoeksvraag behandeld:

Onderzoeksvraag In hoeverre zijn intensiteiten gebaseerd op data uit MoCo-camera's betrouwbaar om gebruikt te worden voor kalibratie van het verkeersmodel?

Bij het beantwoorden van de vraag ligt de focus niet op de kalibratie van het verkeersmodel. Dit is het doel waar de intensiteiten gebaseerd op data uit MoCo-camera's gebruikt voor gaan worden. Daarom zullen we bij het beoordelen van de betrouwbaarheid van de intensiteiten gebaseerd op data uit MoCo-camera's er specifiek op letten of de betrouwbaarheid hoog genoeg is voor kalibratie van het verkeersmodel. Voor het beantwoorden van de onderzoeksvraag worden de volgende deelvragen gesteld:

Deelvraag 1. Hoe betrouwbaar zijn data uit MoCo-camera's en hoe kunnen eventuele fouten in deze data eruit worden gefilterd?

Deelvraag 2. Hoe kunnen intensiteiten worden geschat op wegen waar niet alle rijstroken bemeten zijn?

1.3 Motivatie en doelstelling

In de studie 'vernieuwingstraject afdeling Verkeersonderzoek' kwam naar voren dat "veel bij DIVV in potentie aanwezige databronnen niet of in voldoende mate worden gebruikt." [3]. Een van deze databronnen is de MoCo-camera's. Naar aanleiding van deze studie is het idee ontstaan om dit onderzoek uit te voeren.

Het doel van het onderzoek is om advies uit te brengen over de betrouwbaarheid van intensiteiten op basis van data uit MoCo-camera's voor het verkeersmodel. Het laten uitvoeren van verkeerstellingen kost geld. Inzicht over in welke mate gebruik gemaakt kan worden van intensiteiten op basis van data uit MoCo-camera's stelt Verkeersonderzoek bijvoorbeeld in staat de beslissing te maken om tijdens haar tweejaarlijks telplan op minder punten te tellen, doordat deze telpunten vervangen kunnen worden door MoCo-camera's. Zo kunnen kosten worden bespaard.

1.4 Aanpak

Op de afdeling Verkeerssystemen (VKS) binnen DIVV staat een server met data uit MoCo-camera's. Deze data zal worden bewerkt tot intensiteiten. We zullen een verkennende data-analyse toepassen op deze intensiteiten. Hierbij bekijken we bijvoorbeeld welke dag- en weektrends de intensiteiten laten zien. Vervolgens gaan we onderzoek doen naar de betrouwbaarheid van de intensiteiten. Hierbij maken we gebruik van betrouwbaarheidsdata die wij op hebben gevraagd bij Vialis, de leverancier van MoCo-camera's. We onderzoeken welke uitspraken we kunnen doen over de intensiteiten en hoe we de betrouwbaarheidsdata kunnen toepassen om onze dataset betrouwbaarder te maken. Naast het onderzoeken van betrouwbaarheidsdata bekijken we ook een onderzoek dat door Vialis is uitgevoerd. Dit is een onderzoek dat is uitgevoerd in Rotterdam, maar hier worden dezelfde camera's gebruikt. We onderzoeken hoe we de resultaten uit dit onderzoek kunnen gebruiken.

Daarna stellen we een methode samen om op basis van de data uit MoCo-camera's betrouwbare intensiteiten te bepalen die gebruikt kunnen worden voor kalibratie van het verkeersmodel. Ook onderzoeken we hoe intensiteiten bepaald kunnen worden op wegen waar niet alle rijstroken zijn bemeten. We bekijken welke factoren een rol spelen in het voorspellen van deze intensiteiten.

De methode om betrouwbare intensiteiten te bepalen en de methode om de intensiteiten te bepalen op wegen waar niet alle rijstroken zijn bemeten passen we toe op de dataset. Deze resultaten vergelijken we met eerder gemeten intensiteiten. Er zijn niet op alle locaties waar MoCo-cameras hangen intensiteitsmetingen in het verleden uitgevoerd, maar op de locaties waar dit wel is gedaan kunnen we een vergelijking maken.

Tot slot zullen deze resultaten worden geïnterpreteerd en zullen er conclusies worden getrokken. Op basis van deze conclusies zullen adviezen worden opgesteld.

1.5 Indeling van het verslag

De indeling van het verslag is als volgt:

- In hoofdstuk 2 wordt omschreven wat het MoCo-systeem is en wat het doel van dit systeem is.
- Hoofdstuk 3 gaat over de dataset. Er wordt beschreven hoe de data kan worden verzameld en hoe de metadata eruit ziet. Ook zal een verkennende data-analyse worden uitgevoerd.
- In hoofdstuk 4 wordt de methode beschreven om de data te filteren. In dit hoofdstuk zal eerst een betrouwbaarheidsanalyse worden uitgevoerd. Deze analyse bestaat uit het analyseren van de betrouwbaarheidsdata, en het bekijken van de betrouwbaarheidsstudie van Vialis dat in Rotterdam is uitgevoerd. Op basis van de bevindingen van de betrouwbaarheidsanalyse zal de methode om data te filteren worden opgebouwd.

- Hoofdstuk 5 beschrijft een methode om intensiteiten te schatten op plaatsen waar niet alle rijstroken bemeten zijn. In dit hoofdstuk zal worden beschreven welke factoren een rol spelen in het schatten van deze intensiteiten.
- In hoofdstuk 6 worden de resultaten bekeken van het toepassen van de methodes genoemd in hoofdstuk 4 en hoofdstuk 5. De intensiteiten verkregen na het toepassen van deze methodes zullen worden vergeleken met tellingen uitgevoerd in het verleden.
- Hoofdstuk 7 geeft conclusies van het onderzoek. De deelvragen zullen beantwoord worden en adviezen zullen worden opgesteld.

2 Wat is MoCo?

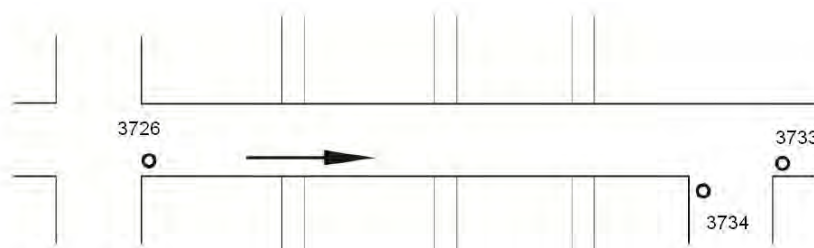
In dit hoofdstuk wordt beschreven wat het MoCo-systeem is. We beginnen met het uitleggen van het doel van het MoCo-systeem. Vervolgens wordt uitgelegd hoe het MoCo-systeem werkt. Hierna komt de architectuur van het MoCo-systeem aan bod. De toevoeging van camera's van de leverancier Gatsometer aan het MoCo-systeem wordt hierna besproken. Het hoofdstuk wordt beëindigd met uitleg over de locaties van de MoCo-camera's.

2.1 Doel

MoCo staat voor Monitoring Corridors. Het MoCo-systeem is door Vialis opgezet in opdracht van DIVV. Het doel van het systeem is om realtime informatie in te winnen over reistijden in Amsterdam. Oorspronkelijk werden alleen de corridors bemeten. Dit zijn de wegen met een grotere verwerkingscapaciteit dan andere wegen en dienen als hoofdontsluiting van de stad [6]. In de loop van de tijd is het MoCo-systeem uitgebreid waardoor het systeem nu informatie inwint over bijna alle S-wegen in Amsterdam. Alleen op de s117 staat er geen MoCo-camera.

2.2 Werking

De werking van het MoCo-systeem is eenvoudig. In figuur 1 is een voorbeeld gegeven van een cameraopstelling. Wanneer een voertuig langs de MoCo-camera op meetlocatie 3726 rijdt, wordt zijn kenteken geregistreerd door een License Plate Recognition camera (LPR-camera). Het voertuig passeert vervolgens meetlocatie 3733 of meetlocatie 3734 waar het kenteken wederom wordt geregistreerd. Omdat op de meetlocaties zowel het kenteken als het tijdstip van herkenning worden geregistreerd kan de reistijd tussen twee meetlocaties worden berekend.



Figuur 1: Werking van het MoCo-systeem; bron: [7].

2.3 Architectuur

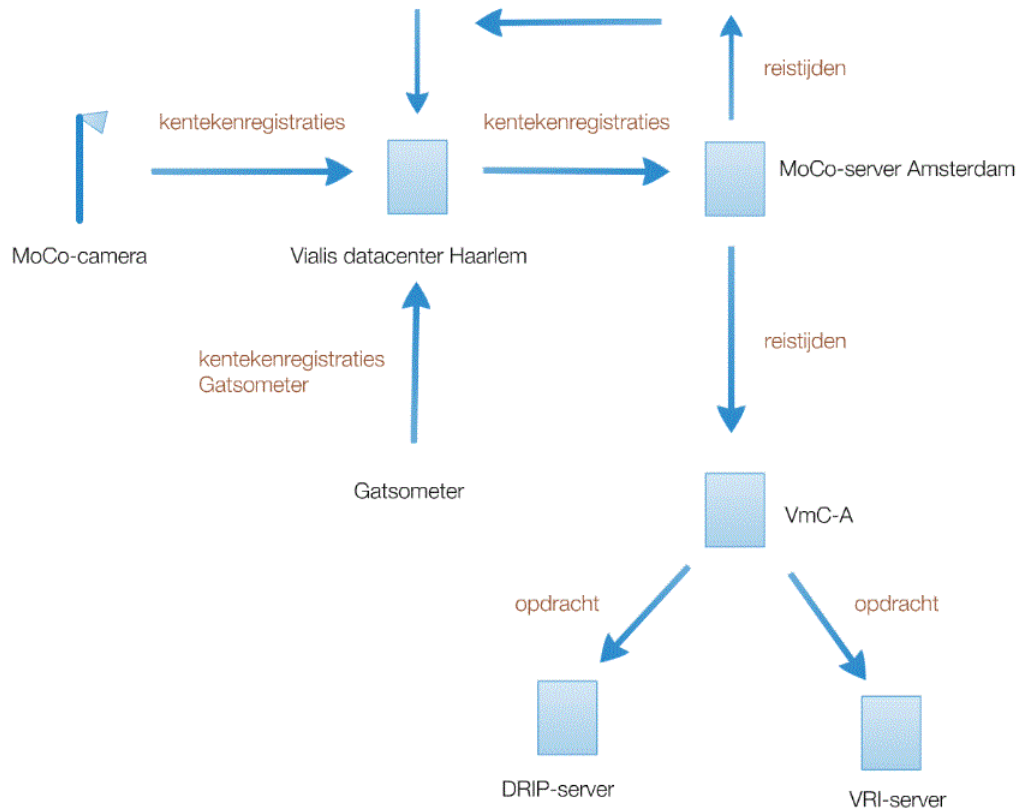
In figuur 2 is de architectuur van het MoCo-systeem weergegeven. Een LPR-camera maakt beelden van een rijstrook. Wanneer de software behorende bij

de LPR-camera vermoedt dat een voertuig met een kenteken langsrijdt, wordt geprobeerd dit kenteken te herkennen. Indien een kenteken wordt herkend, wordt deze met een versleuteling geregistreerd.

Elke minuut wordt een informatiepakket verstuurd van de LPR-camera naar het datacentrum van Vialis in Haarlem. Dit informatiepakket bevat versleutelde kentekens met daarbij behorende tijdstippen van registratie. Vanuit het datacentrum van Vialis in Haarlem worden de versleutelde kentekens met tijdstippen doorgestuurd naar de MoCo-server in Amsterdam. Hier worden de reistijden uitgerekend. De reistijden worden vervolgens doorgestuurd naar het datacentrum van Vialis.

Ook worden de reistijden doorgestuurd naar het Verkeersmanagement Centrale voor de gemeente Amsterdam (VmC-A). VmC-A gebruikt de reistijden voor dynamisch verkeersmanagement. VmC-A stuurt informatie naar onder andere de DRIP-server (dynamische route-informatiepanelen) en de VRI-server (Verkeersregelingsinstallatie). Deze sturen de DRIP's en de VRI's in de stad aan.

Op een deel van de wegen waarover verkeer de stad binnenkomt staan LPR-camera's die zijn neergezet om het vrachtverkeer dat de stad binnenkomt te registreren. Aan de hand van bepaalde criteria wordt bepaald of het toegestaan is voor de vrachtauto om de stad binnen te gaan. Deze camera's worden milieuzonecamera's genoemd. De milieuzonecamera's zijn in het beheer van het bedrijf Gatsometer. Oorspronkelijk maakten de milieuzonecamera's geen deel uit van het MoCo-systeem. Omdat de milieuzonecamera's in staat zijn om alle kentekens te registreren zijn deze camera's ook opgenomen in het MoCo-systeem. De data uit deze milieucamera's komt van Gatsometer het datacentrum van Vialis binnen.



Figuur 2: Architectuur van MoCo-systeem.

2.4 Toevoeging milieuzonecamera's

Wanneer in dit verslag MoCo-camera's genoemd worden, betreft dit alle LPR-camera's in het MoCo-systeem: De LPR-camera's van Vialis en de LPR-camera's van Gatsometer. Indien er toch onderscheid moet worden gemaakt tussen deze camera's, dan wordt dit expliciet aangegeven.

Er zijn drie grote verschillen tussen de MoCo-camera's van Vialis en de milieuzonecamera's:

1. MoCo-camera's van Vialis registreren het kenteken aan de achterkant van een voertuig; milieuzonecamera's registreren het kenteken aan de voorkant van een voertuig;
2. MoCo-camera's van Vialis en milieuzonecamera's hebben een andere locatieaanduiding;

3. MoCo-camera's van Vialis worden beheerd en onderhouden door Vialis; milieuzonecamera's worden beheerd en onderhouden door Gatsometer.

In figuur 3 en figuur 4 zijn respectievelijk een MoCo-camera en een milieuzonecamera weergegeven.



Figuur 3: MoCo-camera van Vialis.



Figuur 4: Milieuzonecamera.

2.5 Meetlocaties

Er zijn verschillende meetlocaties in Amsterdam. Meetlocaties worden als volgt aangeduid: ZX-locatienummer.

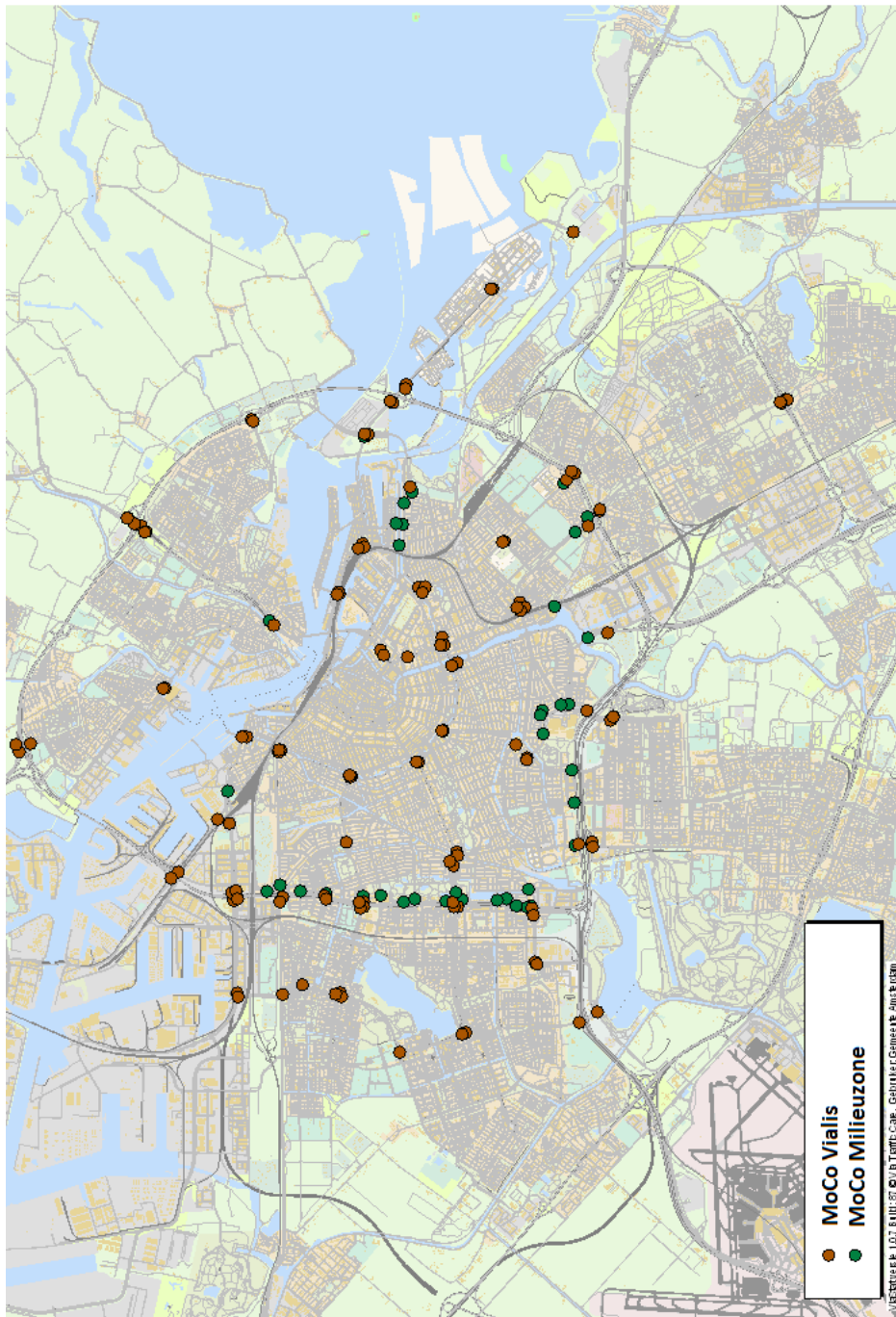
Op een meetlocatie kunnen meerdere rijstroken bemeten zijn. Elke LPR-camera bemeeet slechts één rijstrook. Een LPR-camera heeft een meetlocatie en het rijstrooknummer. De meest rechtse rijstrook wordt aangeduid met 1. De rijstrook één links van de meest rechtse rijstrook wordt aangeduid met 2, etc. Een uitzondering hierop zijn de locaties waar alle rijstroken bemeten zijn. Hier wordt de meest linkse rijstrook aangeduid met 1 en de rijstrook één rechts van de meest linkse rijstrook met 2, etc.

In tegenstelling tot de MoCo-camera's van Vialis hebben de meetlocaties van de milieuzonecamera's geen ZX voor hun locatienummer. De naamgeving van een camera in dit verslag zal als volgt zijn:

meetlocatie_rijstrooknummer

Voor alle MoCo-camera's is bekend wat de coördinaten zijn van de locatie en wat de straatnaam is waar de camera op staat. Een lijst met deze locaties is weergegeven in Appendix C.

In figuur 5 zijn de locaties van MoCo-camera's weergegeven.



Figuur 5: Locaties van MoCo-camera's.

3 Dataset

Dit hoofdstuk gaat over de dataset. De dataset bestaat uit kentekenregistraties en betrouwbaarheidsdata. Ten eerste zal aan worden gegeven hoe de data is verzameld. Daarna zal de metadata worden beschreven. Het hoofdstuk zal worden afgesloten met een verkennende data-analyse.

3.1 Dataverzameling

3.1.1 Kentekenregistraties

Via de afdeling VKS is toegang te verkrijgen tot data uit MoCo-camera's. Wekelijks zet Vialis data uit haar database op een server welke toegankelijk is voor VKS.

Elke week worden vier zip-bestanden geleverd. Deze vier zip-bestanden hebben als naam Werkelijke Reistijden, Separate Reistijden, Basis Reistijden en Matches. In de zip-bestanden van Werkelijke Reistijden, Separate Reistijden en Basis Reistijden is voor een aantal cameraparen een csv-bestand aanwezig. Voor het zip-bestand Matches is voor elke MoCo-camera een csv-bestand aanwezig. In tabel 1 zijn de vier zip-bestanden met omschrijving van de inhoud weergegeven.

Naam	Omschrijving
Werkelijke Reistijden	De gemiddelde reistijd tussen twee camera's in seconden weergegeven per minuut.
Separate Reistijden	De reistijd in seconden tussen twee camera's. Het tijdstip waarop een motorvoertuig de eerste camera is gepasseerd en de daarbij behorende reistijd in seconden.
Basis Reistijden	De voorspelde reistijd tussen twee camera's in seconden weergegeven per minuut.
Matches	Registratie van versleutelde kentekens van gepasseerde motorvoertuigen. Het tijdstip waarop een motorvoertuig de camera's gepasseerd en het daarbij behorende versleutelde kenteken.

Tabel 1: De vier zip-bestanden die wekelijks door Vialis worden geleverd

De data die nodig zijn voor het bepalen van intensiteiten staan in het zip-bestand Matches. Omdat dit onderzoek over intensiteiten gaat, richten we ons alleen op Matches. Op de overige zip-bestanden zal niet meer worden ingegaan.

Elk zip-bestand heeft de volgende bestandsnaamstructuur:

```
Matches_<jaar begin><maand begin><dag begin>_<jaar eind>  
<maand eind><dag eind>.zip
```

Alle kentekenregistraties voor de week van 02-10-2011 tot 09-10-2011 staan bijvoorbeeld in het zip-bestand:

Matches_20111002_20111009.zip

In een zip-bestand is voor elke camera een csv-bestand aanwezig met de kentekenregistraties voor die week. De bestandsnaamstructuur van deze csv-bestanden is als volgt:

```
Matches_<serienummer><rijstrook>_ <jaar begin><maand begin>
      <dag begin>_<jaar eind><maand eind><dag eind>.csv
```

Alle kentekenregistraties voor de camera met serienummer 11 en rijstrooknummer 2 staan in het volgende bestand:

Matches_11002_20111002_20111009.csv

3.1.2 Betrouwbaarheidsdata

Maandelijks brengt Vialis een rapportage uit waarin staat beschreven hoe goed de MoCo-camera's hebben gewerkt. Omdat de milieuzonecamera's in beheer zijn van Gatsometer, doet Vialis in deze rapporten geen uitspraken over de betrouwbaarheid van de milieuzonecamera's. De rapporten worden verstuurd naar de afdeling VKS. In de rapporten wordt gebruik gemaakt van het beschikbaarheidspercentage en het herkeningspercentage. Om het herkeningspercentage te definiëren introduceren we het begrip trigger:

Definitie 3.1 : Een trigger is een poging om een kenteken te registreren en vindt plaats wanneer de software in de camera vermoedt een kenteken in beeld te zien.

In tabel 2 zijn de betrouwbaarheidsindicatoren weergegeven. In hoofdstuk 2 zagen we dat elke minuut een informatiepakket wordt verstuurd van de MoCo-camera naar het datacentrum van Vialis. Per minuut kunnen dus maximaal 60 informatiepakketten verstuurd en ontvangen worden.

Betrouwbaarheidsindicator	Omschrijving	Berekening
Beschikbaarheidspercentage	% van de tijd dat een camera beschikbaar was	Aantal informatiepakketten ontvangen door Vialis datacentrum / Totaal aantal informatiepakketten mogelijk
Herkeningspercentage	% herkende kentekens uit aantal triggers	Aantal kentekens herkend / Aantal keer camera getriggerd

Tabel 2: Betrouwbaarheidsindicatoren.

Omdat de data in de rapporten alleen grafisch is weergegeven, is de betrouwbaarheidsdata bij Vialis opgevraagd.

De betrouwbaarheidsgetallen die geleverd zijn kennen een beperking voor dit onderzoek: De geleverde betrouwbaarheidsgetallen zijn gemiddeldes per maand. Voor nauwkeuriger onderzoek zijn deze betrouwbaarheidsgetallen per uur nodig.

Vialis is verzocht om de betrouwbaarheidsdata per uur te leveren. Vialis heeft echter aangegeven de betrouwbaarheidsdata per uur niet voor dit onderzoek te kunnen leveren, omdat dit veel werk is. Maar Vialis heeft aangegeven dat de technische mogelijkheid bestaat om de storingsdata per uur te leveren.

3.2 Beschrijving metadata

3.2.1 Kentekenregistraties

	A	B	C	L
1	DATUM	KENTEKEN	BETROUWBAARHEID	
2	10-2-2011 0:00	F56280FD55687FA8C4AD16FA3927B671		99
3	10-2-2011 0:00	7125191544A3CBC11FA734AA13AC6A5C		93
4	10-2-2011 0:00	54A1FE51DC0F4C82802615AD06AAC192		90
5	10-2-2011 0:01	5CE2625F6C725E60CD803CE7F017F129		88
6	10-2-2011 0:01	442E26DD257ADCA4D3FB1EDEF7750BF6		92
7	10-2-2011 0:01	EE492D169BCB65ED6DB5D3E585014653		92
8	10-2-2011 0:01	BD85AABE3F6131AFD8CA2A652F0633DC		90
9	10-2-2011 0:02	24DBA7457D9E91FED204A0E9559A0E0B		89
10	10-2-2011 0:02	F366097F96AD92156A4A69AFC4F9E564		88

Figuur 6: Matches csv-bestand.

In figuur 6 is een voorbeeld te zien van de data uit een csv-bestand. In de eerste kolom DATUM staat het tijdstip waarop een motorvoertuig is geregistreerd. De tweede kolom KENTEKEN geeft het versleutelde kenteken van het voertuig weer. In de derde kolom staat een maat voor betrouwbaarheid dat het geregistreerde versleutelde kenteken ook het kenteken was dat door de camera was te zien. Hoe dit getal precies is berekend blijkt niet gemakkelijk te achterhalen.

Het soort voertuig dat de camera passeerde is niet uit dit csv-bestand af te leiden. Deze informatie is bij Vialis wel aanwezig. Op basis van de eerste twee letters van een herkend kenteken bepaalt de camera wat voor soort voertuig is herkend.

3.2.2 Betrouwbaarheidsdata

A	B	C	D
	Maart		April
Meetlocaties	Herkenningspercentage	Beschikbaarheidspercentage	Herkenningspe
3399	0,920		0,996
3400	0,760		0,996
3401	0,970		0,997
3402	0,920		0,970
3403	0,930		0,997
3404	0,950		0,996

Figuur 7: Betrouwbaarheidsdata.

Figuur 7 geeft een deel van de betrouwbaarheidsdata weer. Per meetlocatie zijn per maand het gemiddelde herkenningspercentage en het gemiddelde beschikbaarheidspercentage gegeven.

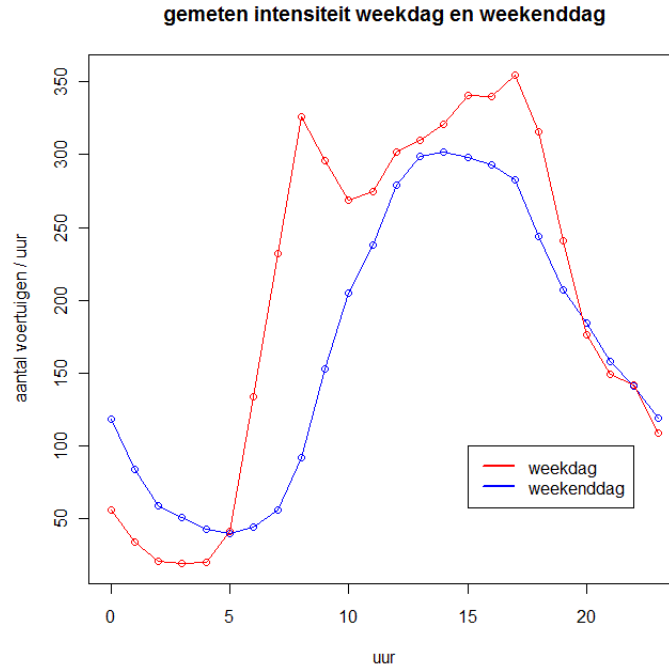
3.3 Verkennende Data-analyse

In dit onderzoek is een steekproef genomen van kentekenregistraties van de volgende vijf weken van het jaar 2011: 10, 14, 19, 38, en 39. Dit zijn respectievelijk een week in maart, een week in april, een week in mei en twee weken in september. Daarnaast is er ook de beschikking over betrouwbaarheidsdata van deze maanden.

3.3.1 Kentekenregistraties

Ten eerste bekijken we de gemeten intensiteiten op een weekday en weekenddag. De gemeten intensiteit is het aantal herkende kentekens in een bepaalde periode op een bepaalde plaats. De weekday en weekenddag delen we op in 24 periodes van een uur. De eerste periode begint om 12 uur 's nachts en eindigt om 1 uur 's nachts. De tweede periode begint om 1 uur 's nachts en eindigt om 2 uur 's nachts, etc. Voor elke periode zijn er voor elke camera 25 waarnemingen voor de weekday en 10 waarnemingen voor de weekenddag. Er zijn 5 weekdays en 2 weekenddagen. Voor 5 weken geeft dit $5 \cdot 5 = 25$ waarnemingen voor de weekday en $5 \cdot 2 = 10$ waarnemingen voor de weekenddag.

Voor elke camera zijn er voor de 24 periodes een gemiddelde gemeten intensiteit. Vervolgens bekijken we voor elke periode wat de gemiddelde gemeten intensiteit is van alle camera's voor die periode. Dit geeft een gemiddelde gemeten intensiteit voor een bepaalde periode. Het resultaat is weergegeven in figuur 8. Op de x-as staat het begin van elke periode. Het einde van de desbetreffende periode is het begin van de periode plus een uur.

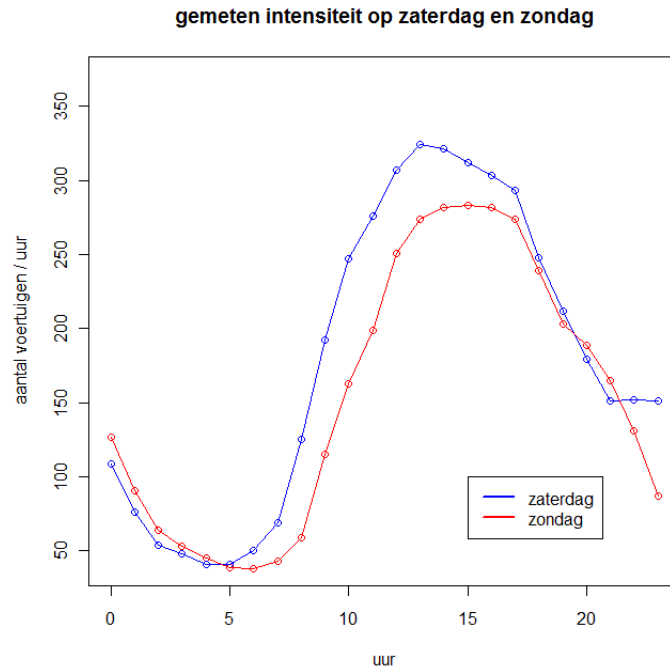


Figuur 8: Gemeten intensiteit op weekdag en weekenddag; De x-as geeft het begin van een periode aan. De periode eindigt een uur na de beginperiode.

In figuur 8 is te zien dat voor de weekdag twee pieken aanwezig zijn. Er is een piek rond de periode die begint om 08:00 uur en een piek rond de periode die begint om 17:00 uur. Tijdens het weekend zijn deze pieken niet aanwezig. In het weekend is het meeste verkeer gemeten rond de periode die begint om 14:00 uur. De piek in het weekend is echter niet zo spits als de pieken van de weekdag.

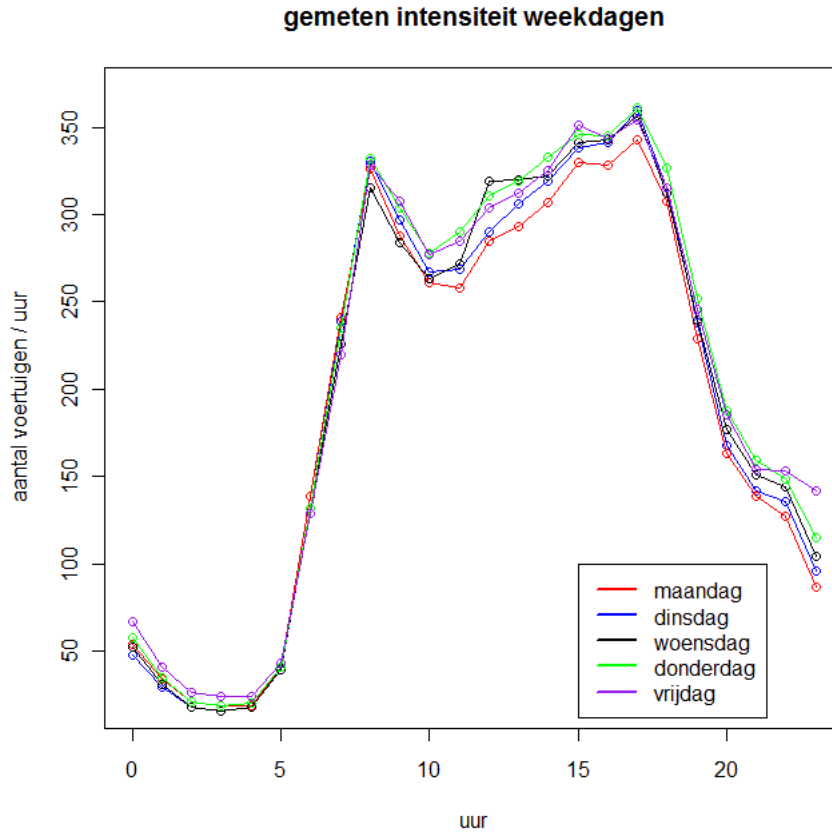
Daarnaast is te zien dat gemiddelde gemeten intensiteit in het weekend bijna altijd lager ligt dan de gemiddelde gemeten intensiteit gedurende de week. Echter in de periodes die beginnen om 20:00 uur tot en met 05:00 uur liggen de gemiddelde gemeten intensiteit in het weekend hoger dan gedurende de week.

In figuur 9 is weergegeven hoe de gemiddelde gemeten intensiteit verdeeld is op de zaterdag en zondag. Op zaterdag en zondag is ongeveer hetzelfde dagpatroon te zien. Op zaterdag zijn echter meer voertuigen gemeten dan op de zondag.



Figuur 9: Gemeten intensiteit op zaterdag en zondag; De x-as geeft het begin van een periode aan. De periode eindigt een uur na de beginperiode.

Ook wordt onderzocht of er verschillen zijn tussen de verschillende dagen in de week. De gemiddelde gemeten intensiteit over de verschillende dagen in de week is weergegeven in figuur 10. In figuur 10 is te zien dat de gemiddelde gemeten intensiteit over de verschillende dagen in de week dicht bij elkaar ligt. Opvallend is dat de gemiddelde gemeten intensiteit op maandag in de periodes die beginnen om 10.00 uur tot en met 17:00 uur lager ligt dan op andere dagen. De vrijdag onderscheidt zich van andere dagen doordat de gemiddelde gemeten intensiteit vanaf de periode die begint om 23.00 uur gemiddeld hoger is dan op andere dagen.

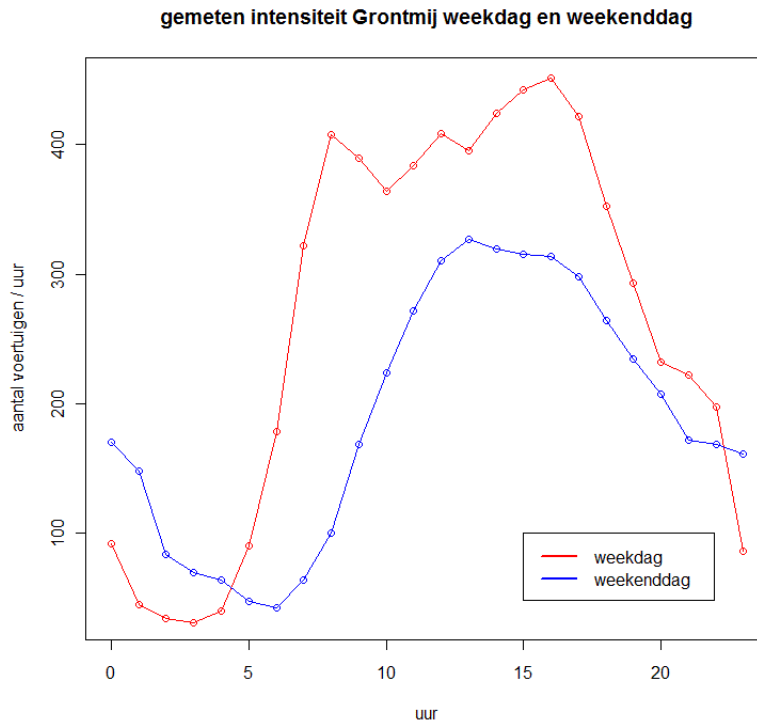


Figuur 10: Gemeten intensiteit op weekdagen; De x-as geeft het begin van een periode aan. De periode eindigt een uur na de beginperiode.

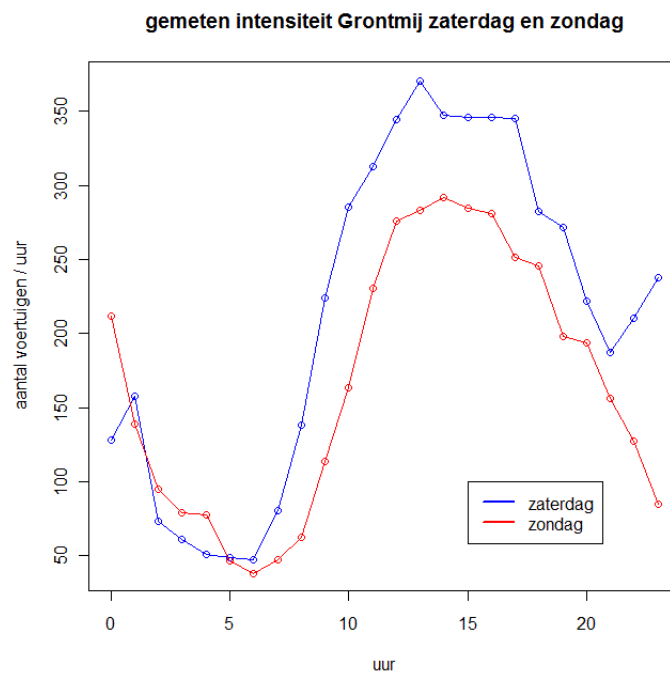
In figuur 8, figuur 9 en figuur 10 zijn trends te zien die te verklaren zijn. De pieken in figuur 8 lijken de ochtend- en de avondspits weer te geven. De ochtendspits is de periode tussen 08:00 en 10:00 en de avondspits is de periode tussen 16:00 en 18:00. Figuur 9 laat het verschil tussen zondag en zaterdag zien. Dat op zaterdag meer verkeer gemeten is op de weg, kan mogelijk verklaard worden doordat de winkels wel op zaterdag open zijn en niet op zondag. In figuur 10 is te zien dat op maandag minder verkeer is gemeten. Een mogelijke verklaring hiervoor is dat veel mensen op maandag een vrije dag hebben. De hogere gemiddelde gemeten intensiteit op vrijdag vanaf die periode die begint om 23:00 is mogelijk een gevolg van het feit dat op vrijdagavond veel mensen uitgaan.

De trends in figuur 8, figuur 9 en figuur 10 kunnen we vergelijken met trends uit andere tellingen. Hiervoor hebben we de meest recente tellingen gebruikt

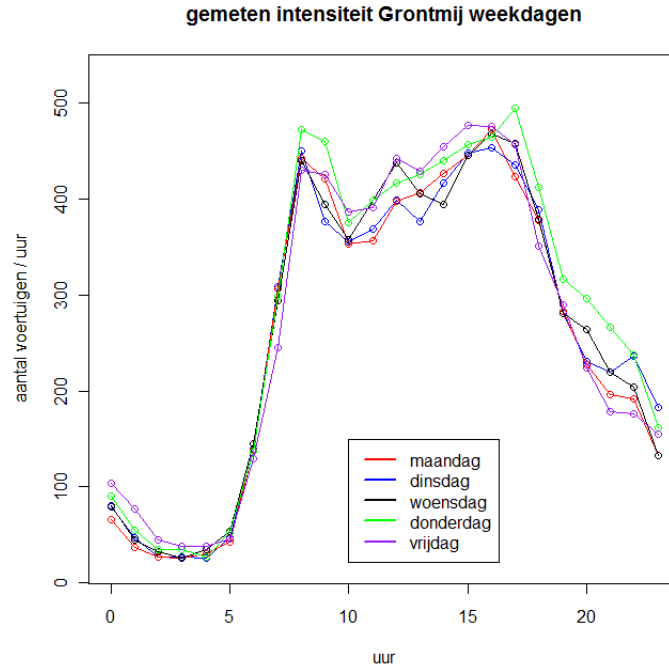
die uitgevoerd zijn door het bedrijf Grontmij. De steekproef die we van deze tellingen hebben is echter kleiner dan de steekproef uit de MoCo-data. Daarom zijn de trends van de meest recente tellingen minder duidelijk. Toch kunnen we zien dat de trends van de gemeten intensiteiten op basis van MoCo-data sterk lijken op de trends van gemeten intensiteiten gebaseerd op Grontmij tellingen. De trends van de Grontmij tellingen zijn te zien in respectievelijk figuur 11, figuur 12 en figuur 13.



Figuur 11: Gemeten intensiteit op basis van Grontmij tellingen op een weekday en weekenddag; De x-as geeft het begin van een periode aan. De periode eindigt een uur na de beginperiode.



Figuur 12: Gemeten intensiteit op basis van Grontmij tellingen op zaterdag en zondag; De x-as geeft het begin van een periode aan. De periode eindigt een uur na de beginperiode.



Figuur 13: Gemeten intensiteit op basis van Grontmij tellingen op weekdagen; De x-as geeft het begin van een periode aan. De periode eindigt een uur na de beginperiode.

Omdat de trends in figuur 8, figuur 9 en figuur 10 te verklaren zijn, kan geconcludeerd worden dat de trends goed overeenkomen met trends in de werkelijkheid. We hebben het hier alleen over de vorm van de grafieken. Over het niveau van de gemiddelde gemeten intensiteit kunnen we nog geen uitspraken doen. In elke periode hebben we het gemiddelde genomen van de gemeten intensiteit van alle camera's. Deze gemiddelde gemeten intensiteit kan per camera verschillen. Het is daarom nog niet te zeggen hoe realistisch deze gemiddelde gemeten intensiteiten zijn.

Cameraniveau Om de data verder te onderzoeken is een subset van de dataset onderzocht. Deze subset van de dataset is zodanig gekozen dat we de resultaten van het toepassen van de methodes in hoofdstuk 4 en 5 op de subset van deze dataset kunnen vergelijken met oude tellingen. In hoofdstuk 6 zal blijken dat veel van deze oude tellingen hebben plaatsgevonden tussen 16:00 - 18:00 op een dinsdag, woensdag of donderdag. Daarom wordt voor alle camera's de gemeten intensiteit tussen 16:00 - 18:00 bekeken op dinsdag, woensdag en donderdag.

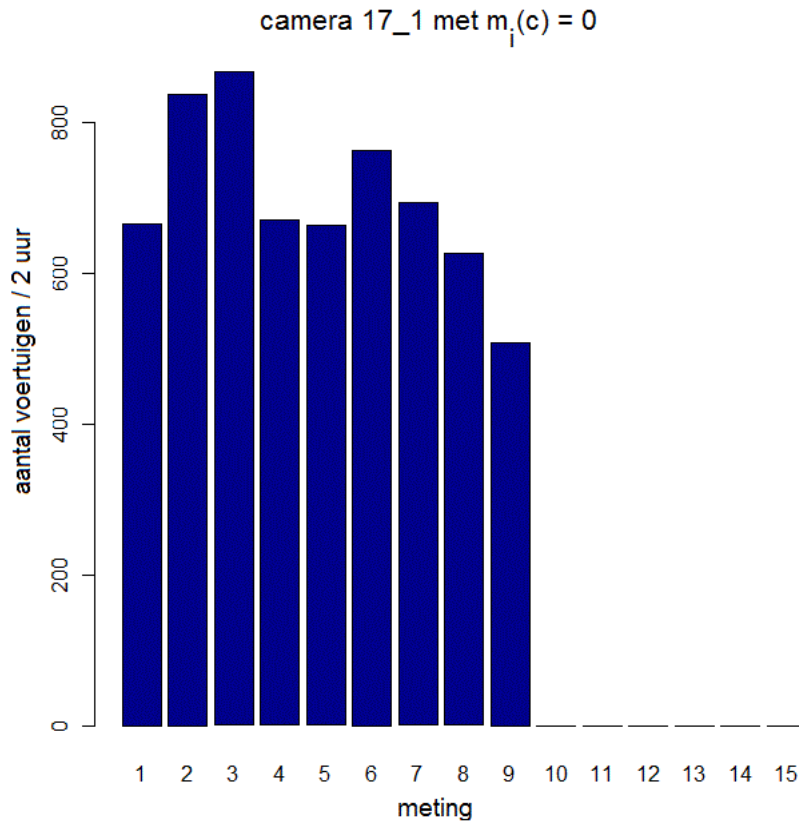
Dit geeft voor elke camera voor vijf weken drie metingen. Deze metingen

definiëren we met $m_i(c)$:

$$m_i(c) = \# \text{ herkende kentekens tussen 16:00 - 18:00 voor camera } c \text{ op de } i\text{-de meting}$$

Dit zijn in totaal per camera vijftien waarnemingen van $m_i(c)$. De dinsdag, woensdag en donderdag in maart worden respectievelijk aangegeven met $i = 1$, $i = 2$ en $i = 3$. De dinsdag, woensdag en donderdag in april worden respectievelijk aangegeven met $i = 4$, $i = 5$, en $i = 6$, etc.

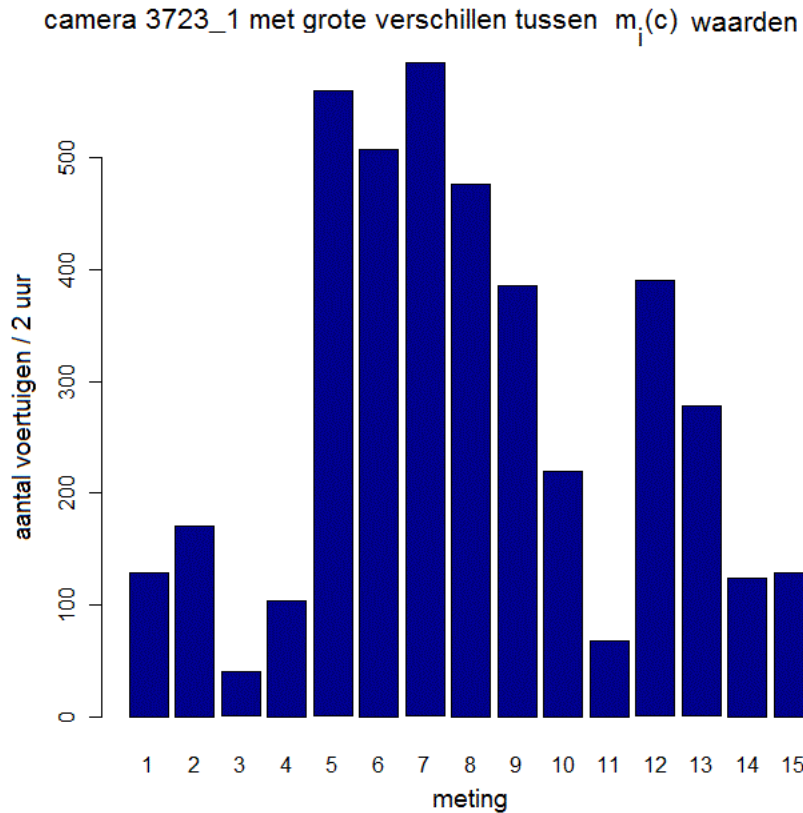
Het eerste dat opvalt is dat sommige camera's de waarden $m_i(c) = 0$ hebben voor zekere i waarden. Een voorbeeld is gegeven in figuur 14. De waarden $m_i(c) = 0$ kunnen er op duiden dat de camera niet werkte, waardoor er geen data naar het Vialis datacenter kon worden verstuurd. Een andere mogelijkheid is dat er geen verkeer reed, bijvoorbeeld door afsluiting van de weg.



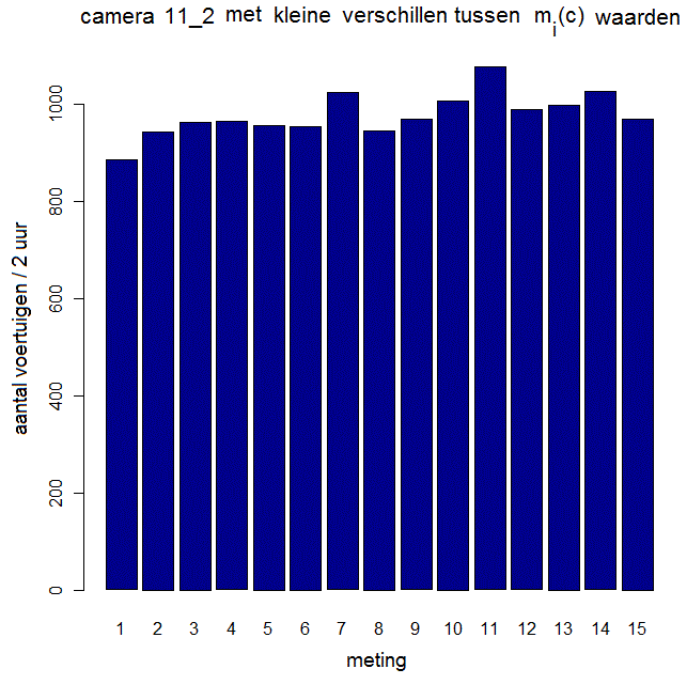
Figuur 14: Camera met $m_i(c) = 0$.

In figuur 15 en figuur 16 zijn de $m_i(3723_1)$ waarden en $m_i(11_2)$ waarden weergegeven. Er is duidelijk een verschil tussen de $m_i(3723_1)$ waarden en $m_i(11_2)$ waarden op te merken. De $m_i(11_2)$ waarden verschillen minder van elkaar dan de $m_i(3723_1)$ waarden. Dit kan erop duiden dat camera 3723_1 minder betrouwbaar was en veel verkeer dat langs de camera reed niet heeft geregistreerd. Op basis van alleen de $m_i(3723_1)$ waarden is deze conclusie niet te trekken. Het is ook mogelijk dat het aantal voertuigen dat langs camera 3723_1 reed tijdens de metingen veel verschilde.

Een ander verschil tussen $m_i(3723_1)$ waarden en de $m_i(11_2)$ waarden is de orde van grootte. Het maximum van de $m_i(3723_1)$ waarden is $m_7(3723_1) = 592$, en het maximum van de $m_i(11_2)$ waarden is $m_{11}(11_2) = 1212$. Dit verschil lijkt verklaarbaar doordat de weg waarop camera 11_2 staat mogelijk drukker is.



Figuur 15: Camera met grote verschillen tussen $m_i(c)$ waarden.



Figuur 16: Camera met kleine verschillen tussen $m_i(c)$ waarden.

Uit figuur 15 en figuur 16 blijkt dat er grote verschillen kunnen zijn in de spreiding van $m_i(c)$ waarden. Om de spreidingen in $m_i(c)$ waarden van verschillende camera's met elkaar te vergelijken moet gebruikt gemaakt worden van een relatieve spreidingsmaat. Hiervoor wordt de steekproef variatiecoëfficiënt $\hat{v}(c)$ gebruikt:

$$\hat{v}(c) = \frac{s_n(m(c))}{\bar{m}_n(c)}.$$

Hierbij is $\bar{m}_n(c)$ gedefinieerd als:

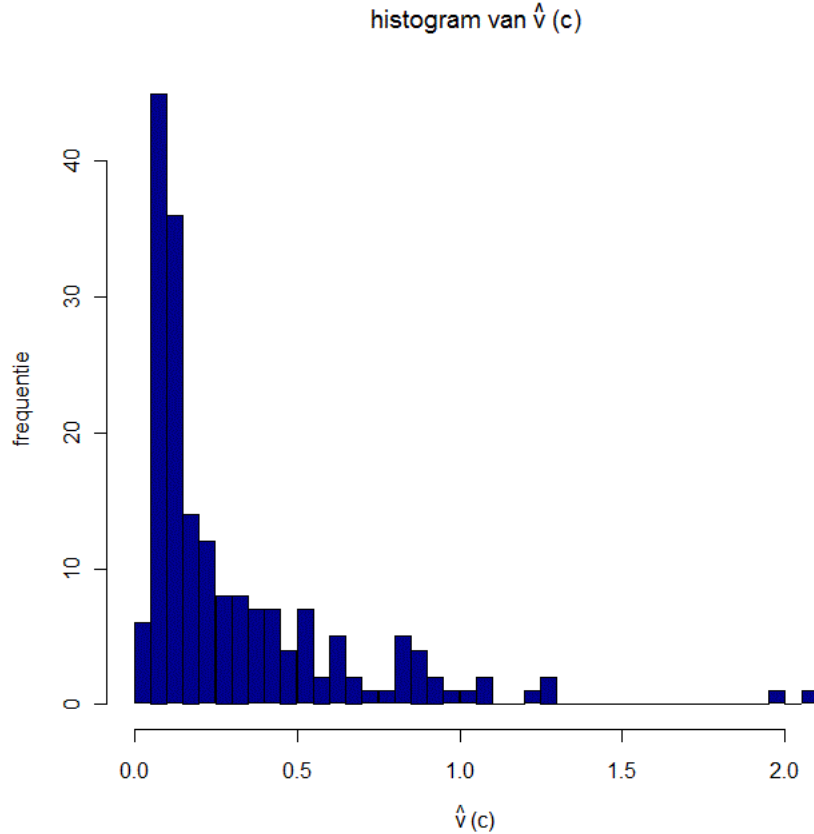
$$\bar{m}_n(c) = \frac{1}{n}(m_1(c) + \dots + m_n(c)).$$

En $s_n(m(c))$ is gedefinieerd als:

$$s_n(m(c)) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (m_i(c) - \bar{m}_n(c))^2}.$$

In deze formule is $s_n(m(c))$ de steekproef standaarddeviatie van de $m_i(c)$ waarden. $\bar{m}_n(c)$ is het steekproefgemiddelde van de $m_i(c)$ waarden.

De $\hat{v}(c)$ waarden van camera 3723_1 en camera 11_2 zijn respectievelijk $\hat{v}(3723_1) = 0,68$ en $\hat{v}(11_2) = 0,05$. In figuur 17 is een histogram weergegeven van de $\hat{v}(c)$ waarden van alle camera's. De meeste $\hat{v}(c)$ waarden zijn kleiner dan 0,2. Toch is te zien dat sommige camera's een $\hat{v}(c)$ waarde groter dan 2 hebben.

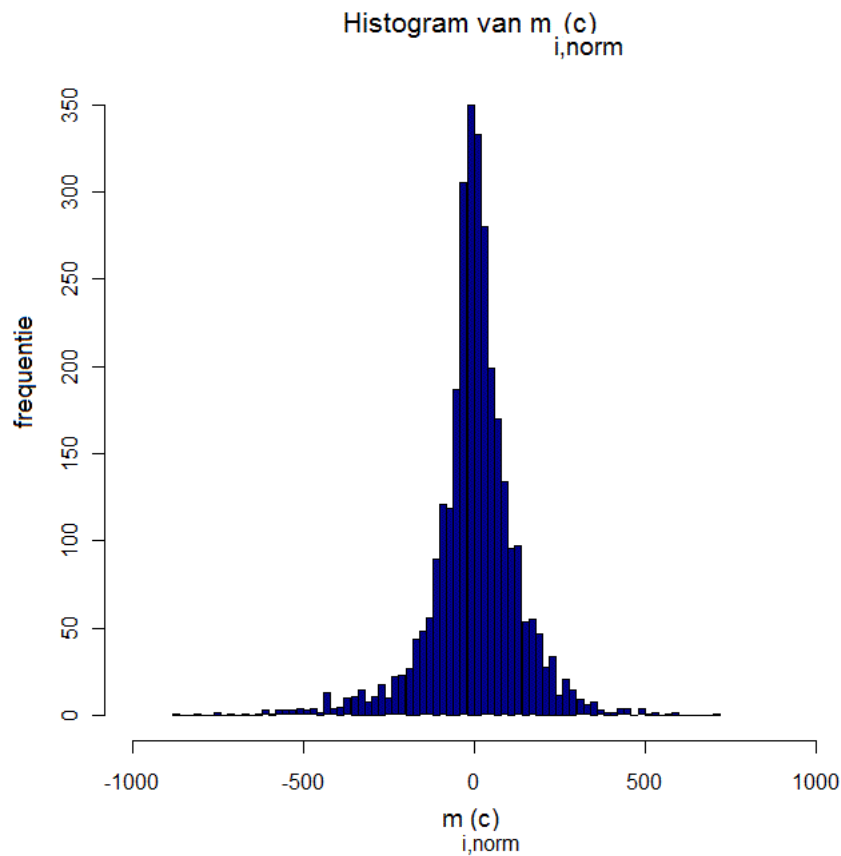


Figuur 17: Verdeling $\hat{v}(c)$ van alle camera's.

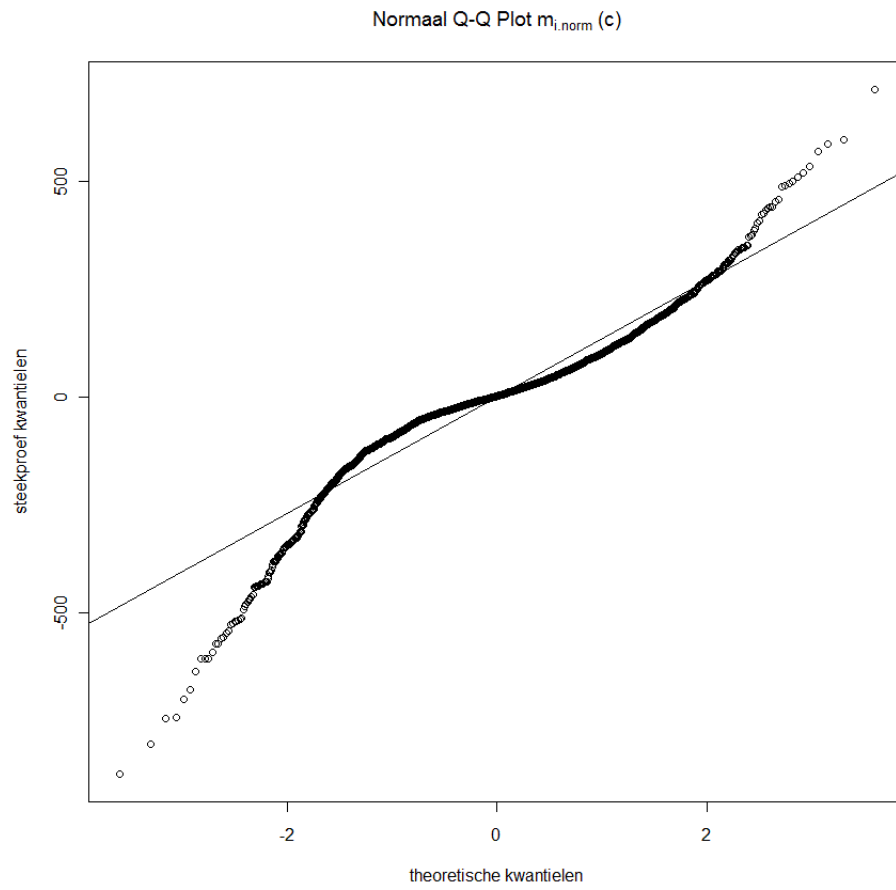
Ook willen we kijken naar de verdeling van de $m_i(c)$ waarden. Zoals te zien in figuur 15 en figuur 16 verschillen de $m_i(c)$ waarden van verschillende camera's in orde van grootte. We willen de verdeling van $m_i(c)$ waarden bekijken, waarbij we het effect van de verschillende orde van grootte van de $m_i(c)$ waarden van de verschillende camera's buiten beschouwing laten. Hiervoor moeten de waarden van $m_i(c)$ genormaliseerd worden. Dit betekent dat de $m_i(c)$ waarden verminderd worden met $\bar{m}_n(c)$. Hierdoor ontstaat de genormaliseerde gemeten intensiteit $m_{i,norm}(c)$:

$$m_{i,norm}(c) = m_i(c) - \bar{m}_n(c).$$

In figuur 18 zijn alle waarden van $m_{i,norm}(c)$ weergegeven. De symmetrische vorm van de verdeling van $m_{i,norm}(c)$ doet denken aan een normale verdeling. Om te bepalen of $m_{i,norm}(c)$ normaal verdeeld is bekijken we het normaal QQ-plot. Deze is weergegeven in in figuur 19. In deze QQ-plot is een referentielijn weergegeven. Als de punten ongeveer op de referentielijn liggen, dan is $m_{i,norm}(c)$ normaal verdeeld. Te zien is dat de punten afwijken van de referentielijn bij de lage theoretische kwantielen ongeveer kleiner dan -2 en bij de hoge theoretische kwantielen ongeveer groter dan 2. De staarten wijken dus af van een normale verdeling. $m_{i,norm}(c)$ is dus niet normaal verdeeld.

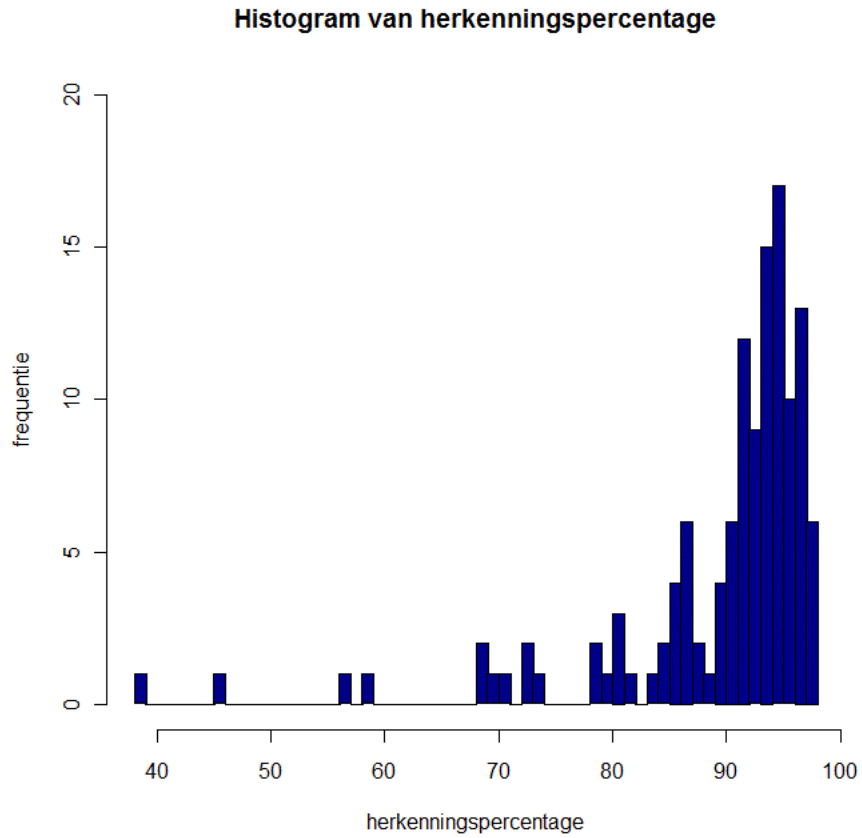


Figuur 18: Histogram van $m_{i,norm}(c)$ waarden.



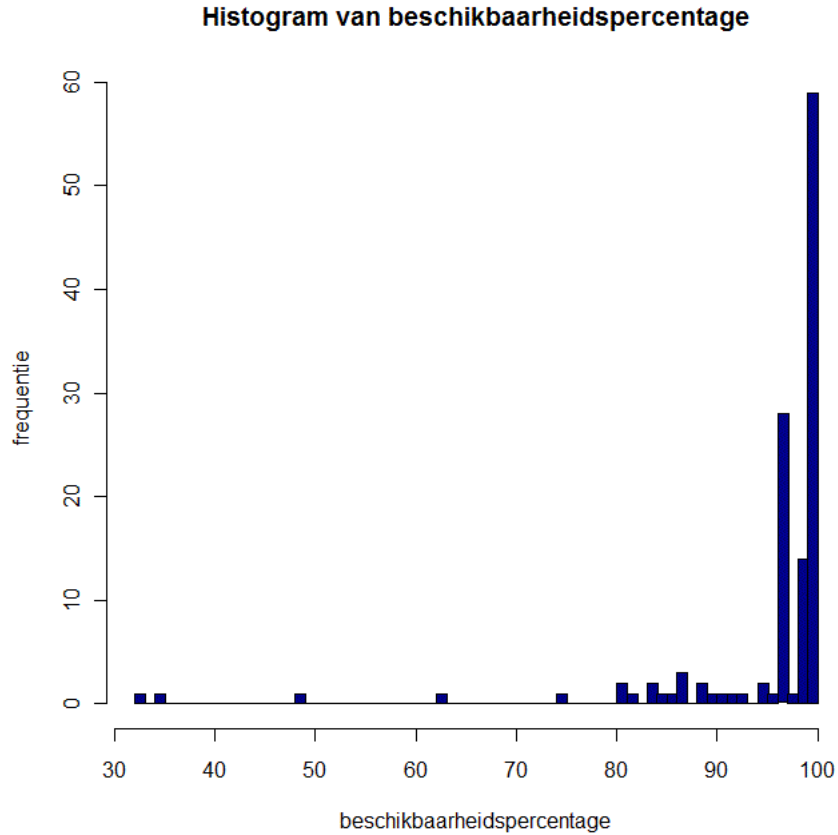
Figuur 19: Normaal QQ-plot van $m_{i, \text{norm}}(c)$.

3.3.2 Storingsdata



Figuur 20: Gemiddelde herkenningspercentages.

In figuur 20 is een histogram gegeven van de gemiddelde herkenningspercentages over de maanden per camera. Het grootste deel van de herkenningspercentages ligt tussen de 90% en de 100%. Het gemiddelde herkenningspercentage is 89%.



Figuur 21: Gemiddelde beschikbaarheidspercentages.

De gemiddelde beschikbaarheidspercentages zijn weergegeven in figuur 21. Veel beschikbaarheidspercentages liggen dichtbij de 100%. Het gemiddelde beschikbaarheidspercentage is 94%

Uit de verkennende data-analyse kan het volgende geconcludeerd worden:

1. De vorm van de weektrend en dagtrend lijken plausibel ;
2. Er zijn een aantal metingen $m_i(c) = 0$ voor zekere i waarden;
3. De meeste $\hat{v}(c)$ waarden liggen onder de 0,2. Toch zijn er $\hat{v}(c)$ waarden boven de 2;
4. De verdeling van $m_{i,norm}(c)$ lijkt symmetrisch, maar is geen normale verdeling;
5. Het gemiddelde herkenningspercentage is 89%; het gemiddelde beschikbaarheidspercentage is 94%.

4 Filtermethode

Uit de verkennende data-analyse bleek dat er grote verschillen zijn in $\hat{v}(c)$ waarden. Mogelijk zijn de grote verschillen te verklaren door onbetrouwbaarheid van de camera's. Het doel van dit hoofdstuk is om een methode te ontwikkelen om op basis van $m_i(c)$ waarden intensiteiten te bepalen die gebruikt kunnen worden voor het verkeersmodel. Het hoofdstuk is als volgt opgedeeld:

- Ten eerste zullen een aantal notaties worden geïntroduceerd.
- Daarna zal het schema van kentekenherkenning worden geïntroduceerd. Hierin is schematisch weergegeven welke situaties zich kunnen voordoen wanneer een voertuig langs een camera rijdt.
- Vervolgens zal een analyse gemaakt worden van de betrouwbaarheid van camera's. Deze betrouwbaarheidsanalyse is opgedeeld in twee delen: Ten eerste zullen de resultaten uit een betrouwbaarheidsonderzoek van het reistijdmeetsysteem in Rotterdam worden geanalyseerd. Daarna zullen de betrouwbaarheidsdata worden geanalyseerd.
- Tenslotte zal op basis van conclusies uit de betrouwbaarheidsanalyse een methode worden opgebouwd om waarden voor intensiteiten in het verkeersmodel te bepalen.

4.1 Notatie

De stochastische variabele Q_w is het aantal gepasseerde voertuigen op een bepaalde plaats over een bepaalde periode Δt . Q_w is als volgt gedefinieerd:

$$Q_w = \frac{N}{\Delta t}.$$

Hierbij is de stochastische variabele N het aantal werkelijk gepasseerde voertuigen en Δt de periode waarover gemeten is. Op basis van de kentekenregistraties kunnen intensiteiten bepaald worden die door de camera zijn gemeten. Dit is gedefinieerd door de stochastische variabele Q_g , welke het aantal gemeten voertuigen weergeeft:

$$Q_g = \frac{N_h}{\Delta t}.$$

De stochastische variabele N_h geeft het aantal herkende kentekens weer. Omdat de aanwezigheid van betrouwbaarheidsdata impliceert dat niet alle camera's altijd betrouwbaar zijn, kan gesteld worden dat Q_g niet gelijk is aan Q_w .

De MoCo-camera's hebben een bepaalde onbetrouwbaarheid. Daarom kan Q_w als volgt worden uitgedrukt:

$$Q_w = Q_g + \varepsilon.$$

Hierbij is de stochastische variabele ϵ het verschil tussen het aantal werkelijke voertuigen dat langs een punt reed en het aantal gemeten voertuigen.

De verdeling van Q_g kunnen we op basis van de MoCo-data niet exact bepalen. Wel kunnen we de $m_i(c)$ waarden gebruiken om te verdeling van Q_g te benaderen. $m_i(c)$ is een gemeten intensiteit. We nemen aan dat de $m_i(c)$ waarden komen uit de verdeling van Q_g . Volgens de wet van de grote aantallen krijgen we dan:

$$\bar{m}_n(c) \rightarrow E(Q_g), \text{ als } n \rightarrow \infty.$$

4.2 Schema van kentekenherkenning

Bij voorkeur worden kentekens van alle voertuigen die langs een MoCo-camera rijden herkend. In de praktijk worden niet alle kentekens herkend. Toch kan bepaald worden van welk percentage van de voertuigen dat is langsgereden het kenteken is herkend. Dit gebeurt op basis van triggers.

In de meeste gevallen wanneer een trigger plaatsvindt wordt een kenteken herkend. Echter, wanneer een voertuig langs een camera rijdt, waarvan de software in de camera het kenteken niet heeft herkend, zal dit voertuig vrijwel altijd worden getriggerd. Een voorbeeld hiervan is als sprake is van een vervuild kenteken. Doordat het kenteken vervuild is kan het kenteken niet door de software in de camera worden herkend. Uit paragraaf 4.3.1 zal blijken dat de kans echter groot is dat de software in de camera getriggerd wordt door de vervuilde kentekenplaat.

Het is mogelijk dat de software in de camera wordt getriggerd door iets anders dan een voertuig. Een voorbeeld hiervan is dat de software in de camera wordt getriggerd door een bewegende tak. De volgende drie situaties zijn mogelijk als de software in de camera wordt getriggerd:

- Er reed een voertuig en het kenteken is herkend;
- Er reed een voertuig en het kenteken is niet herkend;
- Iets anders dan een voertuig heeft de software in de camera getriggerd.

Bovendien is het mogelijk dat een voertuig langs de camera rijdt en dat de software in de camera niet wordt getriggerd. Dit is de vierde situatie:

- Er reed een voertuig en de software in de camera is niet getriggerd.

De vier bovenstaande situaties kunnen schematisch worden gegeven, zoals te zien in figuur 22.

Geen trigger – onherkend voertuig	
Trigger – geen voertuig	Trigger – onherkend voertuig
Trigger – herkend voertuig	

Figuur 22: Schema van kentekenherkenning.

Op basis van het schema in figuur 22 kunnen de verschillende situaties als volgt worden gedefinieerd:

$\#trigger_{hv}$ = aantal triggers dat over een bepaalde periode in de categorie trigger-herkend voertuig valt.

$\#trigger_{ov}$ = aantal triggers dat over een bepaalde periode in de categorie trigger-onherkend voertuig valt.

$\#trigger_{gv}$ = aantal triggers dat over een bepaalde periode in de categorie trigger-geen voertuig valt.

$\#geentri\text{gger}_{ov}$ = aantal voertuigen dat over een bepaalde periode langs de camera is gereden maar niet is getriggerd.

Vervolgens kan het totaal aantal triggers $\#totaaltrigger$ als volgt bepaald worden:

$$\#totaaltrigger = \#trigger_{hv} + \#trigger_{ov} + \#trigger_{gv}.$$

Het herkenningspercentage is dan:

$$herkenningspercentage = \frac{\#trigger_{hv}}{\#totaaltrigger}.$$

Het beschikbaarheidspercentage is het percentage van het aantal informatiepakketten dat aankomt bij het Vialis datacentrum als percentage van het aantal informatiepakketten dat maximaal in die periode ontvangen kon worden. Wanneer het beschikbaarheidspercentage $< 100\%$, dan betekent dit dat over een deel van de periode waarover is gemeten geen data is binnen gekomen. Voertuigen die in dit deel van de periode de camera passeren, vallen dan in de categorie geen trigger - onherkend voertuig.

Voorbeeld 4.1: Gedurende een periode van een uur rijden 1000 voertuigen langs een camera. In de laatste 15 minuten stond de camera uit, en werd er geen informatie verstuurd naar het Vialis datacentrum. In het laatste kwartier reden er 250 voertuigen. Gedurende de eerste 45 minuten reden er 750 voertuigen. De software in de camera werd door alle 750 voertuigen getriggerd. Van 500 van de 750 voertuigen was het kenteken herkend.

$$herkenningspercentage = \frac{500}{750} = 66\%.$$

$$beschikbaarheidspercentage = \frac{45}{60} = 75\%.$$

4.3 Analyse van betrouwbaarheid camera's

4.3.1 Onderzoek Rotterdam

Rotterdam heeft een soortgelijk systeem als MoCo. Het Rotterdamse systeem om reistijden te bemeten heet MoRo (Monitoring Rotterdam). Tijdens dit

onderzoek zijn een aantal meetlocaties onderzocht die niet voldeden aan een herkenningpercentage groter dan 75%. Het doel van het onderzoek was om aanbevelingen te doen om de herkenningpercentages op deze meetlocaties te verhogen naar minimaal 75%.

In dit onderzoek zijn de camerabeelden geanalyseerd. Onder normale omstandigheden worden deze camerabeelden niet bewaard. Voor dit onderzoek zijn de beelden opgeslagen. Op vier meetlocaties waar twee camera's staan zijn twee metingen gedaan van een uur. Vialis heeft aangegeven dat in het onderzoek het niet is voorgekomen dat voertuigen langs een camera reden die niet getriggerd waren. Op basis daarvan doen we de volgende aanname:

Aanname 4.1: Als een camera over een bepaalde periode een beschikbaarheidspercentage heeft van 100%, dan $\#geentri\text{gger}_{ov} = 0$.

Op basis van de camerabeelden heeft Vialis onderzocht welke situaties zich voordoen wanneer de software in de camera wordt getriggerd, maar dat geen kenteken is herkend. Vialis heeft de verschillende situaties opgedeeld in een aantal categorieën [4]. Deze zijn weergegeven in tabel 3.

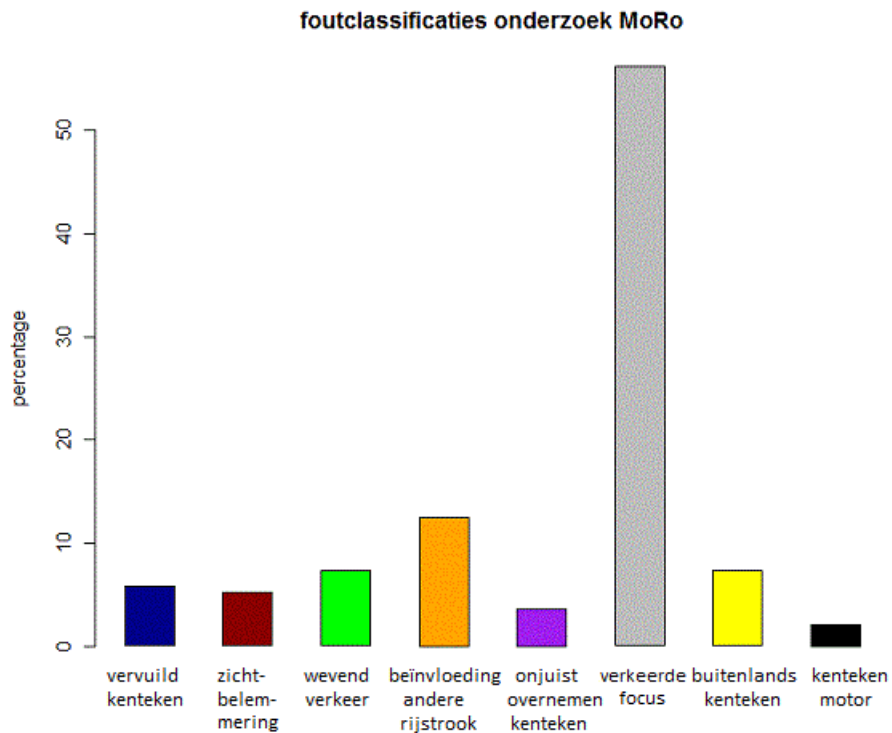
Soort	Omschrijving
Vervuilde kentekens	Kentekens zijn vervuild zodat de camera geen cijfers en letters kan lezen.
Belemmering van zichtveld	Obstakels (aan het voertuig) belemmeren het zicht op de kentekenplaat.
Wevend verkeer	Verkeer gedetecteerd in het midden tussen twee rijstroken.
Beïnvloeding verkeer andere rijstrook	Verkeer van de andere rijstrook wordt gedetecteerd.
Onjuiste overname van het kenteken	Verkeerde kenteken wordt in database opgenomen door softwarefout.
Focus verkeerd	De camera legt geen focus op de kentekenplaat, maar op iets anders (bijvoorbeeld een reflector op het voertuig).
Buitenlandse kentekens	Buitenlandse kentekens worden niet herkend.
Motorrijders	Kenteken van een motor wordt niet herkend.

Tabel 3: Foutclassificaties onderzoek MoRo.

Opvallend is de aanwezigheid van buitenlandse kentekens en motorrijders als foutclassificatie. Dit heeft uitleg. De software is in staat om buitenlandse kentekens te lezen en hierbij een versleuteld kenteken te genereren. Het verschil tussen een buitenlands kenteken en een Nederlands kenteken is het aantal tekens

van de kentekenplaat en het soort tekens. Nederlandse kentekens bevatten altijd zes tekens. Buitenlandse kentekens bevatten vaak meer dan zes tekens en bovendien soms speciale tekens. Vialis geeft aan dat zij ten behoeve van performance alleen kentekens die zes normale tekens bevat als herkend classificeert. Hetzelfde geldt voor motorrijders. Deze hebben niet zes tekens achter elkaar, omdat het kenteken verdeeld is over twee regels. De versleutelde kentekens worden daarom niet opgeslagen.

In figuur 23 is de verdeling weergegeven van triggers waarbij geen sprake was van een herkend kenteken.



Figuur 23: Percentage foutclassificaties onderzoek MoRo.

Tijdens bijna alle situaties in figuur 23 was er sprake van een voertuig dat getriggerd was, maar waarvan het kenteken niet was herkend. De situatie beïnvloeding verkeer andere rijstrook is anders dan de overige situaties. Hier was weliswaar een voertuig getriggerd, maar dit voertuig reed niet over de gemeten rijstrook. Omdat we geïnteresseerd zijn in het aantal gepasseerde voertuigen op de rijstrook die de camera meet, zouden alle triggers die onder beïnvloeding verkeer andere rijstrook vallen geïnterpreteerd moeten worden als $trigger_{gv}$. De rest van de situaties valt onder $trigger_{ov}$.

De situaties die vallen onder $trigger_{ov}$ zijn ongeveer 90% van de totale situaties. Omdat $trigger_{ov}$ hoog is ten opzichte van $trigger_{gv}$ nemen we voor simplificatie aan dat:

Aanname 4.2: $trigger_{gv} = 0$.

Op basis van het schema van kentekenherkenning kan Q_w gedefinieerd worden als:

$$Q_w = \#trigger_{hv} + \#trigger_{ov} + \#geentrieger_{ov}.$$

Op basis van aanname 4.1 kan Q_w gedefinieerd worden als:

$$Q_w = \#trigger_{hv} + \#trigger_{ov}.$$

En op basis van aanname 4.2 kan $\#totaaltrigger$ gedefinieerd worden als:

$$\#totaaltrigger = \#trigger_{hv} + \#trigger_{ov}.$$

Nu kan ε gedefinieerd worden. Alvorens we dit doen definiëren we eerst de stochastische variabelen herkenningspercentage H en het beschikbaarheidspercentage B :

$H = \text{herkenningspercentage over een bepaalde beperiode.}$

$B = \text{beschikbaarheidspercentage over een bepaalde periode.}$

Bij 100% beschikbaarheidspercentage in de periode waarin de intensiteiten gemeten worden is de enige onbetrouwbaarheid het herkenningspercentage. Dan is de gemeten intensiteit gelijk aan het percentage van het aantal werkelijke voertuigen dat herkend is: HQ_w . Wanneer het beschikbaarheidspercentage niet gelijk is aan 100%, dan is er over een gedeelte van de periode waarin wordt gemeten geen data beschikbaar. Er is slechts over een bepaald percentage van de periode voertuigen gemeten. Om Q_g te krijgen moeten we HQ_w nog met B vermenigvuldigen:

$$Q_g = HBQ_w,$$

$$Q_w = \frac{1}{HB}Q_g.$$

Omdat $\varepsilon = Q_w - Q_g$, kunnen we ε als volgt schrijven:

$$\varepsilon = \frac{1}{HB}Q_g - Q_g,$$

$$\varepsilon = Q_g\left(\frac{1}{HB} - 1\right).$$

4.3.2 Betrouwbaarheidsdata

Informatie over ε was in dit onderzoek aanwezig. Voor elke meetlocatie was het herkenningpercentage en beschikbaarheidspercentage per maand beschikbaar.

Aan de hand van een voorbeeld wordt bepaald wat het gevolg is dat alleen de gemiddelde betrouwbaarheidsdata per maand is gegeven. In figuur 24 zijn de drie intensiteiten van maart weergegeven voor camera 3755_1. Het gemiddelde herkenningpercentage en het gemiddelde beschikbaarheidspercentage voor de maand maart waren respectievelijk 78% en 99,5%.

In figuur 24 is te zien dat er een groot verschil is tussen $m_2(3755_1)$ en $m_3(3755_1)$. $m_2(3755_1)$ is meer dan vijf keer zo groot als $m_3(3755_1)$. Beschouw de volgende formule:

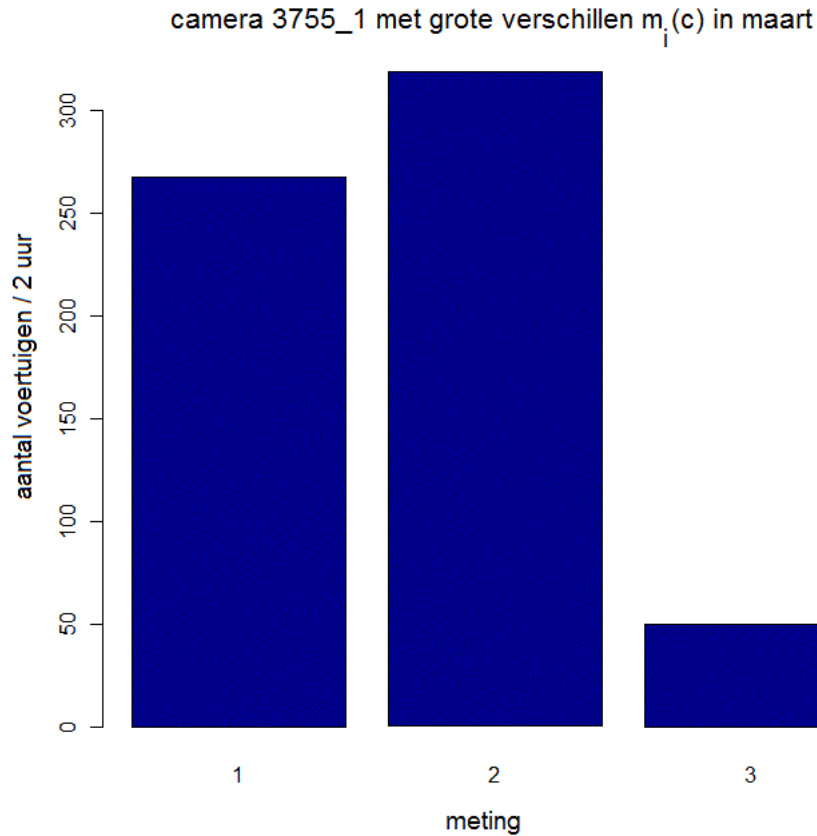
$$Q_g = Q_w - \varepsilon.$$

Invullen van $\varepsilon = Q_g(\frac{1}{HB} - 1)$ geeft

$$Q_g = Q_w - (Q_g(\frac{1}{HB} - 1)).$$

Zoals te zien aan bovenstaande formule worden $m_2(3755_1)$ en $m_3(3755_1)$ bepaald door de werkelijke intensiteit, het herkenningpercentage en beschikbaarheidspercentage tijdens de meting. Aangezien het gemiddelde beschikbaarheidspercentage bijna 100% was voor de hele maand maart, laten we deze factor buiten beschouwing. De grote verschillen tussen $m_2(3755_1)$ en $m_3(3755_1)$ zijn nu verklaarbaar door de werkelijke intensiteit tijdens de meting of het herkenningpercentage tijdens de meting.

Het lijkt onwaarschijnlijk dat tijdens de avondspits meer dan vijf keer zoveel auto's reden op meting 2 dan meting 3. Dit is mogelijk als er tijdens meting 3 sprake was van een extreme verkeersomstandigheid. Er kunnen bijvoorbeeld minder auto's hebben gereden doordat er sprake was van een verkeersongeval. Deze situaties zijn echter zeldzaam. Daarentegen was het gemiddelde herkenningpercentage voor de maand maart met 78% niet hoog. Het lijkt waarschijnlijk dat het herkenningpercentage het grootste gedeelte tussen het verschil tussen $m_2(3755_1)$ en $m_3(3755_1)$ bepaalt. Dit impliceert dat er een grote spreiding lijkt te zijn tussen herkenningpercentages op meting 2 en 3. Dit heeft als gevolg dat op basis van het gemiddelde herkenningpercentage voor de maand maart, niet te bepalen is wat de herkenningpercentages zijn van meting 1, 2 en 3. We weten slechts het gemiddelde herkenningpercentage voor de maand maart.

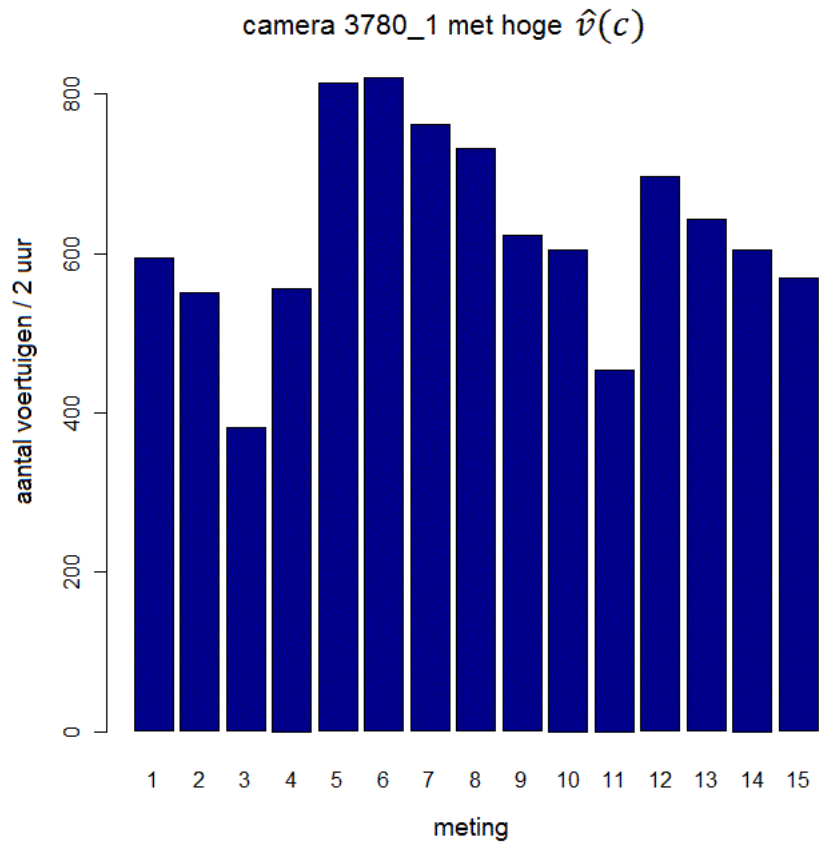


Figuur 24: Camera met grote verschillen tussen $m_i(c)$ waarden.

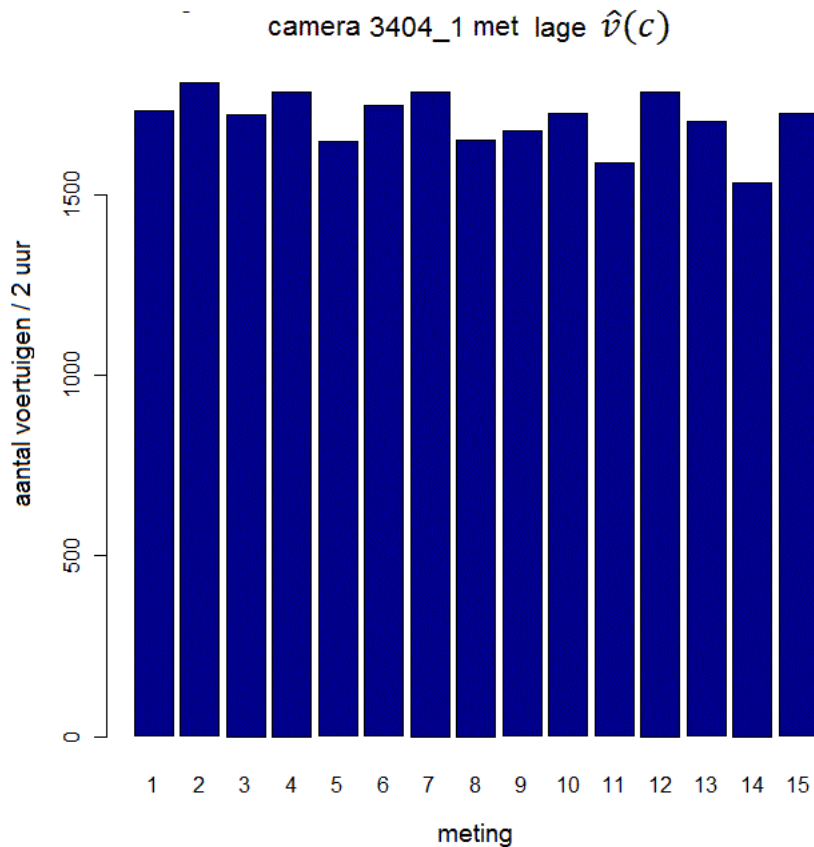
Omdat het beschikbaarheidspercentage net als het herkenningpercentage ook een gemiddelde is van deze maand, kan het exacte beschikbaarheidspercentage van de meting 1, 2 en 3 ook niet bepaald worden.

Deelconclusie 4.1: Op basis van het gemiddelde herkenningpercentage en beschikbaarheidspercentage van een maand kan niet worden aangegeven wat het herkenningpercentage en beschikbaarheidspercentage van de camera voor een meting van 2 uur in die maand.

Toch kan de betrouwbaarheidsdata gebruikt worden om een verband te ontdekken tussen de herkenningpercentages en $\hat{v}(c)$. In figuur 25 en figuur 26 zijn respectievelijk de $m_i(3780_1)$ waarden en $m_i(3404_1)$ waarden gegeven.



Figuur 25: Camera met hoge $\hat{v}(c)$.



Figuur 26: Camera met lage $\hat{v}(c)$.

	3780 herkenningpercentage	3780 beschikbaarheidpercentage	3404 herkenningpercentage	3404 beschikbaarheidpercentage
maart	86%	99 %	95 %	100 %
april	78 %	99 %	94 %	99 %
mei	77 %	99 %	94 %	99 %
september	82 %	92 %	95 %	97 %
gemiddelde van alle maanden	81%	97%	95 %	99 %

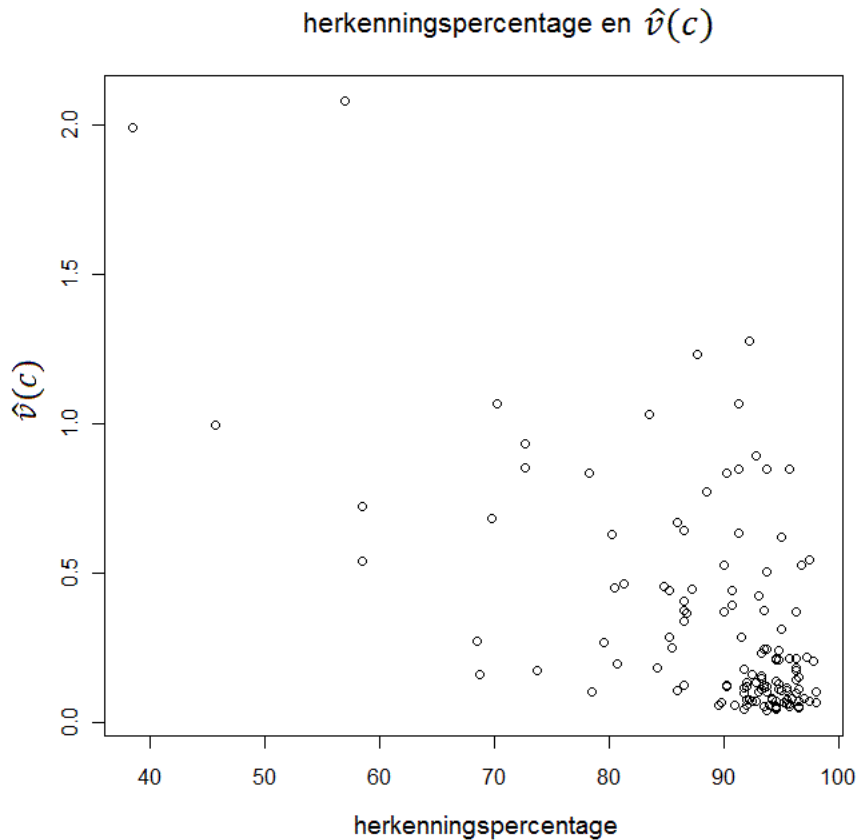
Tabel 4: Herkenningpercentages en beschikbaarheidpercentages van camera met hoge c_v en camera met lage c_v .

In tabel 4 zijn de gemiddelde herkenningpercentages en gemiddelde beschikbaarheidpercentages van camera's 3780_1 en 3404_1 per maand weergegeven.

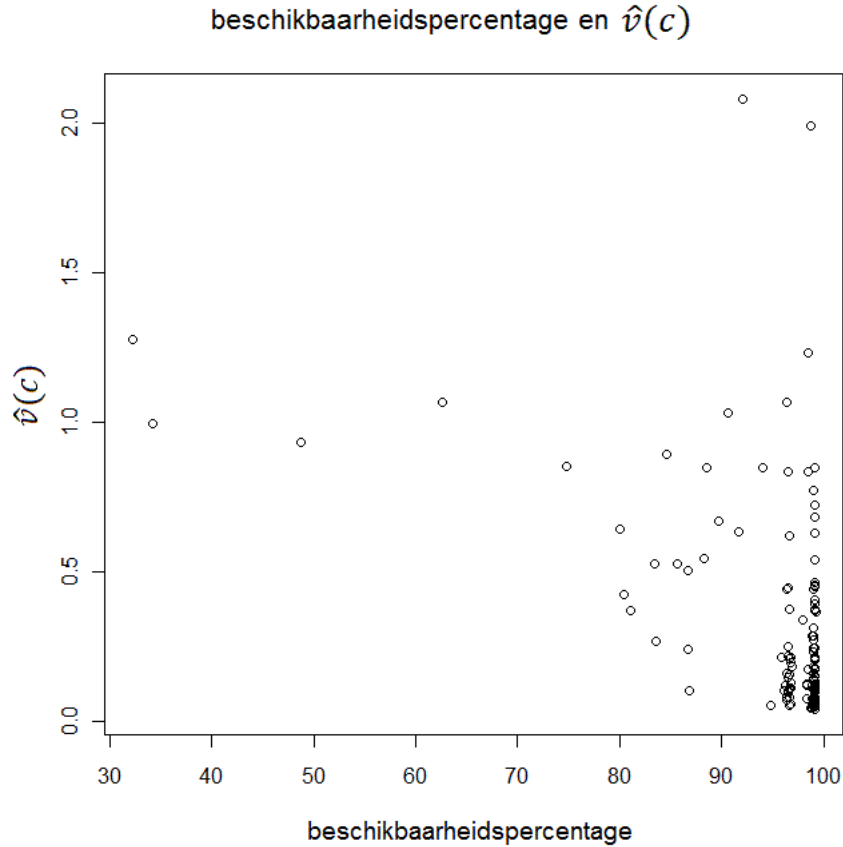
De $\hat{v}(c)$ waarden van deze camera's zijn $\hat{v}(3404_1) = 0,05$ en $\hat{v}(3780_1) = 0,2$. Er zou een verband kunnen zijn tussen het lagere gemiddelde herkenningpercentage van camera 3780_1 en de relatief hoge $\hat{v}(3780_1)$ waarde.

Een verklaring hiervoor is dat bij een hoog gemiddeld herkenningpercentage als 95% per definitie niet sprake kan zijn van een hoge spreiding in herkenningpercentages, terwijl bij een gemiddeld laag herkenningpercentage de herkenningpercentages verspreid uit elkaar kunnen liggen.

We bekijken het verband tussen de gemiddelde herkenningpercentages van alle maanden en $\hat{v}(c)$. Ook bekijken we het verband tussen de gemiddelde beschikbaarheidspercentages van alle maanden en $\hat{v}(c)$. In figuur 27 is het gemiddeld herkenningpercentage van alle maanden van alle camera's uitgezet tegen $\hat{v}(c)$. In figuur 28 zijn de gemiddelde beschikbaarheidspercentages van alle maanden van alle camera's uitgezet tegen $\hat{v}(c)$.



Figuur 27: Verband herkenningpercentage en $\hat{v}(c)$.



Figuur 28: Verband beschikbaarheidspercentage en $\hat{v}(c)$.

De correlatiecoëfficiënt tussen $\hat{v}(c)$ en herkenningspercentage is -0,61. Tussen beschikbaarheidspercentage en $\hat{v}(c)$ is de correlatiecoëfficiënt -0,46.

In figuur 27 en figuur 28 is hetzelfde patroon terug te zien. Rond hoge waarden voor herkenningspercentage of beschikbaarheidspercentage is een hoge dichtheid rond $\hat{v}(c)$. Dit geeft weer wat we eerder zagen in paragraaf 3.3.2: de herkenningspercentages en beschikbaarheidspercentages zijn beide redelijk hoog. Toch is te zien dat rond de hoge waarden voor herkenningspercentage en beschikbaarheidspercentage nog hoge $\hat{v}(c)$ waarden te zien zijn. Het verschil is echter de grotere dichtheid rond de lage $\hat{v}(c)$ waarden.

Hoewel de verbanden te zien zijn, zijn de correlatiecoëfficiënten niet hoog. Dit heeft er waarschijnlijk mee te maken dat de betrouwbaarheidsindicatoren gemiddelden zijn per maand, en niet voor de twee uur van de metingen. Verwacht wordt dat wanneer wordt gewerkt met beschikbaarheidspercentages en herkenningspercentages per uur, dat de correlatiecoëfficiënten hoger zullen zijn.

Toch kan gesteld worden:

Deelconclusie 4.2: Er is een negatieve correlatie tussen de betrouwbaarheids-indicatoren herkenningspercentage en beschikbaarheidspercentage, en $\hat{v}(c)$.

4.4 Methode

4.4.1 Definitie $q_m(l)$

Het doel van het onderzoek is om te onderzoeken in hoeverre de kentekenregistraties geschikt zijn om intensiteiten te bepalen die geschikt zijn voor kalibratie van het verkeersmodel. Laat $q_m(l)$ een intensiteit zijn die gebruikt kan worden voor het verkeersmodel:

$q_m(l)$ = intensiteit op locatie l die gebruikt kan worden voor het verkeersmodel op basis van $m_i(c)$ waarden, waarbij camera c op locatie l staat.

De definitie is niet exact, omdat er geen vaste eisen zijn waar een intensiteit die gebruikt wordt voor het verkeersmodel aan moet voldoen. Met Verkeersonderzoek zijn wel een aantal criteria opgesteld:

- De $m_i(c)$ waarden die worden gebruikt voor bepaling van het verkeersmodel $q_m(l)$ moeten zo klein mogelijke verschillen hebben met werkelijke intensiteiten op die metingen;
- Wanneer een $m_i(c)$ waarde laag is door een bijzondere verkeersomstandigheid (bijvoorbeeld een ongeval op de locatie van de camera), dan moet deze $m_i(c)$ waarde niet gebruikt worden voor de bepaling van $q_m(l)$.

Bij het toepassen van de definitie van $q_m(l)$ moet onderscheid gemaakt worden tussen locaties waar alle rijstroken bemeten zijn en locaties waar niet alle rijstroken bemeten zijn. Wanneer alle rijstroken bemeten zijn, wordt de $q_m(l)$ waarde gebaseerd op de $m_i(c)$ waarden van alle camera's c op locatie l . In het geval dat niet alle rijstroken bemeten zijn wordt de $q_m(l)$ waarde gebaseerd op de $m_i(c)$ waarden van de aanwezige camera's c op locatie l . Voor de rijstroken die niet worden bemeten zal nu een schatting gemaakt worden van de intensiteit. Hierop zal verder worden ingegaan in hoofdstuk 5.

4.4.2 Bepalen $q_m(l)$

We hebben nu gesteld dat we de $q_m(l)$ waarde gaan baseren op de $m_i(c)$ waarden. Hoe we de $q_m(l)$ waarden gaan bepalen zullen we nu bekijken.

Laten we de definitie van ε bekijken:

$$\varepsilon = Q_g\left(\frac{1}{HB} - 1\right).$$

Op basis van deelconclusie 4.2 kan gesteld worden dat de $\hat{v}(c)$ waarden een negatieve correlatie met het herkenningspercentage en het beschikbaarheidspercentage hebben. Hoe hoger $\hat{v}(c)$, hoe lager de herkenningspercentages en

beschikbaarheidspercentages zijn. Hoe lager de herkenningspercentages en beschikbaarheidspercentages zijn, hoe hoger ε is. Eerder zagen we:

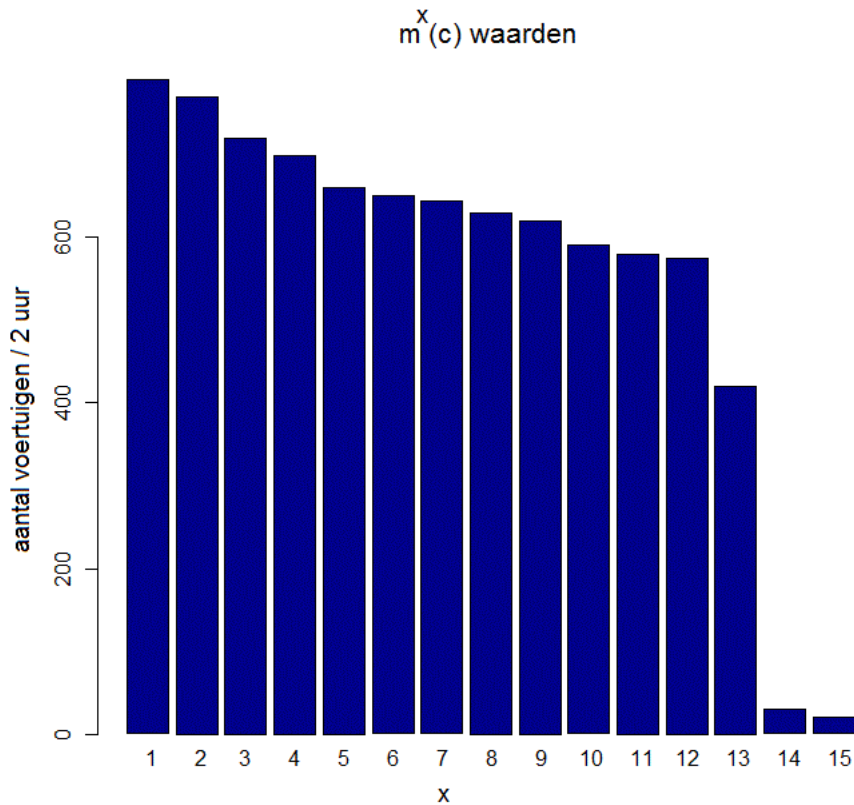
$$Q_g = Q_w - \varepsilon.$$

Hoe hoger ε is, hoe groter het verschil tussen Q_g en Q_w . Omdat we aannemen dat de $m_i(c)$ waarden uit de verdeling van Q_g komen, zullen de $m_i(c)$ waarden meer verschillen van de werkelijke intensiteiten bij lage herkennings- en betrouwbaarheidspercentages. Laten we nu een definitie geven voor een gesorteerde $m_i(c)$ waarde:

$m^x(c)$ = de x -de $m_i(c)$ waarde in de rij waarin $m_i(c)$ waarden zijn gesorteerd van hoog naar laag

Dit betekent dat $m^1(c)$ en $m^{15}(c)$ respectievelijk de hoogste en laagste $m_i(c)$ waarden zijn.

Bekijk figuur 29, waarin $m^x(c)$ waarden van een camera zijn weergegeven.



Figuur 29: $m^x(c)$ waarden.

Bij elke $m^x(c)$ waarde is er een onbekende werkelijke intensiteit, onbekend herkenningspercentage en onbekend beschikbaarheidspercentage. Dat de herkennings- en beschikbaarheidspercentages onbekend zijn volgt uit deelconclusie 4.1. Deze geeft namelijk aan dat we de gemiddelde herkennings- en beschikbaarheidspercentages per maand niet kunnen gebruiken om de herkennings- en beschikbaarheidspercentages te bepalen voor de $m_i(c)$ waarden. Dit geldt ook voor de $m^x(c)$ waarden, omdat de $m^x(c)$ waarden op de $m_i(c)$ waarden gebaseerd zijn.

De verschillen in combinatie van de werkelijke intensiteit, herkennings- en beschikbaarheidspercentage zorgen voor verschillen in $m^x(c)$ waarden. Merk op dat een $m^x(c)$ waarde bijvoorbeeld een hogere werkelijke intensiteit kan hebben dan een andere $m^x(c)$ waarde, terwijl de eerst genoemde $m^x(c)$ waarde alsnog lager is. Een $m^x(c)$ waarde met een werkelijke intensiteit van 600 en herkennings- en beschikbaarheidspercentages van respectievelijk 75% en 100% heeft de volgende waarde:

- $m^x(c) = 600 \cdot 0,75 \cdot 1 = 450$.

Vergelijk dit met een $m^x(c)$ waarde met een werkelijke intensiteit van 500 en herkennings- en beschikbaarheidspercentages van respectievelijk 100%:

- $m^x(c) = 500 \cdot 1 \cdot 1 = 500$.

De $m^x(c)$ waarde hangt dus af van de combinatie van werkelijke intensiteit, herkennings- en beschikbaarheidspercentage. De $m^x(c)$ waarden met een hoge x (bijvoorbeeld $x = 1$ tot en met $x = 5$) hebben een hogere combinatie van werkelijke intensiteit, herkennings- en beschikbaarheidspercentage dan de $m^x(c)$ waarden met een lage x (bijvoorbeeld $x = 11$ tot en met $x = 15$).

Laten we kijken naar de $m^x(c)$ waarden met een hoge x . Nu zijn er verschillende situaties mogelijk:

- De combinatie van herkenningspercentage, beschikbaarheidspercentage en werkelijke intensiteit zorgt ervoor dat de $m^x(c)$ waarden hoger zijn dan de gemiddelde werkelijke intensiteit over de gemeten periode;
- De combinatie van herkenningspercentage, beschikbaarheidspercentage en werkelijke intensiteit zorgt ervoor dat de $m^x(c)$ waarden lager zijn dan de gemiddelde werkelijke intensiteit over de gemeten periode.

Wat precies het geval is bij de $m^x(c)$ waarden met een hoge x is niet te bepalen. Maar door $q_m(l)$ te baseren op $m^x(c)$ waarden met een hoge x kan ervoor gezorgd worden dat meetfouten voor zo veel mogelijk niet worden meegenomen in $q_m(l)$.

Een ander argument om $q_m(l)$ te baseren op $m^x(c)$ waarden met een hoge x , is dat op deze manier $m^x(c)$ waarden niet worden meegenomen, waarbij er een lage werkelijke intensiteit is die voornamelijk bepaald wordt door een extreme verkeersomstandigheid. Zoals eerder aangegeven willen we $q_m(l)$ niet baseren op $m^x(c)$ waarden die laag zijn door een extreme verkeersomstandigheid.

Daarom stellen we de volgende vuistregel voor.

Vuistregel Bepaal $q_m(l)$ op basis van het gemiddelde van m^1, \dots, m^n en kies n laag.

Voor deze vuistregel zullen we nu een aantal voorbeelden geven.

Voorbeeld 4.2: Een weg bestaat uit één rijstrook met een MoCo-camera. We kiezen $n = 3$. $q_m(l)$ is nu gelijk aan het gemiddelde van $m^1(c), \dots, m^3(c)$.

Voorbeeld 4.3: Een weg bestaat uit twee rijstroken. Op beide rijstroken staat een MoCo-camera. We kiezen $n = 3$. $q_m(l)$ is nu gelijk aan het gemiddelde van $m^1(c), \dots, m^3(c)$ van de camera op de rechter rijstrook plus het gemiddelde van $m^1(c), \dots, m^3(c)$ van de camera op de linker rijstrook.

Voorbeeld 4.4: Op een weg staan twee rijstroken. Alleen op de rechter rijstrook staat een MoCo-camera. We kiezen $n = 3$. $q_m(l)$ is nu gelijk aan het gemiddelde van $m^1(c), \dots, m^3(c)$ plus een voorspelling van de intensiteit op de linker rijstrook. Of dit een betrouwbare intensiteit oplevert zal worden behandeld in hoofdstuk 5.

Aan de vuistregel kunnen we nog een restrictie toe voegen. Wanneer we ervoor kiezen om een n te nemen groter dan 1, dan moeten we ervoor zorgen dat $\hat{v}(c)$ van de n $m^x(c)$ waarden niet te hoog is. Wanneer $\hat{v}(c)$ van de n $m^x(c)$ waarden hoog is, dan is dat een indicatie dat er mogelijk alsnog te veel fouten zitten in n $m^x(c)$ waarden. Wanneer we een gemiddelde nemen van de n $m^x(c)$ waarden, verwachten we een gemiddelde te nemen van een aantal metingen die een hoge betrouwbaarheid lijken te hebben. Daarom zouden deze n $m^x(c)$ waarden dichtbij elkaar moeten liggen. Wanneer de $\hat{v}(c)$ waarde van deze n $m^x(c)$ waarden hoog is, dan lijken de n $m^x(c)$ waarden niet betrouwbaar.

Het getal n is niet precies te bepalen. In hoofdstuk 6 zullen we verschillende waarden van n bekijken.

Voordat tot deze vuistregel is gekomen zijn een aantal alternatieve methodes overwogen. Echter, deze alternatieve methodes voegen niks toe aan deze methode. Deze methodes staan kort beschreven in Appendix A.

5 Voorspelling meerdere rijstroken

In dit hoofdstuk wordt een model opgesteld om de intensiteit te bepalen op wegen met twee rijstroken, maar waarvan slechts één rijstrook bemeten wordt. De rijstrook die bemeten is, is op een paar uitzonderingen na de rechter rijstrook. Ook zijn er meetlocaties met wegen met meer dan twee rijstroken, waarvan niet alle rijstroken bemeten zijn. In dit onderzoek is gebleken dat er niet genoeg informatie is om hier een voorspelling voor te doen. Dit zal later in dit hoofdstuk toegelicht worden.

In paragraaf 4.4.1 zagen we dat $q_m(l)$ de intensiteit is op een locatie die geschikt is voor het verkeersmodel. Op wegen waar deze locatie twee rijstroken heeft is $q_m(l)$ de som van de intensiteiten van de linker en rechter rijstrook. In het verkeersmodel kunnen geen intensiteiten voor rijstroken worden ingevoerd. Intensiteiten kunnen alleen voor een bepaalde weg worden ingevoerd. Maar omdat de camera's rijstroken bemeten, zijn de intensiteiten wel per rijstrook aanwezig. Als het in het verkeersmodel mogelijk was geweest om de intensiteit per rijstrook in te voeren, dan was dit met de camera's mogelijk. De intensiteit voor het verkeersmodel op de linker rijstrook zouden we dan definiëren met $q_{m,links}(l)$ en de intensiteit op de rechter rijstrook zouden we definiëren met $q_{m,rechts}(l)$. Merk op dat we $q_{m,links}(l)$ en $q_{m,rechts}(l)$ niet in het huidige verkeersmodel kunnen invoeren. Toch hebben we deze variabelen gedefinieerd omdat deze variabelen ons helpen om een voorspelling te kunnen doen voor $q_m(l)$ op locaties met twee rijstroken, waarvan slechts één rijstrook wordt bemeten. Voor $q_m(l)$ op deze locaties geldt dan:

$$q_m(l) = q_{m,links}(l) + q_{m,rechts}(l).$$

Vervolgens definiëren we p als:

$$p = \frac{q_{m,links}(l)}{q_{m,rechts}(l)}.$$

$p = 1$ geeft aan dat op de linker rijstrook net zo veel voertuigen rijden als op de rechter rijstrook. Wanneer p bekend is kan $q_{m,links}(l)$ als volgt bepaald worden:

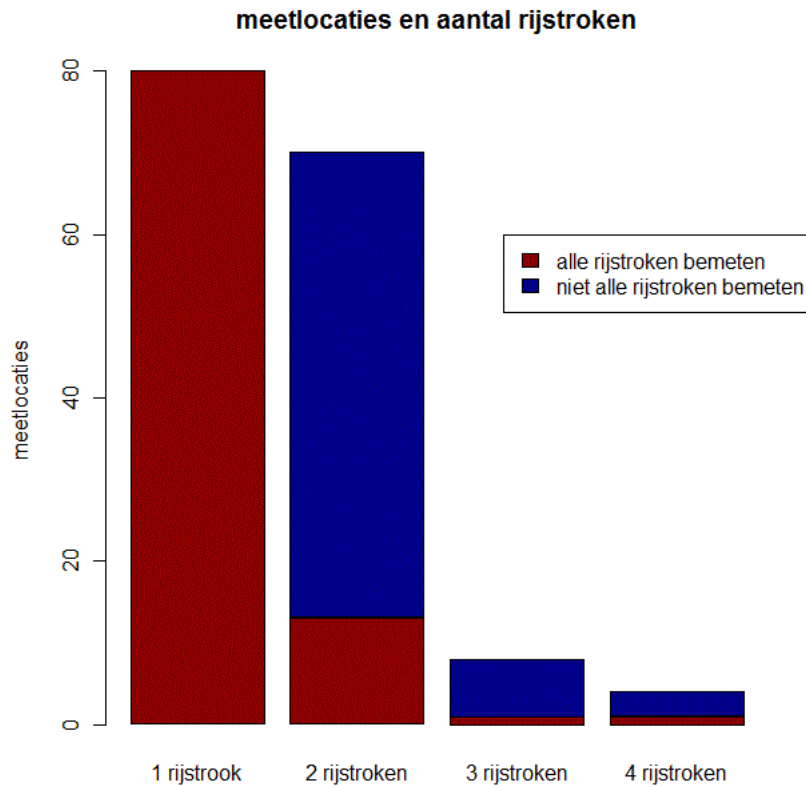
$$q_{m,links}(l) = p \cdot q_{m,rechts}(l).$$

De totale intensiteit op de weg kan dan geschreven worden als:

$$q_m(l) = p \cdot q_{m,rechts}(l) + q_{m,rechts}(l).$$

Wanneer p en $q_{m,rechts}(l)$ bekend zijn kan $q_m(l)$ worden bepaald. Dus wanneer slechts een rijstrook is bemeten en $q_{m,links}(l)$ onbekend is, kan met behulp van p toch $q_m(l)$ worden bepaald. Met behulp van lineaire regressie proberen we p te bepalen voor locaties waar slechts een rijstrook is bemeten. Dit is mogelijk wanneer een aantal meetlocaties met twee rijstroken aanwezig zijn, waarbij beide rijstroken zijn bemeten. Deze meetlocaties zijn aanwezig.

In figuur 30 is te zien hoeveel rijstroken de verschillende meetlocaties hebben. Het aantal rijstroken per meetlocatie is per locatie opgezocht in Google Maps. Van de 165 meetlocaties zijn er 70 meetlocaties met twee rijstroken. Van deze 70 meetlocaties zijn er 10 meetlocaties waarvan beide rijstroken bemeten zijn. Dit betekent dat we van 10 meetlocaties volledige informatie hebben. Op basis van deze 10 meetlocaties kunnen we een uitspraak doen over de meetlocaties waar niet alle rijstroken bemeten zijn.



Figuur 30: Verdeling aantal rijstroken over meetlocaties .

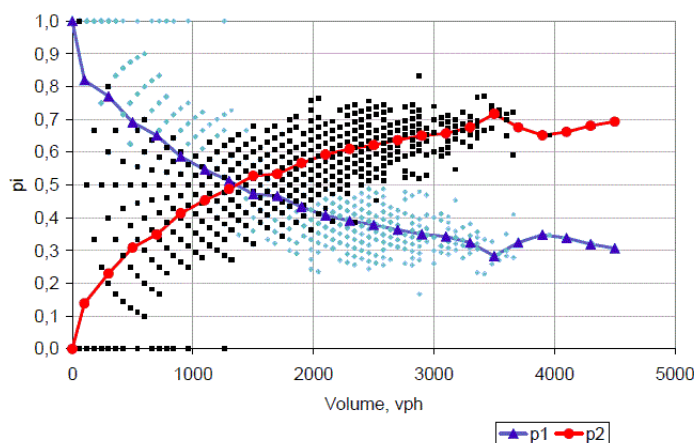
Bij het onderzoeken van de camera's is er een opvallend verschil tussen drie meetlocaties en de overige zeven locaties. Op deze drie meetlocaties staat de camera gericht op een voorsorteervak. Het is te verwachten dat op de voorsorteervakken de voertuigen zich anders zullen verdelen over de linker en rechter rijstrook dan wanneer dit niet het geval is. Op de voorsorteervakken verdelen voertuigen zich volgens de richting waar ze heen moeten.

Er zijn twee verklarende variabelen voor p gevonden. Dit zijn $q_{m,rechts}$ en de afstand meetpunt tot het vorige kruispunt. Met het meetpunt wordt het punt dat de voertuigen in beeld komen bedoeld. Deze twee verklarende variabelen zullen nu verder worden onderzocht en toegelicht.

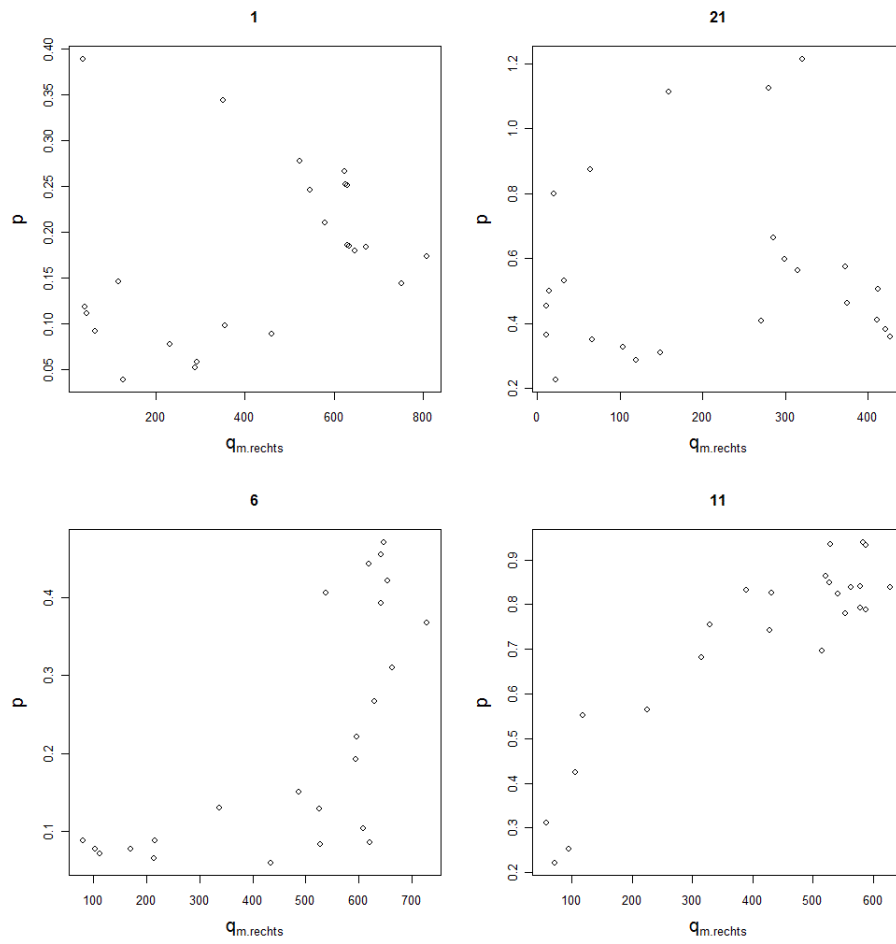
In de vorige hoofdstukken was $m_i(c)$ de avondspitsintensiteit op dinsdag, woensdag en donderdag voor de vijf weken. Voor deze analyse is het interessant om de gemeten intensiteiten te bekijken voor elk uur. Zo ontstaan er uiteindelijk grotere verschillen tussen de $q_{m,rechts}(l)$ waarden. Dit maakt het makkelijker om te onderzoeken of er een verband is tussen de $q_{m,rechts}(l)$ waarden en p . We definiëren $m_i(c)$ nu als de gemeten intensiteit in een bepaald uur. Voor de 15 dagen is voor elke periode van een uur in de dag de $m_i(c)$ waarde bepaald. De eerste periode is van 12 uur 's nachts tot 1 uur 's nachts. De tweede periode is van 1 uur 's nachts tot 2 uur 's nachts, etc. Zo ontstaan er voor elke periode van een uur 15 $m_i(c)$ waarden. Vervolgens is voor elke periode van een uur $q_{m,rechts}$ bepaald volgens de methode in paragraaf 4.4.2 door het gemiddelde te nemen van m^1, \dots, m^n met $n = 1$. Dit is gelijk aan de hoogste waarde te kiezen van de 15 $m_i(c)$ waarden.

5.1 $q_{m,rechts}(l)$ en p

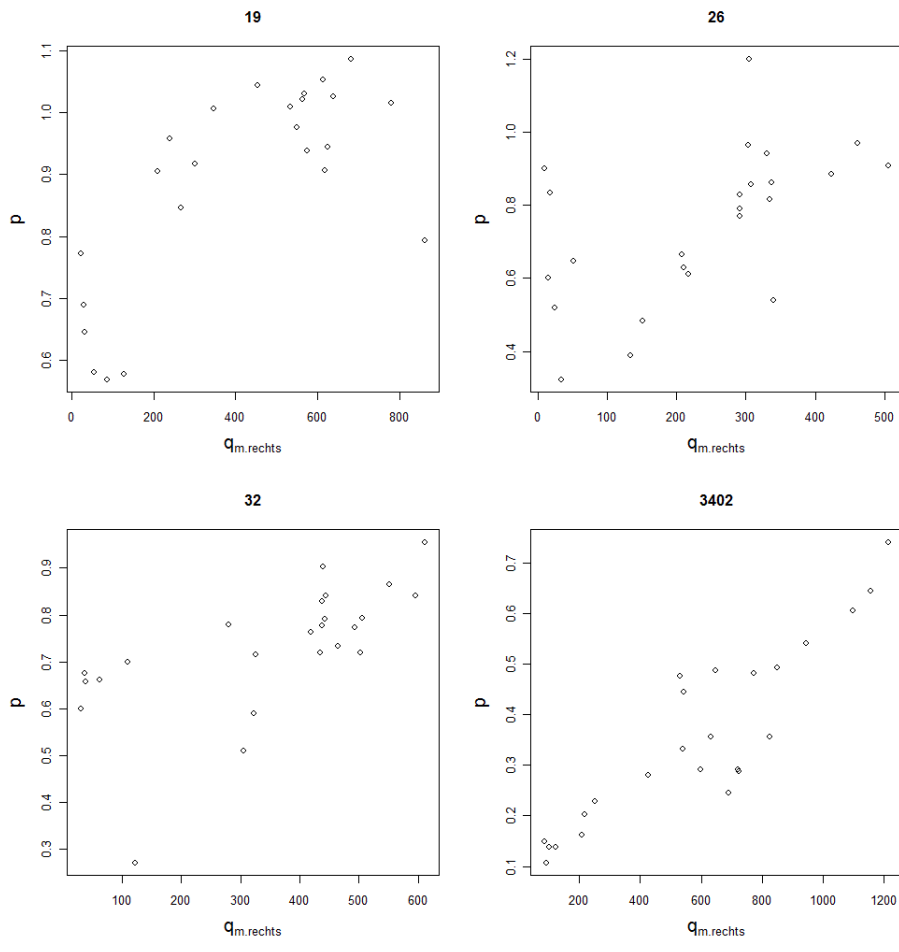
Wu [5] laat zien dat er een verband is tussen intensiteiten op de linker en rechter rijstrook op verschillende snelwegen in het buitenland. Een verklaring hiervoor zou kunnen zijn dat als het drukker wordt op de rechter rijstrook, meer auto's links gaan rijden om de weg optimaal te benutten. In figuur 31 is het verband te zien tussen intensiteit en de verhouding van de linker rijstrook, en rechter rijstrook. Dit verband was gemeten op een Duitse snelweg.



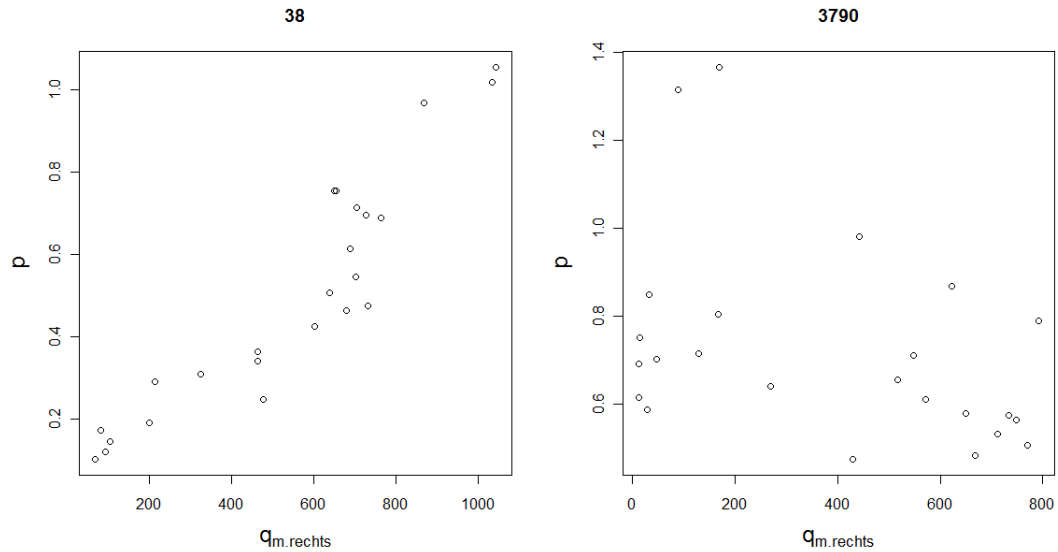
Figuur 31: Verband Duitse snelweg; de p_i waarde is anders dan p die wij hebben gedefinieerd; p_i is hier de intensiteit van rijstrook i als percentage van de totale intensiteit van de weg; bron. [5]



Figuur 32: Verbanden $q_{m,rechts}(l)$ en p voor de 10 meetlocaties (1 van 3).

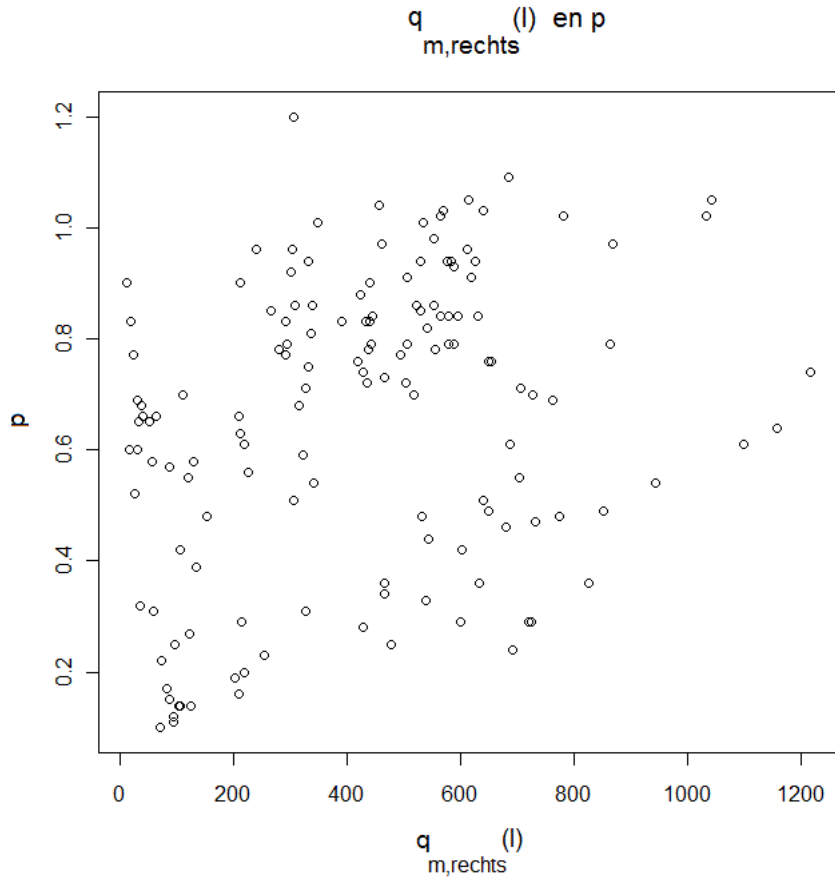


Figuur 33: Verbanden $q_{m,rechts}(l)$ en p voor de 10 meetlocaties (2 van 3).



Figuur 34: Verbanden $q_{m,rechts}(l)$ en p voor de 10 meetlocaties (3 van 3).

In figuur 32, figuur 33 en figuur 34 zien we voor de 10 camera's het verband tussen $q_{m,rechts}(l)$ en p . Voor de camera's op meetlocatie 1, meetlocatie 21 en meetlocatie 6 is geen lineair verband te zien. Op de overige meetlocaties met uitzondering van meetlocatie 3790 is wel een positief verband te zien. Er is voor gekozen om meetlocatie 1, meetlocatie 21 en meetlocatie 6 niet op te nemen in het model. Deze meetlocaties onderscheiden zich van de rest doordat het meetpunt op een voorsorteervak staat. Het regressiemodel is daarom alleen toepasbaar voor meetpunten die niet op een voorsorteervak staan. Verder vertoont meetlocatie 3790 een afwijkend patroon. Om een goed regressiemodel te krijgen is deze ook uit de dataset gefilterd. In figuur 35 is het algemeen verband tussen $q_{m,rechts}(l)$ en p weergegeven.



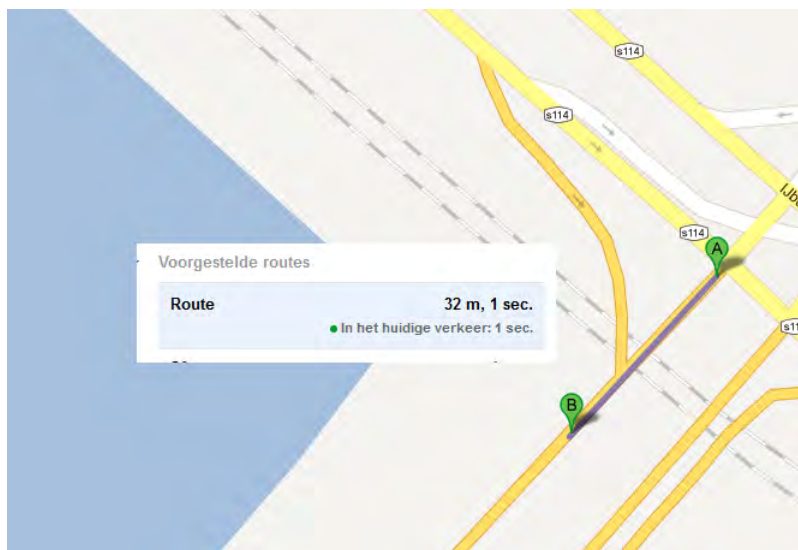
Figuur 35: Algemeen verband $q_{m,rechts}(l)$ en p

Er is een positief verband te zien tussen intensiteit en verhouding. De correlatiecoëfficiënt is 0,5.

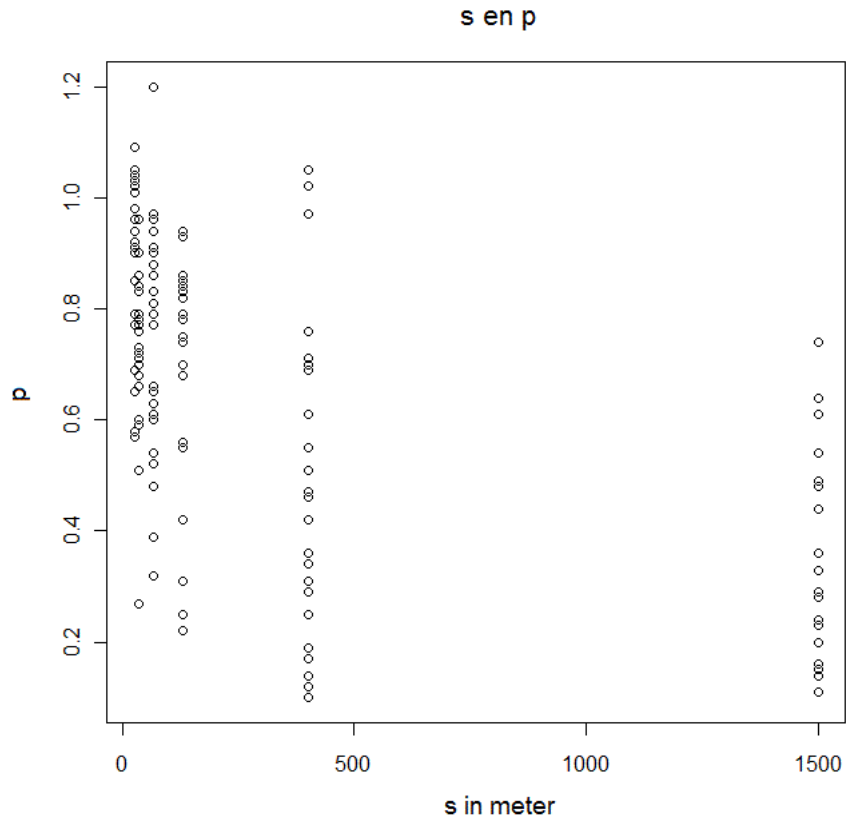
5.2 Afstand vorige kruispunt en verhouding

Voor elke locatie is de afstand van het vorige kruispunt tot de meetpunt van het camera genomen. De reden dat we deze afstand bekijken is omdat we vermoeden dat wanneer een camera dicht tot het vorige kruispunt staat, de kans groot is dat een voertuig op de linker en rechter rijstrook van het kruispunt tegelijk vertrekken en tegelijk de camera passeren. Wanneer de camera verder van het kruispunt staat, hebben de voertuigen op de linker rijstrook de tijd om zich naar de rechter rijstrook te verplaatsen. Omdat alleen de coördinaten van de camera gegeven zijn, hebben we Vialis gevraagd om een schatting te geven van de afstand tussen het meetpunt en de camera. Vialis heeft aangegeven dat dit

ongeveer 15 meter is. Voor de milieuzonecamera's houden we ook 15 meter aan. De afstand tot het vorige kruispunt is in Google Maps bepaald. De schatting tot het vorig kruispunt in Google Maps is niet heel precies. Elk kruispunt ziet er anders uit dus een exact punt tot dit kruispunt is niet te bepalen. Daarom moet deze afstand geïnterpreteerd worden als een schatting. In figuur 36 is hier een voorbeeld van gegeven.

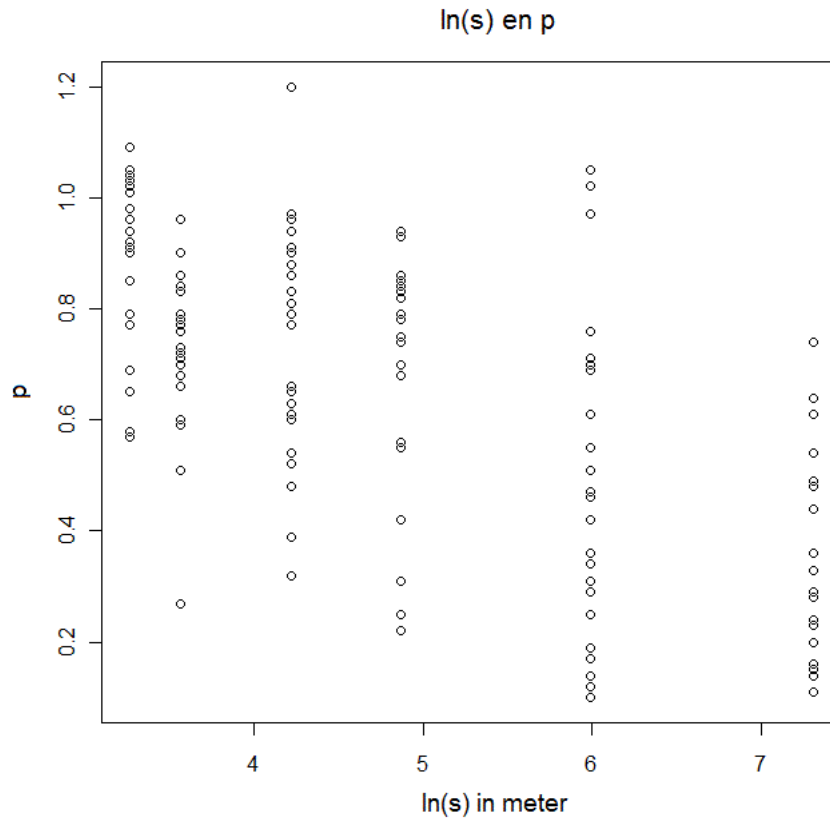


Figuur 36: Google Maps bepaling afstand vorig kruispunt.



Figuur 37: Verband s en p .

Laat s de afstand zijn van het meetpunt van de camera tot het vorig kruispunt. In figuur 37 is te zien dat afstand en verhouding een logaritmisch verband lijken te vertonen. Wanneer de afstand tot het kruispunt groter wordt, heeft een verhoging in deze afstand minder effect op p . Om de afstand tot het vorige kruispunt in het lineaire regressiemodel op te nemen, gaan we deze variabele lineariseren.



Figuur 38: Verband $\ln(s)$ en p

In figuur 38 is het verband te zien tussen $\ln(s)$ en p . Het verband tussen $\ln(s)$ en p lijkt lineair te zijn.

Op basis van bovenstaande verbanden zijn verschillende regressiemodellen bekeken. Ten eerste zijn de verbanden bekeken met één verklarende variabele:

$$p = 0,39 + 0,00037 \cdot q_{m,rechts}(l),$$

$$p = 0,68 - 0,00028 \cdot s,$$

$$p = 1,14 - 0,11 \cdot \ln(s).$$

De waarden voor de verklaarde variantie R^2 waren respectievelijk 0.25, 0.35, en 0.43. Het beste model met een verklarende variabele is met $\ln(s)$. Op deze manier kan 43% van de variantie van de gebruikte dataset worden verklaard.

Vervolgens zijn de modellen met beide verklarende variabelen bekeken:

$$p = 0,46 + 0,00046 \cdot q_{m,rechts}(l) - 0,00033 \cdot s,$$

$$p = 0,98 + 0,00047 \cdot q_{m,rechts}(l) - 0,13 \cdot \ln(s).$$

De waarden voor R^2 waren respectievelijk 0.72 en 0.8. Het beste model is het model met $q_{m,rechts}(l)$ en $\ln(s)$. Dit model verklaart 80% van de variantie in de dataset.

5.3 Meer dan twee rijstroken

In figuur 30 is te zien dat de er 13 meetlocaties zijn met meer dan twee rijstroken. Van 2 van deze 13 is de volledige informatie bekend. Dit aantal is te weinig om lineaire regressie op toe te passen. Daarnaast staan van veel van deze 13 meetlocaties het meetpunt op een voorsorteervak. Zoals we eerder zagen kunnen we het regressiemodel niet toepassen op deze locaties. Voor de meetlocaties met meer dan twee rijstroken en waarvan niet al deze rijstroken bemeten zijn, zal daarom geen oplossing worden gegeven.

6 Resultaten

In dit hoofdstuk worden de resultaten weergegeven van het toepassen van de filtermethode en de regressiemethode in de vorige twee hoofdstukken. De $q_m(l)$ waarden die bepaald zijn door toepassing van deze methodes zijn vergeleken met eerder uitgevoerde intensiteitsmetingen op de exacte meetlocaties. Het vergelijken van onze waarden voor $q_m(l)$ met deze tellingen zal niet leiden tot exacte uitspraken over de betrouwbaarheid van onze gevonden waarden voor $q_m(l)$. De eerder uitgevoerde intensiteitsmetingen zijn namelijk ook steekproeven en kunnen meetfouten bevatten. Bovendien moet bij het interpreteren van de vergelijking rekening mee worden gehouden dat de intensiteitsmetingen waarmee we onze waarden voor $q_m(l)$ vergelijken, metingen zijn over een andere periode dan onze waarden van $q_m(l)$. Echter geeft de vergelijking wel een indicatie van de betrouwbaarheid van de $q_m(l)$ waarden.

Er wordt gebruik gemaakt van twee tellingen:

Periodieke Visuele Tellingen Elk jaar is hier een meting uitgevoerd op een avondspits op een dinsdag, woensdag of donderdag. De metingen die wij gebruiken komen uit het jaar 2008.

Grontmij tellingen Gedurende een week is 24 uur lang gemeten op locaties in het najaar 2011. Voor vergelijking met deze telling hebben we een gemiddelde genomen van de avondspits op de dinsdag, woensdag en donderdag.

De Grontmij tellingen zijn recenter dan de PerVisTel telling en zouden de voorkeur moeten krijgen. Er zijn twee redenen waarom de PerVisTel data toch zijn gebruikt:

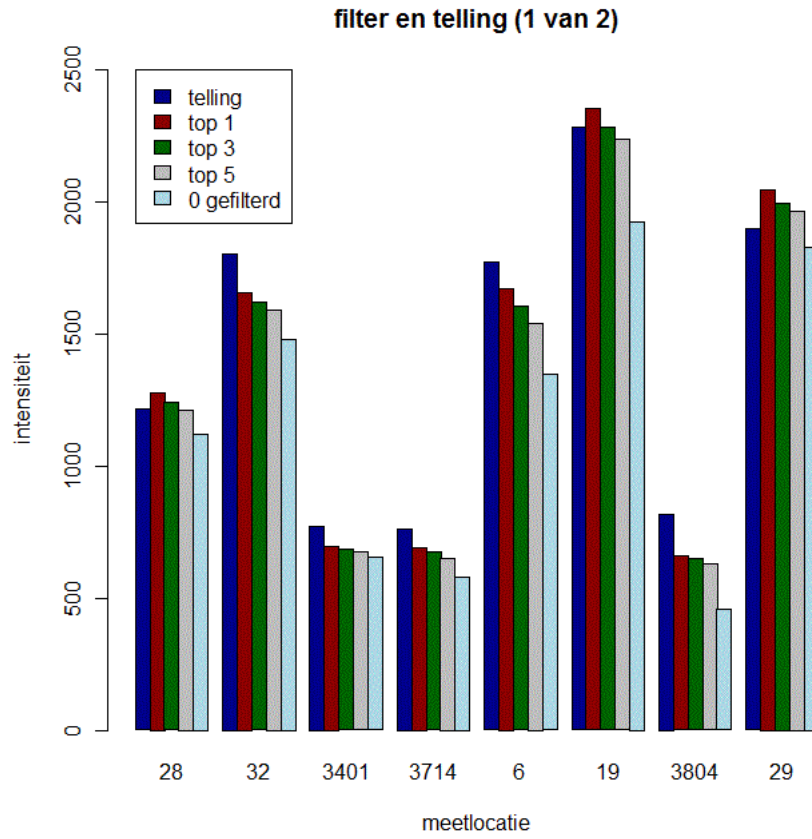
- Een aantal Grontmij tellingen op PerVistTel punten konden niet op tijd worden geleverd;
- Op een aantal PerVisTel punten is geen Grontmij telling uitgevoerd.

De resultaten zijn opgedeeld in twee delen: In eerste deel worden de resultaten gegeven van de meetlocaties waar alle rijstroken zijn bemeten. Dit zijn alle meetlocaties met één rijstrook en meetlocaties met meerdere rijstroken, waarvan alle rijstroken bemeten zijn. In het tweede deel worden de resultaten gegeven van de meetlocaties met twee rijstroken, waarvan slechts één rijstrook is bemeten.

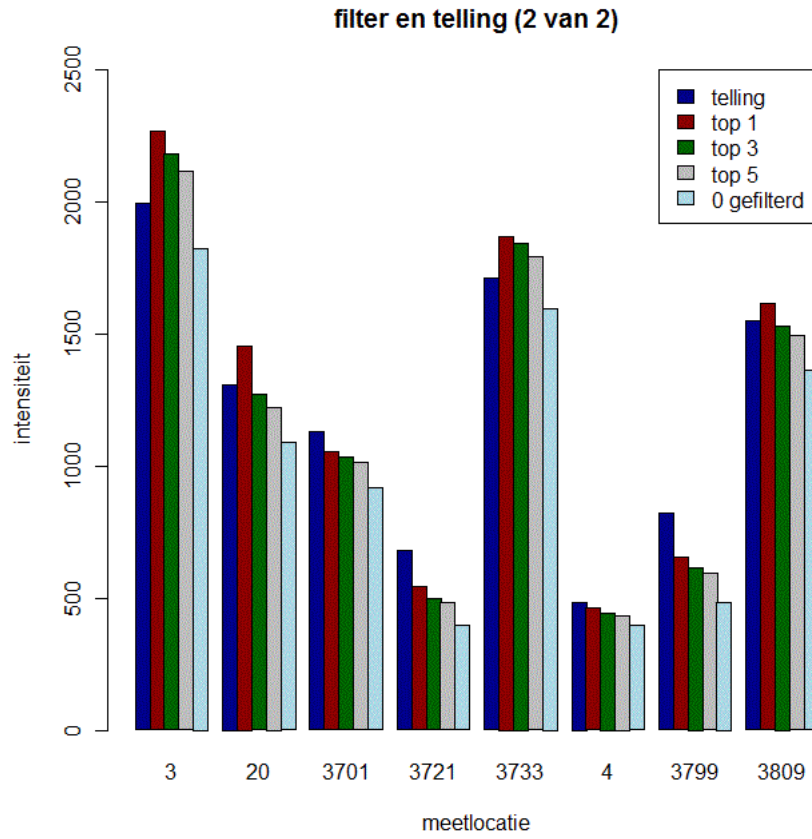
6.1 Alle rijstroken bemeten

Meetlocatie	Straatnaam	Pervistel	Grontmij	telling	top 1	top 3	top 5	0 gefilterd
28	Spaklerweg	1064	1216	1216	1278	1241	1212	1121
32	Pietheijntunnel	1805	-	1805	1659	1620	1592	1481
3401	Kattenburgerstraat	700	774	774	696	684	674	654
3714	Houtmankade	763	-	763	693	674	651	580
6	Jan van Galenstraat	1771	-	1771	1672	1605	1541	1346
19	Amstelveenseweg	2317	2282	2282	2353	2282	2237	1926
3804	Wielingenstraat	819	-	819	660	649	628	458
29	Nieuwe Utrechtseweg	1897	-	1897	2047	1997	1967	1829
3	Haarlemmerweg	1995	-	1995	2268	2181	2115	1825
20	Parnassusweg	1177	1309	1309	1455	1270	1223	1090
3701	Van Diemenstraat	1148	1130	1130	1054	1034	1014	916
3721	Nassaukade	679	-	679	542	497	482	397
3773	IJburglaan	1710	-	1710	1869	1845	1791	1597
4	Wiltzanghlaan	529	485	485	461	443	433	396
3799	Middenweg	822	-	822	654	617	593	484
3809	Klaprozenweg	1550	-	1550	1614	1531	1493	1363

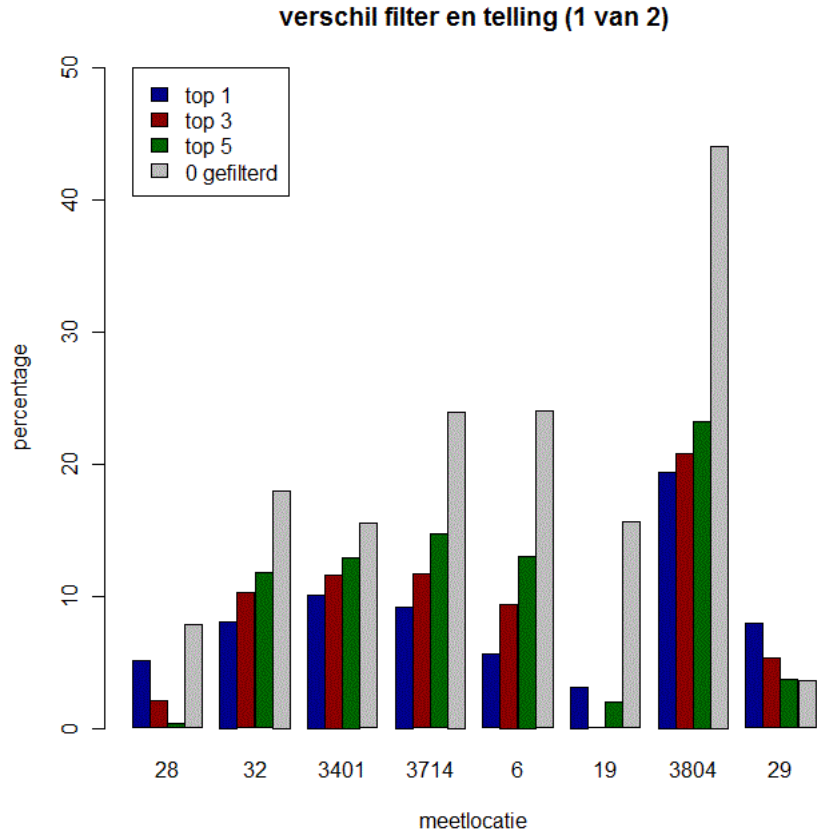
Tabel 5: Meetlocaties met Pervistel telling, Grontmij tellingen en verschillende waarden van $q_m(l)$ voor locaties met alle rijstroken bemeten.



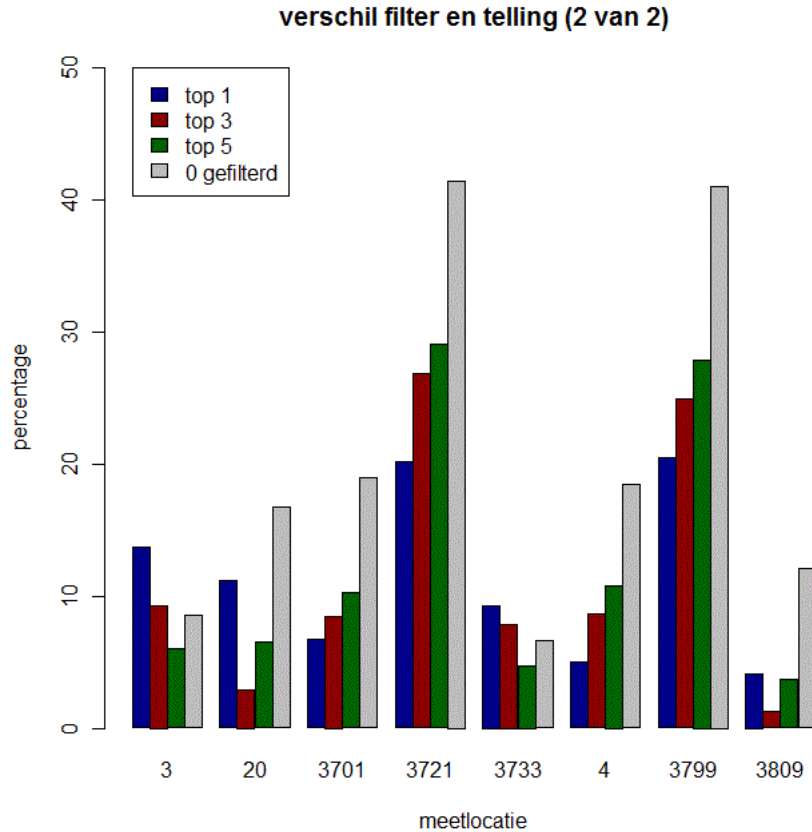
Figuur 39: Vergelijking telling en $q_m(l)$ voor locaties met alle rijstroken bemeten (1 van 2).



Figuur 40: Vergelijking telling en $q_m(l)$ voor locaties met alle rijstroken bemeten (2 van 2).



Figuur 41: percentuele verschillen telling en $q_m(l)$ voor locaties met alle rijstroken bemeten (1 van 2).



Figuur 42: Percentuele verschillen telling en $q_m(l)$ voor locaties met alle rijstroken bemeten (2 van 2).

$q_m(l)$	Gemiddelde afwijking telling
top 1	10%
top 3	10%
top 5	11%
0 gefilterd	20%

Tabel 6: Gemiddelde percentuele afwijking telling en $q_m(l)$ voor locaties met alle rijstroken bemeten.

In tabel 5 zijn de Pervistel en Grontmij tellingen weergegeven. Op de locaties waar Grontmij tellingen aanwezig zijn, wordt deze waarde gebruikt om $q_m(l)$ mee te vergelijken. Op de locaties waar de Grontmij tellingen ontbreken wordt de PerVisTel telling gebruikt. De kolom telling geeft de telwaarde aan die wordt

gebruikt. Daarnaast zijn ook waarden voor $q_m(l)$ te zien. Voor $q_m(l)$ bepalen we per locatie drie verschillende waarden. We nemen het gemiddelde van m^1, \dots, m^n met $n = 1, n = 3$, en $n = 5$. Deze geven we respectievelijk aan met top 1, top 3 en top 5. Ook is het gemiddelde genomen van alle $m_i(c)$ waarden, waarbij de waarden $m_i(c) = 0$ niet in het gemiddelde zijn opgenomen. Op deze manier wordt de intensiteit verkregen waarbij alleen de duidelijkste fouten niet worden meegenomen in de intensiteit. Deze intensiteit geven we aan met 0 gefilterd.

In figuur 39 en figuur 40 zijn de intensiteiten met elkaar vergeleken. Te zien is dat de waarden van top 1, top 3 en top 5 dicht in de buurt liggen van telling. De percentuele verschillen zijn weergegeven in figuur 41 en figuur 42. Hier is duidelijk te zien dat waarde 0 gefilterd een aantal keer aanzienlijk hoger ligt dan de overige intensiteiten.

In tabel 6 zijn de gemiddelde verschillen in procenten weergegeven. top 1, top 3 en top 5 zijn redelijk laag met hun percentages rond de 10%. Waarde 0 is met 20% duidelijk hoger dan de rest.

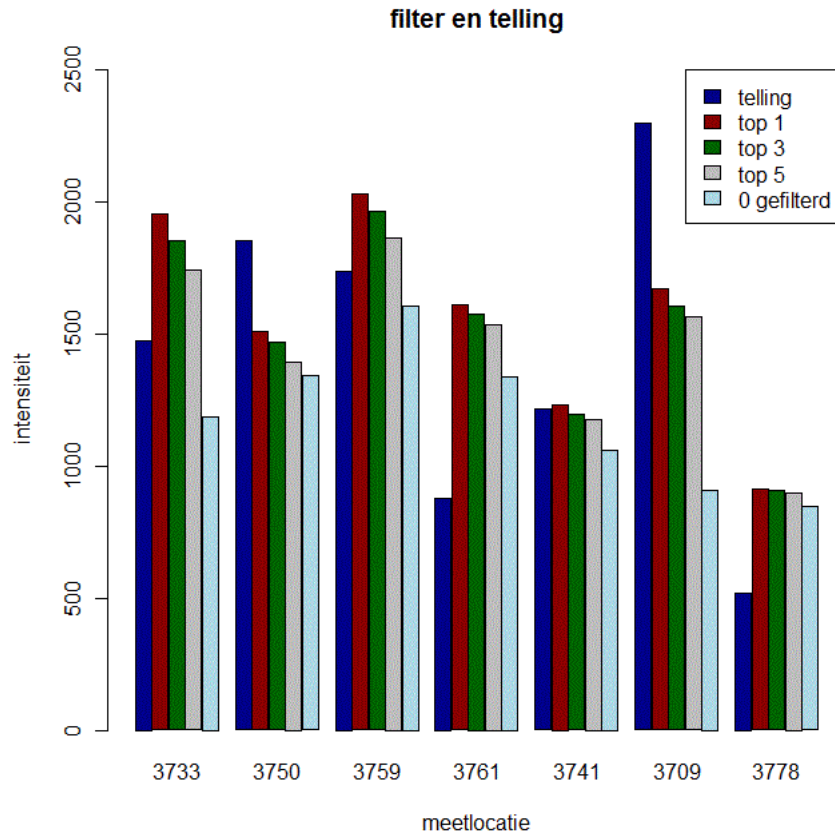
6.2 Twee rijstroken, waarvan een bemeten

Meetlocatie	Straatnaam	Pervis- tel	Gront- mij	telling	top 1	top 3	top 5	0 gefil- terd
3733	Stadhouderskade – s100	1477	-	1477	927	887	842	606
3750	Rhijnspoorplein – s112	1852	-	1852	725	708	677	656
3759	Gaasperdammerweg	1740	-	1740	977	951	911	804
3761	Oprit Gaasperdammerweg	879	-	879	805	792	774	688
3741	Europaboulevard – s109	1216	-	1216	641	624	616	562
3709	Basisweg – s102	2297	-	2297	851	823	805	505
3778	Verlengde Stellingweg – s118	520	-	520	524	521	517	490

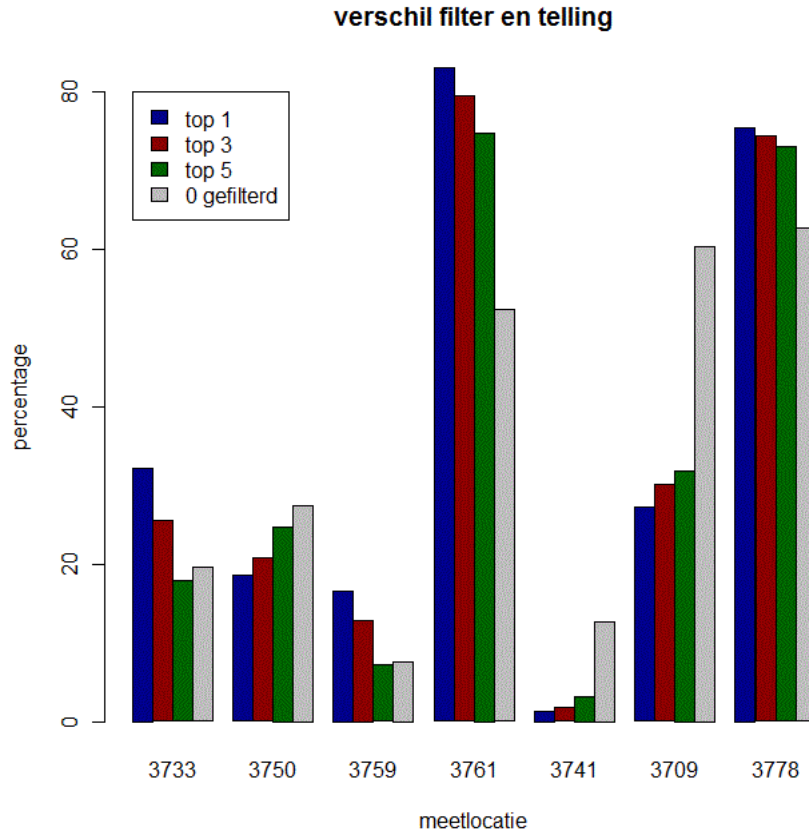
Tabel 7: Meetlocaties met Pervistel telling, Grontmij tellingen en verschillende waarden van $q_{m, rechts}(l)$.

Meetlocatie	Straatnaam	Pervis- tel	Gront- Mij	telling	top 1	top 3	top 5	0 gefil- terd
3733	Stadhouderskade – s100	1477	-	1477	1953	1853	1741	1188
3750	Rhijnspoorplein – s112	1852	-	1852	1508	1468	1394	1343
3759	Gaasperdammerweg	1740	-	1740	2030	1963	1865	1607
3761	Oprit Gaasperdammerweg	879	-	879	1609	1577	1537	1339
3741	Europaboulevard – s109	1216	-	1216	1233	1194	1177	1061
3709	Basisweg – s102	2297	-	2297	1671	1605	1564	910
3778	Verlengde Stellingweg – s118	520	-	520	913	907	899	846

Tabel 8: Meetlocaties met Pervistel telling, Grontmij tellingen en verschillende waarden van $q_m(l)$.



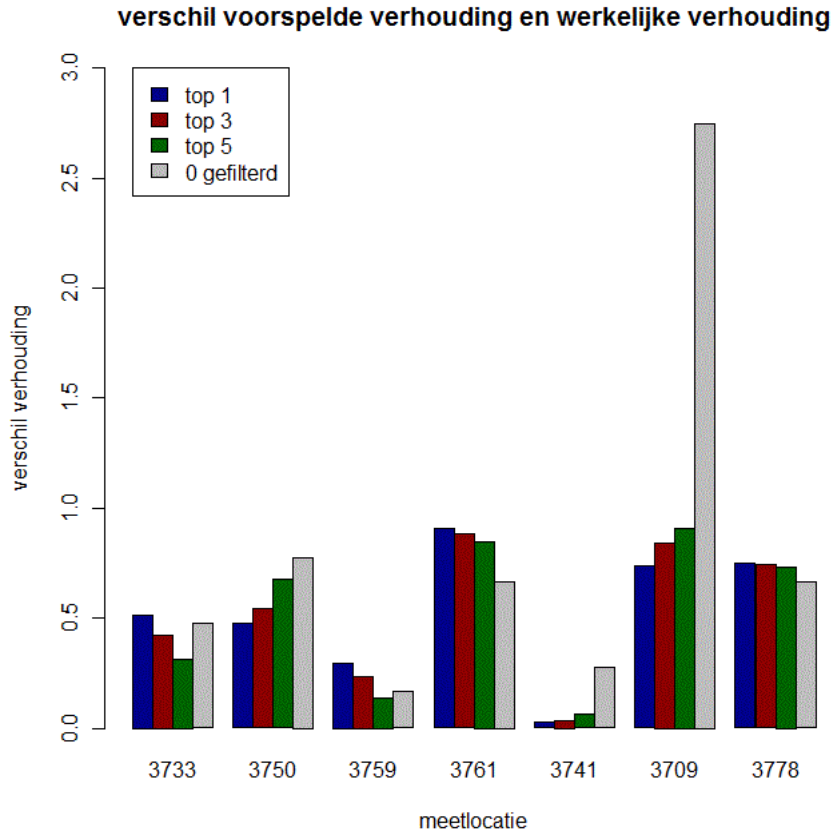
Figuur 43: Vergelijking telling en $q_m(l)$ voor locaties met twee rijstroken waarvan één bemeten.



Figuur 44: Percentuele verschillen telling en $q_m(l)$ voor locaties met alle twee rijstroken waarvan één bemeten.

$q_m(l)$	gemiddelde afwijking telling
top 1	36%
top 3	36%
top 5	35%
0 gefilterd	35%

Tabel 9: Gemiddelde percentuele afwijking telling en $q_m(l)$ voor locaties met twee rijstroken waarvan één bemeten.



Figuur 45: Verschillen voorspelde p met werkelijke p .

$q_m(l)$	gemiddelde afwijking p met telling
top 1	0,5
top 3	0,5
top 5	0,5
0 gefilterd	0,8

Tabel 10: Gemiddelde afwijking tussen p en werkelijke p .

In tabel 7 is hetzelfde weergegeven als in tabel 5, maar voor meetlocaties met twee rijstroken, waarvan één rijstrook bemeten is. Deze tabel bevat niet de $q_m(l)$ waarden maar de $q_{m,rechts}(l)$ waarden. In tabel 8 is dezelfde tabel weergegeven als tabel 7 maar in plaats van $q_{m,rechts}(l)$ waarden, bevat de tabel nu $q_m(l)$ waarden. $q_m(l)$ bevat de voorspelling voor de intensiteit op de linker rijstrook.

In figuur 43 zijn de intensiteiten uit tabel 8 met elkaar vergeleken. De verschillen zijn groter dan wanneer alle rijstroken bemeten waren. De verschillen zijn duidelijk te zien in figuur 44, die de percentuele verschillen weergeeft. In een aantal gevallen was het verschil met de telling bijna 80%. Een verschil met de telling van bijna 80% is redelijk hoog en de $q_m(l)$ waarde lijkt niet betrouwbaar genoeg voor het verkeersmodel.

De gemiddelde afwijkingen zijn weergegeven in tabel 9. De gemiddelde afwijkingen liggen allemaal rond de 35%. Deze gemiddelde afwijkingen kunnen we vergelijken met de gemiddelde afwijking van de locaties met alle rijstroken bemeten in tabel 6. Omdat de intensiteiten op meetlocaties met twee rijstroken, waarvan één bemeten, een voorspelling bevatten van de linker rijstrook, is te verwachten dat de gemiddelde afwijkingen hoger liggen dan voor de meetlocaties met alle rijstroken bemeten.

In figuur 45 zijn de verschillen tussen de voorspelde verhouding en de werkelijke verhouding weergegeven. De werkelijke verhouding is de verhouding zoals die had moeten zijn zodat $q_m(l)$ gelijk was aan de telling. In tabel 10 zijn de gemiddelden van deze verschillen tussen voorspelde verhouding en werkelijke verhouding weergegeven. We zien dat deze rond de 0,5 liggen. Een verschil van 0,5 tussen voorspelde verhouding en werkelijke verhouding geeft aan dat de voorspellingen niet heel exact zijn. Waarde 0 heeft een verschil van 0,8, maar zoals te zien is in figuur 45, wordt dit verschil grotendeels bepaald door één meting.

7 Conclusie en aanbevelingen

In dit hoofdstuk zal worden beschreven wat de conclusies zijn en welke aanbevelingen hieruit volgen. De conclusies volgen uit het beantwoorden van de twee deelvragen. Ten eerste zullen de twee deelvragen behandeld worden. Na het beantwoorden van de deelvragen zullen aanbevelingen worden gedaan. Deze aanbevelingen delen we op in aanbevelingen die direct volgen uit het onderzoek en algemene aanbevelingen.

7.1 Deelvraag 1

Deelvraag 1 was:

1. Hoe betrouwbaar zijn data uit MoCo-camera's en hoe kunnen eventuele fouten in deze data eruit worden gefilterd?

De data uit de MoCo-camera's bestaat uit herkende kentekens. Door het aantal herkende kentekens in een bepaalde periode bij elkaar op te tellen ontstaat er een gemeten intensiteit voor een bepaalde periode. Echter, niet alle kentekens van voertuigen die langs een camera rijden worden herkend. Daarom is de gemeten intensiteit niet gelijk aan het aantal voertuigen dat werkelijk de camera heeft gepasseerd in een bepaalde periode. De leverancier van de MoCo-camera's, Vialis, heeft een betrouwbaarheidsindicator bedacht voor het percentage van de voertuigen waarvan het kenteken herkend wordt: herkenningspercentage. Ook is er een betrouwbaarheidsindicator voor het percentage van de tijd dat de camera data genereert: beschikbaarheidspercentage. Gemiddelden van deze betrouwbaarheidsindicatoren zijn per maand per locatie geleverd.

Voor vijf weken op de dinsdag, woensdag en donderdag hebben we de gemeten intensiteit tijdens de avondspits bekeken voor alle camera's. Het is opgevallen dat gemeten intensiteiten van sommige camera's aanzienlijk meer verschillen van elkaar dan bij andere camera's. Er bleek een negatieve correlatie te zijn tussen de verschillen in gemeten intensiteiten van camera's en de herkenningspercentages en beschikbaarheidspercentages. Hoe hoger de herkenningspercentages en beschikbaarheidspercentages waren, hoe lager de verschillen tussen gemeten intensiteiten van een camera. Uit de aanwezigheid van deze verschillen tussen gemeten intensiteiten blijkt dat er meetfouten zitten in de data.

Om betrouwbare intensiteiten te bepalen voor kalibratie van het verkeersmodel moeten deze intensiteiten gebaseerd worden op zo min mogelijk meetfouten. Wanneer de gemeten intensiteiten per camera van hoog naar laag gesorteerd worden dan is de grootste kans op meetfouten bij de lagere gemeten intensiteiten. De hoogste gemeten intensiteiten hebben de grootste kans op hoge herkenningspercentages en beschikbaarheidspercentages. Daarom kan voor een betrouwbare intensiteit voor het verkeersmodel een gemiddelde worden genomen van de n hoogste gemeten intensiteiten. Hoeveel n precies moet zijn kan niet exact bepaald worden. Wel kunnen we aangeven dat we dit aantal niet te groot moeten nemen, om zo het aantal meetfouten zo klein mogelijk te houden.

Uit de resultaten bleek dat voor de meetlocaties waar alle rijstroken bemeten waren een gemiddelde nemen van n hoogste intensiteiten betrouwbare intensiteiten lijkt te geven. Dit zijn in totaal 95 van de 165 meetlocaties. We hebben een vergelijking gemaakt tussen intensiteiten voor het verkeersmodel op basis van de MoCo-camera's en eerder uitgevoerde tellingen. De intensiteiten voor het verkeersmodel verschilden gemiddeld 10% van de tellingen. We hebben intensiteiten voor het verkeersmodel bepaald met waarde van $n = 1$, $n = 3$ en $n = 5$. De intensiteiten met $n = 1$ zijn uiteraard wat hoger dan de intensiteiten met $n = 5$. Gemiddeld gezien was er echter geen groot verschil met de telling tussen de intensiteiten met $n = 1$ en de intensiteiten met $n = 5$. Wel is te zien dat als een gemiddelde wordt genomen van alle gemeten intensiteiten, waarbij de gemeten intensiteiten die nul waren niet worden meegenomen in dit gemiddelde, dat deze gemiddelde intensiteiten meer verschillen van de oude tellingen. Gemiddeld was het verschil nu 20%. Het bepalen van betrouwbare intensiteiten voor het verkeersmodel is dus mogelijk door een gemiddelde te nemen van de n hoogste gemeten intensiteiten.

7.2 Deelvraag 2

Deelvraag 2 was:

2. Hoe kunnen intensiteiten worden geschat op wegen waar niet alle rijstroken bemeten zijn?

Op wegen met twee rijstroken is de intensiteit gelijk aan de som van de intensiteit op de linker rijstrook en de intensiteit op de rechter rijstrook. Veel van de wegen met twee rijstroken hebben alleen een camera op de rechter rijstrook. Op basis van de intensiteit op de rechter rijstrook hebben we geprobeerd met lineaire regressie een voorspelling te doen van de intensiteit op de linker rijstrook. Dit lineaire regressiemodel konden we opstellen omdat er een aantal wegen waren met twee rijstroken waarbij een camera stond op zowel de linker als de rechter rijstrook. Op deze wegen hebben we gekeken wat de intensiteit op de linker rijstrook is als percentage van de intensiteit op de rechter rijstrook. Zo ontstaat een verhouding die we p noemen. Er bleek een positief verband te zijn tussen p en twee verklarende variabelen: de intensiteit op de rechter rijstrook en de afstand van het meetpunt van de camera tot het vorig kruispunt. Het lineaire regressiemodel gaf een verklarende variantie van 80% voor de meetpunten waarop het model was gebaseerd. De twee verklarende variabelen lijken p goed te voorspellen.

Uit de resultaten bleek dat er redelijk grote verschillen waren tussen de oude tellingen en de intensiteit voor het verkeersmodel op de locaties waar de linker rijstrook is voorspeld. Op een aantal punten was het verschil zelfs rond de 80%. Op basis hiervan concluderen wij dat het gebruikte regressiemodel geen voorspelling geeft voor p met een hoge betrouwbaarheid.

Wij vermoeden dat de meetpunten waar alle rijstroken bemeten waren, niet representatief genoeg waren voor alle meetpunten. Waarschijnlijk hangt p sterk af van hoe de meetlocatie er precies uit ziet. De meetpunten met volledige

informatie op alle rijstroken was echter beperkt tot 10 meetpunten. Op basis van de beschikbare data kan geen betrouwbare uitspraak worden gedaan over meetlocaties waar niet alle rijstroken zijn bemeten.

7.3 Aanbevelingen

Neem de meetlocaties waar alle rijstroken bemeten zijn in gebruik
Uit de resultaten blijkt dat door een gemiddelde te nemen van de hoogste n gemeten intensiteiten er betrouwbare intensiteiten voor het verkeersmodel ontstaan. Dit is echter alleen het geval bij de meetlocaties waar alle rijstroken zijn bemeten.

Om de betrouwbaarheid van intensiteiten voor het verkeersmodel te verhogen, overleg met Vialis of betrouwbaarheidsindicatoren per uur en camera kunnen worden geleverd. In dit onderzoek was geen beschikking over de betrouwbaarheidsdata per uur. Wanneer er beschikking is over de herkenningspercentages en beschikbaarheidspercentages per camera en per uur, dan kan per gemeten intensiteit met een grote betrouwbaarheid bepaald worden hoeveel voertuigen werkelijk langs de camera reden.

Onderzoek de kosten / baten van het plaatsen van een camera op rijstroken waar niet alle rijstroken zijn bemeten. Omdat het regressiemodel geen betrouwbare waarden voor p opleverde voor meetlocaties, waar niet alle rijstroken waren bemeten, adviseren we om de data op deze meetlocaties niet te gebruiken. Wanneer een camera geplaatst wordt op de rijstroken waar nog geen camera staat, dan zijn alle rijstroken bemeten. Van meetlocaties waarvan alle rijstroken bemeten zijn weten we dat op basis van de data uit camera's op deze locaties betrouwbare intensiteiten voor het verkeersmodel kunnen worden bepaald. Het is interessant om te onderzoeken hoeveel dit precies kost, en wat de toegevoegde waarde is van de intensiteiten op deze meetlocaties.

7.4 Algemene aanbevelingen

Sla de kentekenregistraties op in een database De kentekenregistraties bestaan uit veel records. Deze worden op dit moment wekelijks geleverd in csv-bestanden. Om deze data op een makkelijke manier te kunnen gebruiken is een database nodig. Er zou een script worden kunnen geschreven waarbij de data die Vialis levert automatisch wordt geïmporteerd in de database. Een optie is ook om Vialis te verzoeken om de kentekenregistraties in databaseformaat te leveren, waardoor de data makkelijker kan worden geïmporteerd.

Overleg met Vialis of voertuigtype aan data kan worden toegevoegd
Zoals in hoofdstuk 3.2.1 staat beschreven, slaat Vialis het voertuigtype van een kenteken op. Een verzoek kan worden gedaan aan Vialis om deze informatie bij de kentekenregistraties te voegen. Wanneer deze informatie beschikbaar

is kunnen intensiteiten worden ingedeeld naar voertuigtype. Overwogen moet worden of deze informatie voor Verkeersonderzoek toegevoegde waarde heeft.

Referenties

- [1] <http://www.amsterdam.nl/gemeente/organisatie-diensten/ivv/divv-organisatie/missie-visie/>, datum 30-03-2012
- [2] Achterberg, W., van den Elshout, J., Moeskops, L., Teekamp, R. (2010): Rapportage Actualisatie Basisprognoses, blz. 8
- [3] Harms, F., Schrijver, J., Grooten, J., van der Lof, M. (2010): Eindverslag vernieuwingstraject afdeling Verkeersonderzoek, blz. 10
- [4] van Sleuwen, K. (2011): Onderzoek MoRo systeem gemeente Rotterdam, blz. 6
- [5] Wu, N.: Impact of traffic regulation on lane flow - distribution and capacity of motorways, blz. 3
- [6] Dienst Infrastructuur Verkeer en Vervoer (2005): Beleidskader Hoofdnetten, blz. 22
- [7] Hek, L. (2010): Categorisering hoofdnet Auto Amsterdam, blz. 10

A Alternatieve filtermethoden

In dit onderzoek zijn een aantal oplossingsmethoden overwogen. Echter, wanneer de filtermethode zoals beschreven in paragraaf 4.4 wordt toegepast om de intensiteit voor het verkeersmodel te bepalen zijn deze alternatieve oplossingsmethoden overbodig. Gemeten intensiteiten die kunnen worden gefilterd met deze alternatieve methoden, zullen waarschijnlijk al worden gefilterd door de filtermethode. De volgende methoden zijn daarom niet volledig uitgewerkt. Ze dienen slechts als hint voor eventuele verdere toepassingen van de data. De twee methodes zijn filteren op basis van volgtijden en filteren op basis van relaties tussen camera's.

A.1 Volgtijden

Kentekenregistraties zijn op secondeniveau beschikbaar. Door de tijdstippen van kentekenregistraties te bekijken, kan de betrouwbaarheid van een camera worden onderzocht. Op een bepaalde locatie voor een bepaalde periode zijn m voertuigen gemeten. Voor elk van deze m voertuigen is precies het tijdstip bekend wanneer ze de camera zijn gepasseerd. Er kan een analyse worden gedaan door de verschillen in deze tijdstippen te onderzoeken.

Definieer de m tijdstippen van registratie als t_1, \dots, t_m . Nu definiëren we de volgtijden als het verschil tussen het tijdstip van registratie van een voertuig en het tijdstip van registratie van het voertuig dat voor hem de camera passeerde. De volgtijden in deze periode noteren we als v_2, \dots, v_m . Bekijk nu het maximum van alle volgtijden, $\max(v_i, \dots, v_m)$.

Als we de intensiteit meten in een periode van een uur dan vermoeden we geen onbetrouwbaarheid als we bijvoorbeeld zien dat $\max(v_i, \dots, v_m) = 3$ min. Maar wanneer bijvoorbeeld $\max(v_i, \dots, v_m) = 30$ min betekent dit dat gedurende een half uur geen voertuigen zijn geregistreerd. Dit kan er op duiden dat de camera niet goed werkte.

A.2 Relaties tussen camera's

Voertuigen kunnen meerdere camera's passeren. Met behulp van het aantal auto's dat twee opeenvolgende camera's passeert kan een uitspraak gedaan worden over de gemeten intensiteiten van die twee camera's. Beschouw camera A en camera B die dichtbij elkaar in de buurt liggen. Voertuigen kunnen alleen van camera A naar camera B rijden. Bekijk in een bepaalde periode de kentekens van beide camera's die overeen komen. Dit zijn het aantal voertuigen die zowel camera A als camera B zijn gepasseerd in die periode en definiëren we als n . Wanneer n constant blijkt te zijn voor een bepaalde periode, dan kunnen uitspraken worden gedaan over gemeten intensiteiten van camera A en camera B. Beschouw een voorbeeld waar een gemeten intensiteit slechts twee waarden aan kan nemen: hoog of laag. Nu kan worden gesteld:

Als n hoog is, dan zijn de gemeten intensiteit van camera A en
gemeten intensiteit van camera B ook hoog;

Als n laag is, dan zijn de gemeten intensiteit van camera A laag, gemeten intensiteit van camera B laag, of gemeten intensiteit van beide camera's laag.

Zo kan op basis van n een uitspraak worden gedaan over de gemeten intensiteiten van camera A en camera B.

B Uitdagingen tijdens het onderzoek

Net als elk ander onderzoek kende ook dit onderzoek een aantal uitdagingen. De belangrijkste uitdagingen zullen in deze Appendix worden besproken.

B.1 betrouwbaarheid van coördinaten van meetlocaties

Voor het onderzoek was het van belang om de exacte plaatsing in de stad van de camera's te achterhalen. De locaties van de camera's staan vast door middel van coördinaten. Zie Appendix C voor de locaties van de camera's. De eerste lijst met coördinaten werd geleverd door een contactpersoon van IBA (Ingenieursbedrijf Amsterdam). Deze contactpersoon had al eerder gewerkt met de MoCo-camera's. Tevens heeft deze contactpersoon ook een lijst geleverd met de straatnaam behorende bij de locaties.

Al snel werd duidelijk dat niet alle geleverde coördinaten precies waren. Door met behulp van Google Maps te kijken waar een bepaalde coördinaat naar verwijst, konden we opmerken dat bepaalde coördinaten niet duidelijk naar een bepaalde weg verwezen. In deze gevallen waren een aantal wegen aanwezig waar de coördinaat naar zou kunnen verwijzen.

Vervolgens hebben we aan Vialis gevraagd om een lijst met coördinaten te leveren. Vialis heeft aan dit verzoek voldaan. De coördinaten in deze lijst bleken nauwkeuriger te zijn dan de coördinaten in de oude lijst. Deze coördinaten verwezen consistent naar de plek waar de camera hangt. Toch zagen we dat ook deze lijst niet helemaal nauwkeurig is. Zo was er bijvoorbeeld voor twee meetlocaties hetzelfde coördinaat gevonden.

Later in het onderzoek kwamen we achter het bestaan van zogenaamde locatieonderzoeken. Dit zijn rapporten waarin voor elk van de MoCo-camera's een foto staat van de exacte plaatsing. Door deze rapporten naast de coördinatenlijst van Vialis te leggen kon voor elke locatie de juiste coördinaat worden bepaald.

B.2 bepaling van het aantal rijstroken

Van elke meetlocatie wilden we het aantal rijstroken op de meetlocatie weten. Een lijst met het aantal rijstroken per meetlocatie was niet aanwezig. Daarom hebben we besloten om het aantal rijstroken voor iedere meetlocatie handmatig op te zoeken. Dit hebben we gedaan door in Google Maps de locaties op te zoeken, en te kijken hoeveel rijstroken aanwezig zijn. Voor de meeste locaties was dit duidelijk te zien. Voor andere locaties was dit minder duidelijk. Het exacte meetpunt van de camera is belangrijk voor de bepaling van het aantal rijstroken op deze weg. Met meetpunt bedoelen we het punt op de weg waar de kentekens in beeld komen. Soms gaat een weg van twee rijstroken bijvoorbeeld over in drie rijstroken.

Een lange tijd was er onzekerheid over het aantal rijstroken op een aantal locaties. Maar ook hier brachten de locatieonderzoeken uitkomst. In de locatieonderzoeken is het exacte meetpunt van de camera vastgelegd. Met behulp

van de locatieonderzoeken konden we voor elke meetlocatie het aantal rijstroken bepalen

B.3 het verwerken van kentekenregistraties tot gemeten intensiteiten

De MoCo-camera's genereren veel data. De data van alle camera's van enkele weken bestaat uit miljoenen records. Per week staat de data van alle camera's verdeeld over bijna 200 Excel bestanden. De uitdaging was om de data uit al deze Excel bestanden te verwerken tot gemeten intensiteiten. Gezien de hoeveelheid van de data is het praktisch om te werken met een database. Het leek echter te lastig om een server te draaien op onze werkplek bij DIVV. Daarom hebben we ervoor gekozen om een script te maken in Visual Basic for Excel. Het maken van een werkend script heeft wel enige tijd in beslag genomen.

B.4 interpretatie van betrouwbaarheid van gemeten intensiteiten

Nadat het script in Excel werkte kon voor elke camera een aan gemeten intensiteiten worden bepaald. Zonder een verkeerskundige achtergrond is het in eerste instantie lastig te beoordelen of deze gemeten intensiteiten realistisch lijken. Maar door veel prognoses in GenMod te bestuderen kon uiteindelijk een globaal idee verkregen worden van wat realistische intensiteiten zijn en wat niet.

Een grote vooruitgang in het onderzoek werd geboekt toen de contactpersoon van IBA ons wees op het bestaan van kwaliteitsrapporten van de MoCo-camera's. Door het bestuderen van deze rapporten konden we zien dat de MoCo-camera's over het algemeen goed werken. Nadat we de betrouwbaarheidsdata waarop deze rapporten gebaseerd zijn hadden opgevraagd, konden we aan de analyses beginnen.

Een andere grote vooruitgang in het onderzoek was het moment dat we een interview hadden met een contactpersoon van Vialis. Deze leverde ons het betrouwbaarheidsonderzoek van het reistijdmeetsysteem in Rotterdam. Uit dit onderzoek konden inzichten worden verkregen over wat voor soort meetfouten mogelijk zijn. Het betrouwbaarheidsonderzoek van Rotterdam alleen was voor ons niet duidelijk genoeg. Het contact met Vialis bleek erg waardevol om onduidelijkheden in dit onderzoek te verhelderen.

C Meetlocaties

In deze Appendix zijn alle meetlocaties weergegeven. Bij elke meetlocatie staan de straatnaam en de coördinaten van de meetlocatie. Ook is de informatie over de rijstroken weergegeven:

- het aantal rijstroken dat bemeten is;
- het aantal rijstroken dat aanwezig is;
- het percentage van de aanwezige rijstroken dat bemeten is.

Meetlocatie	Straat	Y-coördinaat	X-coördinaat	# bemeten	# aanwezig	% bemeten
3	Haarlemmerweg	52,384825	4,847137	4	4	100%
27	Nieuwe Utrechtseweg	52,337823	4,910781	3	3	100%
1	Spaarndammerdijk	52,392984	4,871022	2	2	100%
11	Cornelis Lelylaan	52,357837	4,845479	2	2	100%
19	Amstelveenseweg	52,339553	4,85774	2	2	100%
21	Beethovenstraat	52,34012	4,876975	2	2	100%
26	Kennedylaan	52,340662	4,89384	2	2	100%
29	Gooiseweg	52,339827	4,937756	2	2	100%
32	Pietheintunnel	52,372332	4,961724	2	2	100%
3402	Nieuwe Leeuwarderweg s116	52,38606944	4,913469444	2	2	100%
3409	Pampusweg	52,34043056	5,014252778	2	2	100%
3790	Henk Sneevlietweg s107	52,34588889	4,839866667	2	2	100%
38	Nieuwe Leeuwarderweg s116	52,38675	4,914616	2	2	100%
3805	Zuiderzeeweg s115	52,38971111	4,965972222	2	2	100%
6	Jan van Galenstraat	52,372088	4,844396	2	2	100%
13	Heemstedestraat	52,351466	4,843662	1	1	100%
15	Voorburgstraat	52,348488	4,84224	1	1	100%
16	Vlaardingenlaan	52,346806	4,841887	1	1	100%
17	Henk Sneevlietweg	52,346632	4,846497	1	1	100%
18	Naaldwijkstraat	52,346257	4,842278	1	1	100%
20	Parnassusweg	52,339701	4,868672	1	1	100%
2	Sloterdijkweg	52,386831	4,845423	1	1	100%
22	Wielingenstraat	52,344595	4,886168	1	1	100%
23	Scheldestraat	52,345012	4,890988	1	1	100%
25	Europaplein	52,341839	4,893579	1	1	100%
28	Spaklerweg	52,342944	4,918738	1	1	100%
31	Middenweg	52,341654	4,950234	1	1	100%

33	Flevoweg	52,364924	4,947703	1	1	100%
3400	Valkenburgerstraat Z-N	52,36961667	4,907275	1	1	100%
34	Zeeburgerdijk	52,366202	4,944876	1	1	100%
3401	Kattenburgstraat N-Z	52,37601667	4,921638889	1	1	100%
3404	Piet Heinkade – s100	52,37616389	4,922019444	1	1	100%
3406	Ringweg noord A10, oprit buitenring	52,40763611	4,939158333	1	1	100%
3407	Piet Heinkade – s100	52,37638889	4,921425	1	1	100%
3410	Valkenburgerstraat N-Z	52,36919444	4,905991667	1	1	100%
3447	Nieuwe Hemweg – s101	52,39452778	4,863802778	1	1	100%
35	Zeeburgerpad	52,367374	4,939621	1	1	100%
36	Molukkenstraat	52,36639	4,939427	1	1	100%
3701	Van Diemenstraat – s100	52,39056111	4,885086111	1	1	100%
3702	Houtmankade – s100	52,39012778	4,884863889	1	1	100%
3703	Tasmanstraat – s101	52,39083333	4,884838889	1	1	100%
3704	Nieuwe Hemweg – s101	52,40050833	4,850247222	1	1	100%
3705	Nieuwe Hemweg – s101	52,40158333	4,848641667	1	1	100%
3712	Nassaukade – s100	52,38486667	4,881475	1	1	100%
3713	Haarlemmerweg – s103	52,38519722	4,881044444	1	1	100%
3714	Houtmankade – s100	52,38515556	4,881622222	1	1	100%
3715	Burg. De vluchtlaan s104	52,38125278	4,821711111	1	1	100%
3719	Nassaukade – s100	52,37396111	4,874994444	1	1	100%
3720	2e Hugo de Grootstraat – s106	52,37424167	4,874802778	1	1	100%
3721	Nassaukade – s100	52,37425833	4,875272222	1	1	100%
3727	Overtoom – s106	52,36373333	4,878716667	1	1	100%
3728	Nassaukade – s100	52,36400556	4,878841667	1	1	100%
3734	Hobbemakade – s108	52,36003056	4,886641667	1	1	100%
3739	Burg. Roelstraat s105	52,37615833	4,819397222	1	1	100%
3745	Amsteldijk s110	52,35786667	4,904155556	1	1	100%
3746	Stadhouderskade – s100	52,35872222	4,903405556	1	1	100%
3748	Jan van Galenstraat	52,365864	4,804143	1	1	100%
3752	Julianaplein	52,34751667	4,918472222	1	1	100%
3760	Langbroekdreef / Gooiseweg	52,30738333	4,971811111	1	1	100%
3763	Mauritskade – s100	52,36390278	4,923447222	1	1	100%
3764	Lineausstraat – s113	52,36282222	4,923527778	1	1	100%
3765	Mauritskade – s100 / s113	52,36328056	4,922041667	1	1	100%
3766	Hartveldseweg – s113	52,33997222	4,952783333	1	1	100%
3768	Meer en Vaart s106	52,35670833	4,809338889	1	1	100%
3769	Oprit Ringweg-Oost – A10	52,34048056	4,953219444	1	1	100%
3772	Piet Heinkade – s100	52,373175	4,933230556	1	1	100%
3773	IJburglaan – s114	52,36596389	4,975175	1	1	100%

3777	Oprit Ringweg-Noord – A10	52,38605278	4,913436111	1	1	100%
3783	Jan van Galenstraat s105	52,374634	4,858969	1	1	100%
3784	Jan van Galenstraat s105	52,37469444	4,858188889	1	1	100%
3785	Overtoom s106	52,35772222	4,8558	1	1	100%
3786	Amstelveenseweg	52,35756944	4,854988889	1	1	100%
3788	Surinameplein s106	52,35879444	4,853461111	1	1	100%
3791	Henk Sneevlietweg s107	52,34552778	4,827938889	1	1	100%
3793	Oude Haagseweg s107	52,33589167	4,815291667	1	1	100%
3795	Stadionweg s108	52,3488	4,883430556	1	1	100%
3796	Diepenbrockstraat s109	52,34701389	4,880083333	1	1	100%
3797	Stadionweg s108	52,34705833	4,879516667	1	1	100%
3798	Middenweg s113	52,35071389	4,935222222	1	1	100%
3799	Middenweg s113	52,35101944	4,935072222	1	1	100%
3800	Zuiderzeeweg	52,36526111	4,949008333	1	1	100%
3801	Zuiderzeeweg	52,37173889	4,962377778	1	1	100%
3802	Zuiderzeeweg	52,37230278	4,962436111	1	1	100%
3803	IJburglaan s114	52,35272778	4,999675	1	1	100%
3804	IJburglaan s114	52,35303889	4,999661111	1	1	100%
3806	Oprit A10 binnenring	52,38947222	4,965575	1	1	100%
3809	Klaprozenweg s118	52,40279722	4,897286111	1	1	100%
3810	Klaprozenweg s118	52,40303333	4,897113889	1	1	100%
4	Wiltzanghlaan	52,381679	4,845596	1	1	100%
7	Jan Evertsenstraat	52,369313	4,844578	1	1	100%
8	Piet Mondriaanstraat	52,365846	4,842976	1	1	100%
9	Postjesweg	52,364129	4,843789	1	1	100%
14	Rijswijkstraat	52,350048	4,844042	1	1	100%
24	Geulstraat	52,344729	4,892097	1	1	100%
37	Zeeburgpad	52,366905	4,934119	1	1	100%
3408	Nieuwe Leeuwarderweg s116	52,40581944	4,93715	2	3	67%
3743	Europaboulevard – s109	52,33793	4,892091	2	3	67%
5	Bos en Lommerplein	52,377743	4,845061	2	3	67%
3405	Slochterweg N246 / s116	52,40868889	4,940733333	2	4	50%
12	Nachtwachtlaan	52,356859	4,843706	1	2	50%
3399	Weesperstraat	52,36547222	4,905561111	1	2	50%
3403	Ringweg noord A10, oprit binnenring	52,40654167	4,938663889	1	2	50%
3443	Basisweg - s102	52,3913	4,819477778	1	2	50%
3445	Spaarndammerdijk	52,39114444	4,818416667	1	2	50%
3446	Transformatorweg– s102	52,39266944	4,862752778	1	2	50%
3448	Cornelis Lelylaan – s106	52,35615556	4,809813889	1	2	50%

3708	Oprit Einsteinweg (A10) linksom	52,39136944	4,843577778	1	2	50%
3709	Basisweg – s102	52,39197222	4,843077778	1	2	50%
3710	Oprit Einsteinweg (A10), rechtsom	52,39222222	4,844980556	1	2	50%
3711	Haarlemmerplein – s103	52,38499444	4,881680556	1	2	50%
3716	Oprit Einsteinweg (A10) linksom	52,38443611	4,844025	1	2	50%
3717	Haarlemmerweg N200	52,38486667	4,842775	1	2	50%
3718	Haarlemmerweg – s103 – N200	52,38431389	4,819227778	1	2	50%
3722	Burg. Roelstraat s105	52,37546944	4,819786111	1	2	50%
3723	Oprit Einsteinweg A10 ri. zuid	52,371975	4,841708333	1	2	50%
3724	Jan van Galenstraat – s105	52,37244722	4,841372222	1	2	50%
3725	Oprit Einsteinweg A10 ri. noord	52,372537	4,842765	1	2	50%
3726	Stadhouderskade – s100	52,36373333	4,878716667	1	2	50%
3729	Slotermeerlaan s207	52,37524722	4,818683333	1	2	50%
3730	Oprit Einsteinweg A10 ri. zuid	52,35751111	4,841930556	1	2	50%
3732	Oprit Einsteinweg A10 ri. noord	52,35824722	4,843027778	1	2	50%
3733	Stadhouderskade – s100	52,35996667	4,88695	1	2	50%
3735	Stadhouderskade – s100	52,36015833	4,886769444	1	2	50%
3736	Oprit Ringweg zuid A10 ri. Oost	52,33697778	4,858680556	1	2	50%
3738	Oprit Ringweg zuid A10 ri. west	52,33901944	4,858125	1	2	50%
3740	Oprit Ringweg-Zuid – A10	52,33380833	4,890613889	1	2	50%
3741	Europaboulevard – s109	52,334099	4,89001	1	2	50%
3742	Oprit Ringweg-Zuid – A10	52,337609	4,891382	1	2	50%
3744	Mauritskade s100	52,36022222	4,910638889	1	2	50%
3747	Nieuwe Utrechtseweg - A2 – s110	52,33475833	4,912036111	1	2	50%
3749	Rhijnspoorplein – s112	52,35988611	4,908633333	1	2	50%
3750	Rhijnspoorplein – s112	52,36043333	4,908611111	1	2	50%
3753	Mr. Treublaan	52,34817222	4,917583333	1	2	50%
3754	Wiboutstraat – s112	52,34871944	4,918458333	1	2	50%
3755	Oprit ringweg Oost – A10	52,33614722	4,943113889	1	2	50%
3757	Oprit Ringweg Zuid	52,33788056	4,939233333	1	2	50%
3758	Gooiseweg	52,36626389	4,804652778	1	2	50%
3759	Gaasperdammerweg	52,30765278	4,971930556	1	2	50%

3761	Oprit Gaasperdammerweg	52,30831389	4,9709	1	2	50%
3762	Gooiseweg – s112	52,30845278	4,971169444	1	2	50%
3767	Oprit Ringweg-Oost – A10	52,34123333	4,951072222	1	2	50%
3770	Piet Heintunnel (inrit) – s114	52,37250556	4,934386111	1	2	50%
3771	Panamalaan – s100	52,37231944	4,933633333	1	2	50%
3774	Oprit Ringweg-Oost	52,36799722	4,970488889	1	2	50%
3775	IJburglaan – s114	52,3684	4,970894444	1	2	50%
3776	Oprit Ringweg-Oost	52,36608611	4,974094444	1	2	50%
3778	Verlengde Stellingweg – s118	52,42333889	4,882758333	1	2	50%
3779	Oprit Ringweg-Noord – A10	52,42508333	4,880641667	1	2	50%
3780	Verlengde Stellingweg – s118	52,4256	4,882552778	1	2	50%
3781	Bos en Lommerplein s104	52,37798056	4,844433333	1	2	50%
3782	Bos en Lommerplein s104,	52,37773333	4,843627778	1	2	50%
3787	Surinameplein s106	52,358225	4,852297222	1	2	50%
3789	Einsteinweg oprit A10 binnenring	52,34611944	4,841652778	1	2	50%
3792	Johan Huizingalaan s107	52,34528611	4,827397222	1	2	50%
3794	Nieuwe Haagseweg oprit A10 binnenring	52,33863889	4,812611111	1	2	50%
3731	Cornelis Lelylaan – s106	52,3581	4,841875	1	3	33%
3737	Amstelveenseweg – s108	52,33684722	4,857405556	1	3	33%
3751	Gooiseweg – s112	52,34828889	4,9197	1	3	33%
3756	Gooiseweg	52,3361	4,943525	1	3	33%
3808	Midenweg N247	52,421658	4,952665	1	3	33%
3707	Transformatorweg – s102	52,39156667	4,845377778	1	4	25%
3807	Kanaaldijk N235	52,422005	4,953132	1	4	25%