

VU Research Portal

Betrouwbaarheid reistijden

Koomen, Eric; Jacobs-Crisioni, C.G.W.; König, Yosef

2022

document version

Other version

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Koomen, E., Jacobs-Crisioni, C. G. W., & König, Y. (2022). *Betrouwbaarheid reistijden: analyse op basis van floating car data*. Vrije Universiteit / Afd. Ruimtelijke Economie.

General rights

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

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

Take down policy

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

E-mail address:

vuresearchportal.ub@vu.nl

Betrouwbaarheid reistijden; analyse op basis van floating car data

Eric Koomen, Chris Jacobs-Crisioni, Yosef König
10 juni, 2022



Opdracht voor Rijkswaterstaat, Water Verkeer en Leefomgeving

Contactinformatie:
Vrije Universiteit Amsterdam
School of Business and Economics
Department of Spatial Economics/ Spatial Information Laboratory (SPIN/*lab*)
De Boelelaan 1105
1081 HV Amsterdam
Netherlands
Phone: +31 20 5986095
Email: e.koomen@vu.nl
Website: <https://spinlab.vu.nl/>

Samenvatting

De BetrouwbaarheidsTool van het Landelijk Model System (LMSBT2) wordt gebruikt om de (on)betrouwbaarheid van autoreistijden van deur-tot-deur te ramen in maatschappelijke kosten-baten analyses. Het wordt onder meer gebruikt om de mogelijke ontwikkeling van reistijd-onbetrouwbaarheid op de lange termijn (2040-2050) in te kunnen schatten. De voorliggende studie draagt bij aan het verder ontwikkelen van het bestaande instrument en beoogt inzicht te geven in de factoren die van invloed zijn op de betrouwbaarheid van reistijden, waaronder in het bijzonder de interactie tussen hoofdwegennet (HWN) en onderliggend wegennet (OWN). Het doel is om uitspraken te kunnen doen over de betrouwbaarheid van reistijden van deur tot deur.

Voortbouwend op eerdere pilotstudies die de Vrije Universiteit uitvoerde in opdracht van Rijkswaterstaat, gebruiken we floating car data voor het beschrijven van 'deur-tot-deur' reistijden voor een set van routes, om vervolgens de betrouwbaarheid van deze reistijden te bepalen. Voor deze studie gebruiken we de periode tussen de zomer- en kerstvakantie van 2019. Deze periode van 15 weken beschrijft een doorsnee deel van het jaar zonder uitgebreide vakanties en extreme weersomstandigheden. Recentere gegevens zijn minder relevant omdat deze sterk door de Corona-pandemie beïnvloed zijn.

Met behulp van de floating car data en de trajectoriënmethode reconstrueren wij reistijden voor 120 routes bij een verondersteld vertrek per minuut tussen 6:00 uur 's ochtends en 7:00 uur 's avonds voor alle afzonderlijke werkdagen in de studieperiode. Voor de minuten tussen 7 uur 's avonds en 6 uur 's ochtends gaan we ervanuit dat de onbetrouwbaarheid verwaarloosbaar is. Dit levert in theorie bijna 7 miljoen afzonderlijke reistijden op die we gebruiken om verwachte reistijden per route te bepalen. Een deel van de waarnemingen valt echter weg omdat enkele dagen onbruikbaar zijn en we ervoor kiezen extreme vertragingen buiten beschouwing te laten. In navolging van eerder onderzoek gaan we uit van goed geïnformeerde reizigers die per werkdag en vertrekkwartier een bepaalde reistijd verwachten. De gemiddelde tijd die de reis op die werkdag en in dat vertrekkwartier langer duurt dan de snelst mogelijke (freeflow) reistijd noemen we reguliere vertraging. De standaardafwijking van de verwachte reistijd noemen we onbetrouwbaarheid.

De reistijd-onbetrouwbaarheid van een route is vervolgens verklaard aan de hand van verschillende variabelen zoals de reguliere vertraging op die route, het deel van die vertraging dat op het onderliggend wegennet wordt ondervonden, de over de gehele route gemiddelde snelheid bij de snelst mogelijk reis op de route (totale afstand gedeeld door de totale reistijd), de gemiddelde hoeveelheid verkeer op een route en een dagdeel-specifiek gemiddeld effect. Deze resultaten kunnen gebruikt worden om LMSBT2 te actualiseren. Het rapport gaat uitgebreid in op de overeenkomsten en verschillen met een eerdere studie uit 2014 die nu gebruikt wordt voor onderbouwing van de tool. Daarnaast worden diverse alternatieve specificaties besproken die onder meer zicht geven op het effect van weersomstandigheden, wegwerkzaamheden en variatie tijdens de dag, tussen werkdagen en over de studieperiode. Daarbij gaan we dieper in op de econometrische achtergrond van de gekozen specificatie.

Inhoudsopgave

Samenvatting	3
1. Inleiding.....	6
2. Literatuuronderzoek	8
3. Vaststellen routeset	10
4. Samenstellen reistijd database voor de vastgestelde routeset.....	16
5. Vergelijk met voorgaande studie	20
5.1. Reistijden, vertraging en onbetrouwbaarheid.....	20
5.2. Herhalen verklarende analyse huidige rekenmodule LMS-BT.....	27
5.3. Mogelijke oorzaken voor verschil in uitkomsten ten opzichte van de eerdere analyse ..	29
5.4. Totale onbetrouwbaarheid per route.....	37
6. Verdiepende analyse.....	40
6.1. Geïntegreerde benadering onbetrouwbaarheid	40
6.2. Effect rijrichting.....	41
6.3. Aanvullende verklarende variabelen	42
6.4. Verklaar natuurlijk logaritme onbetrouwbaarheid	47
6.5. Alternatieve specificatie om aan te sluiten bij de datastructuur	49
7. Voorstel aanpassing LMS-BT	52
7.1. Dagdeel-specifieke dummy's	52
7.2. Splitsing van HWN en OWN bijdragen voor rapportages	54
Referenties.....	56
Bijlage 1 Koppeling LMS en floating car data netwerken	57
Bijlage 2 Verschillende selectiemethoden voor bepalen gemiddelde reistijden	60
Bijlage 3 Complete schattingsresultaten voor de modellen met tijdsdummies.....	61

1. Inleiding

Rijkswaterstaat is een project gestart dat zich richt op het verder ontwikkelen van het bestaande instrument Landelijk Model System Betrouwbaarheids Tool (LMSBT2, beschreven in Kouwenhoven en Bel, 2014). Dit instrument wordt gebruikt om de (on)betrouwbaarheid van autoreistijden van deur-tot-deur te ramen in maatschappelijke kosten-baten analyses. Het instrument is een postprocessing tool op het Landelijk Model Systeem. Hiermee is het mogelijk om de ontwikkeling van reistijdontbetrouwbaarheden op de lange termijn (2040-2050) te kunnen inschatten. Naast dit hoofddoel heeft het project ook tot doel om meer (wetenschappelijk) inzicht te krijgen in de invloedsfactoren van betrouwbaarheid van reistijden, waaronder eventuele beleidsmaatregelen, en in het bijzonder de interactie tussen hoofdwegennet (HWN) en onderliggend wegennet (OWN). Het doel is om uitspraken te kunnen doen over de betrouwbaarheid van reistijden van deur tot deur.

Het Rijkswaterstaat project kent twee fases: (1) data-analyse en modelschatting voor de onbetrouwbaarheid van reistijden op routes die uit zowel HWN als OWN deeltrajecten bestaan, en die voldoende representatief zijn om de betrouwbaarheid van reistijden op een landelijk niveau te analyseren; (2) updaten van het bestaande instrument en toepassen van het vernieuwde instrument voor een case study. Deze rapportage richt zich alleen op de werkzaamheden van de eerste fase.

Het voorliggende onderzoek bouwt voort op eerdere pilotstudies die zijn uitgevoerd door de Vrije Universiteit in opdracht van Rijkswaterstaat en waarin de bruikbaarheid van floating car data werd onderzocht (Koomen en Jacobs-Crisioni, 2021; Koomen et al., 2018). Deze eerdere studies toonden aan dat de relatie tussen vertraging en (on)betrouwbaarheid op het OWN anders is dan op het HWN. Met deze nieuwe data kan het effect van vertraging op onbetrouwbaarheid van van deur-tot-deur bepaalde reistijden, in samenhang met eventuele aanvullende verklarende variabelen, in beeld worden gebracht. Met dit opnieuw geschatte verband kan vervolgens het instrument LMS-BT in de tweede fase verder worden ontwikkeld voor toepassing op deur-tot-deur reistijden.

In deze eerste fase project richt de VU zich op het beschrijven van 'deur-tot-deur' reistijden op basis van floating car data om de betrouwbaarheid van reistijden op het wegennet voor een set van routes te bepalen. Hiervoor wordt een van BeMobile afkomstige en door NDW beschikbaar gestelde dataset voor het jaar 2019 gebruikt, die deels ook in de pilot-studie door de VU verwerkt is (Koomen en Jacobs-Crisioni, 2021). In die studie gebruikten we data over een periode van vier weken in september. Voor deze studie gebruiken we de periode tussen de zomer- en kerstvakantie van dat jaar: 2 september tot en met 13 december 2019. Deze periode van 15 weken beschrijft een doorsnee deel van het jaar zonder uitgebreide vakanties en extreme weersomstandigheden. De herfstvakantie (19 oktober t/m 27 oktober 2019 in regio's Noord en Midden, 12 oktober t/m 20 oktober 2019 in regio Zuid) is gemarkeerd in de analyse zodat een eventueel effect op betrouwbaarheid onderzocht kan worden. In navolging van het onderzoek van Significance, waar wij op voortbouwen, wordt voor deze vakantieperiode geen aangepaste reistijdverwachting opgesteld¹. Data van voor de zomer was deels incompleet zoals we in onze vorige rapportage beschreven. Vanaf 2020 leiden de beperkingen als gevolg van de Corona-pandemie tot sterk afwijkende reispatronen.

Zoals beschreven in het plan van aanpak (Rijkswaterstaat, 2021), bestaan de activiteiten in de eerste fase van het project uit een vijftal stappen:

- 1) Literatuuronderzoek;
- 2) Vaststellen routeset;
- 3) Samenstellen reistijd database voor de vastgestelde routeset;

¹ Voor de zomer- en kerstvakantie is dat destijds wel gedaan omdat de reistijden duidelijk afweken van het normale patroon (zie pagina 8 in Kouwenhoven en Bel, 2014). Ook in onze dataset wijken de reistijden in de herfstvakantieperiode niet significant af van de omliggende weken, zie daarvoor Figuur 5.

- 4) Verklarende analyse variantie in reistijden;
- 5) Documentatie en overdracht resultaten,

De navolgende hoofdstukken doen verslag van de eerste vier bovengenoemde stappen. Voor de leesbaarheid zijn de resultaten van stap 4 in drie opeenvolgende hoofdstukken beschreven. Hoofdstuk 5 beschrijft de initiële analyse waarin we ons vooral richten op het reproduceren van de analyses die in 2014 voor LMS-BT zijn gedaan. Daarin besteden we veel aandacht aan de overeenkomsten met de eerdere studie en mogelijke oorzaken voor de verschillen die we vinden. In hoofdstuk 6 breiden we de verklarende analyse uit met een aantal nieuwe factoren die niet in de eerdere analyses gebruikt zijn. In hoofdstuk 7 doen we op basis van de verkennende analyses vervolgens een voorstel hoe LMS-BT aangepast kan worden.

2. Literatuuronderzoek

De eerste stap in deze studie betrof het maken van een overzicht van de factoren die mogelijk van invloed zijn op reistijdbetrouwbaarheid. Hiervoor is een beknopte literatuurstudie gedaan waarvoor studies verzameld zijn die afkomstig zijn uit diverse disciplines en uiteenlopende onderzoeksopzetten kennen. Er is vooral gezocht naar studies die variatie in autoreistijden verklaren op basis van verschillende factoren die ook voor onze studie van belang kunnen zijn. Om zo veel mogelijk studies te kunnen betrekken is niet alleen gekeken naar studies die zich richten op het gebruik van floating car data, maar ook naar studies die gebruik maakten van andere bronnen voor reistijden.

De uiteenlopende bevindingen uit eerdere studies zijn gerubriceerd naar een aantal hoofdthema's en de (meetbare) indicatoren waarmee de thema's geoperationaliseerd zijn. Het gaat hierbij allereerst om relatief statische factoren die bekend zijn voor de ervaren weggebruiker, zoals verkeersintensiteit en wegkenmerken. Deze thema's zijn nu als variabelen onderzocht en meegenomen in LMS-BT. Daarnaast zijn er ook meer dynamische factoren die voor onverwachte vertragingen kunnen zorgen. Denk hierbij aan weersomstandigheden en incidenten zoals ongelukken of werkzaamheden. Tabel 1 geeft een overzicht van thema's en indicatoren die in meerdere studies worden genoemd. Enkele belangrijke referenties naar studies die hier dieper op ingaan zijn in de tabel opgenomen. Op basis van deze inventarisatie specificeren we verklarende variabelen die de analyse van reistijdbetrouwbaarheid (Hoofdstuk 4) kunnen verrijken. De vetgedrukte indicatoren in de tabel lijken ons met name relevant om op te nemen in de verklarende analyse. Deze selectie is gebaseerd op: verwacht effect, beschikbaarheid van data, en toevoeging van nieuwe inzichten ten opzichte van bestaande studies.

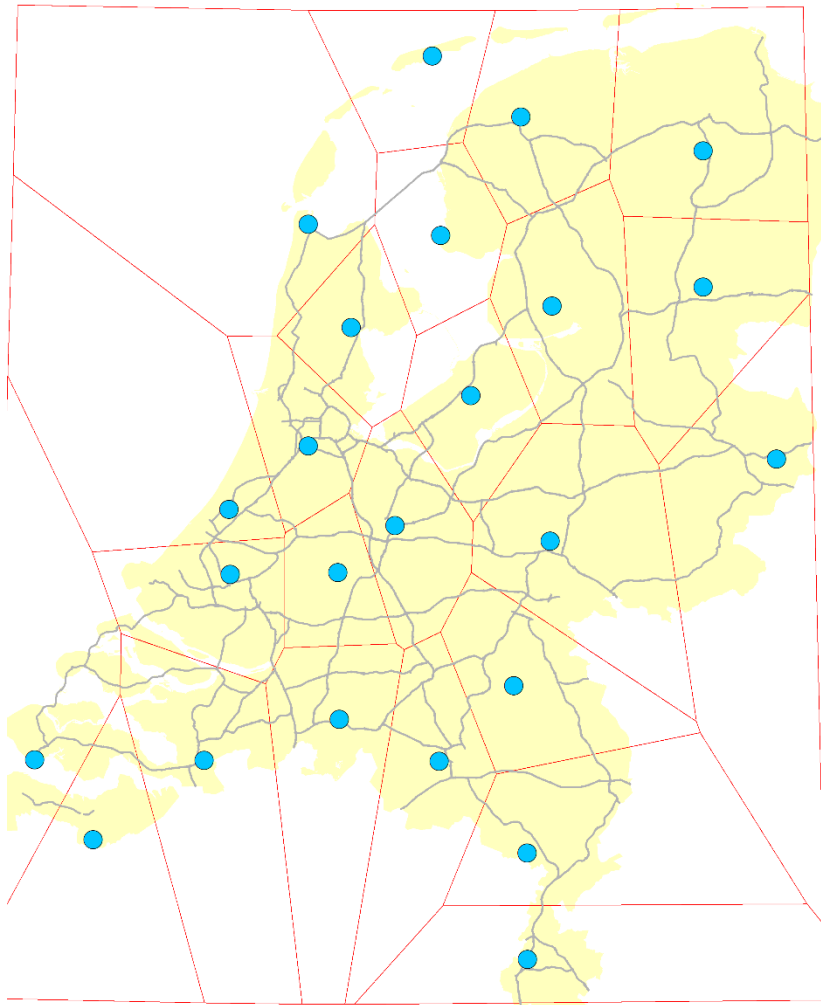
Tabel 1 Overzicht belangrijkste thema's en indicatoren in onderzoek naar reistijdbetrouwbaarheid. **Vetgedrukt** de indicatoren die we in de verklarende analyse willen gebruiken.

Thema	Indicator
Vertraging	Structurele vertraging
Verkeersintensiteit (Javid & Javid, 2017)	Intensiteit/ capaciteit
	Tijdstip (ochtend-, avondspits)
	Dag (werkdag, weekeinde)
Wegkenmerken	Lengte route (Caceres, Hwang & He, 2016)
	Aantal rijstroken (Carrion & Levinson, 2013)
	Verkeerslichten, tunnels, oprit
	Maximumsnelheid, snelheidscontrole, etc.
Weersomstandigheden (Caceres e.a., 2016; Maze e.a., 2006)	Onbewolkt, regen, sneeuw
	Zicht/ mist
	Hoge windsnelheden
	Nachtvorst
Incidenten (Li e.a., 2006)	Incidenten (ongelukken etc.) op de route
	Incident opruimtijd (Javid & Javid, 2017)
Werkzaamheden	Werkzaamheden op de route
	Verkeershinderklasse

Tot de variabelen waarvan effect verwacht wordt op reistijden en daarmee betrouwbaarheid behoren onder meer weersomstandigheden (zie onder meer Caceres e.a., 2016 en Maze e.a., 2006). Weergelateerde variabelen zijn ook aantrekkelijk omdat er in Nederland relatief veel data over beschikbaar is. Figuur 1 geeft een overzicht van de ruimtelijke spreiding van KNMI-weerstations waarvoor gegevens over regen, vorst en mist beschikbaar zijn op een hoge temporele (uur-basis) resolutie tijdens de onderzoeksperiode. Dergelijke data kan aan de onderzochte routes worden

toegekend op basis van nabijheid en beschrijft in zijn simpelste vorm het al dan niet voorkomen van specifieke omstandigheden die samenhangen met vertraging (zoals hevige regenval).

De dag-specifieke inventarisatie van werkzaamheden op het rijkswegennet is een andere, mogelijk relevante toevoeging aan de analyse van reistijdbetrouwbaarheid. Bijvoorbeeld om speciale omstandigheden te markeren en van de analyse te kunnen uitsluiten. Daarnaast zijn enkele indicatoren interessant om op te nemen omdat deze nog maar weinig aandacht krijgen in de wetenschappelijke literatuur. Het gaat hierbij met name om wegkenmerken als aantal rijstroken, verkeerslichten, snelheidsbeperkingen en snelheidscontroles. Zeker de implementatie van trajectcontrole is een wegkarakteristiek die nog weinig onderzocht is.



Figuur 1 Spreiding KNMI-weerstations waarvoor een complete set weersgegevens beschikbaar is voor de onderzoeksperiode.

3. Vaststellen routeset

Om een voor Nederland zo representatief mogelijk beeld te schetsen van deur-tot-deur reistijd-betrouwbaarheid hebben we in overleg met RWS-WVL een aantal criteria opgesteld waaraan de te selecteren routes moeten voldoen. Dit keuzeproces kent drie schaalniveaus. Allereerst kiezen we een set van 40 herkomst-bestemming (HB) relaties op het niveau van LMS-zones. Tussen deze relaties zoeken we de snelste routes over het LMS-netwerk en vervolgens bepalen we een of meer routealternatieven over het wegennetwerk waaraan de floating car data zijn gekoppeld.

Een belangrijk aspect van het gebruik van floating car data is, dat deze aan een eigen, op Open Street Map gebaseerd, netwerk is gekoppeld, dat elke paar weken vernieuwd wordt. Voor de gehanteerde onderzoeksperiode (2 september tot en met 13 december 2019) zijn dit vier verschillende versies. Deze netwerken wijken in detailniveau en ruimtelijke weergave af van het LMS-netwerk, dat gebaseerd is op het nationaal wegenbestand. Bijlage 1 beschrijft hoe de koppeling tussen beide bestanden tot stand gebracht is. Naast de geometrische weergave van het wegennet bestaat de floating car data uit tabellen die per minuut en per wegvak (van gemiddeld 40 meter) informatie bevatten over de gemiddelde snelheid en het aantal voertuigen (probes) waar die snelheid op gebaseerd is. Dat aantal wordt uitgedrukt in de zogeheten probe-coverage (dekkingsgraad) die varieert tussen 0 en 10. Dat getal geeft aan van hoeveel probes er snelheidsinformatie is opgenomen in de actuele snelheidsschatting. De gebruikte database middelt snelheidsschattingen op basis van de maximaal 10 meest recent gepasseerde probes. De gemiddelde snelheid per segment wordt bij voorkeur bepaald over waarnemingen binnen dezelfde minuut. Bij minder dan 10 waarnemingen per minuut worden waarnemingen uit voorliggende meetminuten gebruikt totdat er totaal 10 waarnemingen zijn of er meer dan 30 minuten teruggekeken moet worden. Bij geen probes in de voorgaande 30 minuten is de snelheidsschatting opgesteld op basis van de geregistreerde maximumsnelheid.

De volgende criteria zijn gedefinieerd om de representativiteit van de routes te waarborgen:

- 1) Ruimtelijke representativiteit van de HB-relaties;
- 2) Voldoende variatie in triplengte;
- 3) Preferentie voor de aandelen HWN en OWN per route;
- 4) Preferentie voor routes met verwachte vertragingen
- 5) Voldoende beschikbaarheid van floating car data;
- 6) Beperking van overlap tussen geselecteerde routes, door bij te houden of een link nog niet is gepasseerd door op het iteratie-moment k geselecteerde routes (P_l^k).

Voor het onderbouwen van ruimtelijke representativiteit (criterium 1) en variatie in triplengte (criterium 2) hebben we gebruik gemaakt van het 2019 databestand van het Onderzoek Onderweg in Nederland (ODiN, beschikbaar gesteld via DANS en beschreven in CBS 2020). Dit onderzoek beschrijft op basis van steekproeven de dagelijkse mobiliteit² van Nederlanders ouder dan 6 jaar op Nederlands grondgebied. De gegevens over 2019 hebben betrekking op 53.380 respondenten verdeeld over heel Nederland. Uit dit onderzoek selecteren we alle reguliere verplaatsingen (trips) met een motorvoertuig (inclusief motoren, bestelauto's en vrachtwagens, exclusief brommers, taxi's en touringcars) voor alle dagen van de week. Hieruit hebben we twee overzichten gecreëerd: de herkomsten van verplaatsingen naar provincie (Tabel 2) en de verdeling van verplaatsingen over verschillende lengteklassen (Tabel 3). Voor beide aspecten zijn de wegingsfactoren toegepast (FactorV) die het aantal verplaatsingen corrigeren voor oververtegenwoordiging van bepaalde doelgroepen, perioden en regio's in de steekproef. Beide tabellen geven de randvoorwaarden in

² De dagelijkse mobiliteit is door CBS (2020) gedefinieerd als "de reguliere mobiliteit inclusief vakantie-mobiliteit en exclusief (werk gerelateerde) serieverplaatsingen en mobiliteit met zware vrachtvoertuigen en exclusief vliegtuigen".

termen van aantallen HB-relaties aan die we op basis van ODiN vaststellen voor zowel de verdeling van herkomsten over provincies als de verdeling van verplaatsingen over verschillende afstandsklassen.

Representatieve aandelen HWN en OWN (criterium 3) en voldoende variatie in verwachte drukte/vertraging (criterium 4) leiden we af uit het LMS-netwerk van het jaar 2018. Dit netwerk beschrijft het type weg (waaruit aandeel OWN per route bepaald kan worden), en voertuigverlies-uren per dag en maximale verkeersintensiteiten tijdens de ochtend- en avondspits op basis van modelresultaten. De beschikbaarheid van floating car data (criterium 5) wordt afgeleid uit het attribuut probe-coverage dat per wegsegment aangeeft op hoeveel waarnemingen de gemiddelde snelheid gebaseerd is. We gebruiken hier alleen probe coverage tijdens de spitsperiodes, berekend voor alle vier de gebruikte floating car data netwerken. We passen geen harde minimumwaarde toe, maar selecteren bij voorkeur segmenten met betere dekking. Op deze wijze kunnen we logischer routes selecteren. Om overlap tussen routes zoveel mogelijk te voorkomen (criterium 6) wordt bijgehouden of een link al is gepasseerd door een eerdere route. Die link wordt daarna minder aantrekkelijk gemaakt voor mogelijk later te selecteren routes. In de volgende alinea gaan we nader in op de exacte specificatie van dit routeselectieproces.

Tabel 2 Aantal autoverplaatsingen onderverdeeld naar herkomst per provincie in ODiN-2019 en daarop gebaseerde aantal HB-relaties in deze studie. Betreft zowel trips met bestemmingen binnen dezelfde provincie als naar bestemmingen daarbuiten.

Provincie	Verplaatsingen	Aandeel (%)	HB-relaties
Groningen	2.017	3%	1
Friesland	2.507	3%	1
Drenthe	2.025	3%	1
Overijssel	4.618	6%	2
Flevoland	1.997	3%	1
Gelderland	8.830	12%	5
Utrecht	5.048	7%	3
Noord-Holland	10.855	15%	6
Zuid-Holland	18.605	25%	10
Zeeland	1.689	2%	1
Noord-Brabant	10.583	14%	6
Limburg	5.211	7%	3
<i>Totaal</i>	<i>73.985</i>	<i>100%</i>	<i>40</i>

Tabel 3 Aandeel triplengtes in totaal aantal trips in ODiN-2019 en daarop gebaseerde aantal HB-relaties in deze studie (in tegenstelling tot Tabel 2, zijn hier alleen trips langer dan 5 km opgenomen).

Afstand (km)	Frequentie	Aandeel (%)	HB-relaties
5 - 10	2.000	9%	4
10 - 20	7.517	33%	13
30 - 50	5.310	23%	9
> 50	7.912	35%	14
<i>Totaal</i>	<i>22.739</i>	<i>100%</i>	<i>40</i>

De selectie van HB-relaties voor deze studie is opgezet als optimalisatievraagstuk waarin het aantal herkomsten per provincie en het aantal trips per afstandsklasse als randvoorwaarden zijn opgenomen. Primaire routes worden geselecteerd uit de snelste routes tussen herkomst (*i*) en bestemming (*j*) LMS zones, volgens de maximumsnelheden in de avondspits volgens het LMS. Die routes zijn opgebouwd uit een sequentie van links / zodat elk herkomst-bestemmingspaar een eigen

verzameling heeft van links in haar route, $l \in R_{ij}$. Het selectieproces is vervolgens een iteratieve procedure waarbij, in k iteraties, op herkomst-bestemming (deur-tot-deur) niveau steeds de hoogst scorende route wordt geselecteerd uit de volgende berekening:

$$S_{ij} = Z_i^k A_{ij}^k D_{ij} V_{ij} (\%OWN_{ij}) (C_{ij}^k)^2 M_{ij}^{fcd} \quad (1)$$

Voor alle 1,975,430 HB-relaties die op basis van de 1406 LMS-zones gekozen kunnen worden berekenen we achtereenvolgens:

1. Of het aantal routes uit de herkomstprovincie nog niet uitgeput is in de iteratie (Z_i^k)
2. Of het aantal routes in de afstandsklasse nog niet uitgeput is in de iteratie (A_{ij}^k)
3. De maximale stroomdichtheid (in PersonenAutoEquivalent, PAE, per uur) van de route in de in de ochtend- of avondspits ($D_{ij} = (\sum_{l \in R_{ij}}^{l=1} \max(os_{druk_l}^{LMS}, as_{druk_l}^{LMS}) / (\sum_{l \in R_{ij}}^{l=1} L_l^{LMS})$) en de vertragingdichtheid (uitgedrukt in voertuigverliesuren, h) van de route per etmaal $V_{ij} = (\sum_{l \in R_{ij}}^{l=1} VVU_l^{LMS}) / (\sum_{l \in R_{ij}}^{l=1} L_l^{LMS})$, waarbij V altijd minimaal 1 is en L de lengte van de link is;
4. Het aandeel OWN in de route volgens $\%OWN_{ij} = (\sum_{l \in R_{ij}}^{l=1} L_l^{LMS} OWN_l^{LMS}) / (\sum_{l \in R_{ij}}^{l=1} L_l^{LMS})$
5. Het gemiddelde percentage probe-coverage ten opzichte van de maximale probe coverage in de floating car data-netwerkenrepresentatie van de kortste route per HB-relatie zoals die gematcht kon worden met het LMS-netwerk³ ($C_{ij}^k = \sum_{l \in ij}^{l=1} COVG_l^{fcd} P_l^k$), alsmede het criterium dat bijhoudt of een link nog niet gepasseerd is door een route die eerder is geselecteerd in het iteratieve proces (P_l^k) in voorgaande formule.
6. De matching score M_{ij} die aangeeft welk deel van de floating car data links waarschijnlijk accuraat gematcht zijn⁴ aan de links die deel uitmaken van het kortste pad tussen de HB relaties in ij op het LMS-netwerk, op basis van $M_{ij} = (\sum_{l \in ij}^{l=1} L_l^{fcd}) / (\sum_{l \in ij}^{l=1} L_l^{LMS}) \leq 1$. Dit criterium operationaliseert een preferentie voor routes waarvan de LMS links met veel zekerheid gekoppeld kunnen worden aan links in het FCD-netwerk.

Vervolgens kiezen we de 40 HB-relaties met de hoogste scores. Tabel 3 geeft een overzicht van de 40 geselecteerde HB-relaties. Van de gekozen HB-relaties bewaren we de bijbehorende kortste route volgens het LMS-netwerk, en de route in tegengestelde richting om ook tegenovergestelde richtingen in kaart te brengen voor de tegenovergestelde spits.

Keuze route alternatieven

Na selectie van de 40 primaire routes zijn voor 20 herkomst-bestemmingsparen ook alternatieve routes gezocht. Daarvoor is voor alle combinaties van die geselecteerde paren een tweede sequentie van links gedefinieerd, op basis van wederom de kortste reistijden in de avondspits volgens het LMS-netwerk. Echter, op alle links die in de geselecteerde primaire routes zijn gebruikt, is de reistijd artificeel met 25% verhoogd om zo naar plausible alternatieven te zoeken. Na selectie van route alternatieven is deze verhoging weer ongedaan gemaakt. Vervolgens is de optimalisatiestrategie toegepast die ook bij het definiëren van de primaire routes is bepaald. De selectie gebeurt op basis van de 40*40 mogelijke herkomst-bestemmingsparen uit de geselecteerde primaire routes. Het overlapcriterium wordt nog steeds bewaakt door middel van P_l^k (zie stap 5 in de voorgaand beschreven methode) welke wordt overgenomen van de laatste iteratie van de primaire selectie. Een additioneel criterium dat wordt toegepast is dat een specifiek herkomst-bestemmingspaar ook is geselecteerd als primair pad. De eerdere criteria met betrekking tot herkomstprovincie en afstandsklasse zijn in deze stap buiten beschouwing gelaten. Figuur 2 geeft een overzicht van de gekozen routes.

³ We beschouwen hiervoor alle vier de floating car dataversies voor de complete onderzoeksperiode.

⁴ Het matching proces is beschreven in Bijlage 1, sectie "Koppeling van FCD dekkingsgraad aan LMS netwerk".

Tabel 4 Overzicht van de geselecteerde herkomst-bestemmingscombinaties.

ID	Herkomst	Bestemming	Afst. ¹	Alt. ²
1	Leek, Westerkwartier	Het Hogeland	48,5	2
2	Oosterparkwijk, Groningen	Dokkumer Ie e.o., Leeuwarden	65,6	2
3	Noordoost, Groningen	Tweede Exloermond, Borger-Odoorn	54,4	2
4	Staphorst (kern), Staphorst	Sint Jansklooster, Steenwijkerland	25,0	2
5	Balkbrug, Hardenberg	Steenwijk, Steenwijkerland	42,9	2
6	Schuytgraaf, Arnhem	Velp, Rheden	10,8	4
7	Wehl, Doetinchem	Alverna, Wijchen	54,5	4
8	Warnsveld, Zutphen	Buitengebied Heteren, Overbetuwe	48,4	2
9	Alverna, Wijchen	Rossum, Maasdriel	52,3	2
10	Appel, Driedorp, Kruishaar, Prinsenkamp en Slichtenhorst, Nijkerk	Vuren, West Betuwe	62,3	2
11	Barneveld, Barneveld	Bilthoven Noord, De Bilt	40,8	4
12	Waardenburg en Opijnen, West Betuwe	Lopik, Benschop en Polsbroek, Lopik	35,2	2
13	Vleuten-De Meern, Utrecht	Zeist-Noord, Zeist	17,2	2
14	Nieuw West, Zaanstad	Banne Buiksloot, Amsterdam	10,2	4
15	Amsterdamse Bos, Amstelveen	Grachtengordel-Zuid, Amsterdam	15,6	2
16	Buitenveldert-West, Amsterdam	Slotervaart Noord, Amsterdam	6,2	4
17	Apollobuurt, Amsterdam	Sloter-/Riekerpolder, Amsterdam	6,3	2
18	Apollobuurt, Amsterdam	Bedrijventerreinen, Diemen	8,0	2
19	Noordelijke IJ-oever West, Amsterdam	Schiphol, Haarlemmermeer	24,6	4
20	Heemstede-Centrum, Heemstede	Merenwijkdistrict, Leiden	33,3	4
21	Bos- en Gasthuisdistrict, Leiden	Oegstgeest, Oegstgeest	5,0	2
22	Aalsmeer, Aalsmeer	Leimuideren, Kaag en Braassem	13,5	4
23	Kraayenstein en Vroondaal, 's-Gravenhage	Rijswijk	9,2	2
24	Wateringse Veld, 's-Gravenhage	Voorburg Nrd, Leidschendam-Voorburg	10,4	4
25	Keizer Karelpark, Amstelveen	Aarlanderveen, Alphen aan den Rijn	35,9	2
26	Indische Buurt Oost, Amsterdam	Zevenhoven, Nieuwkoop	38,0	4
27	Bedrijventerrein Schieveen, Rotterdam	Prins Alexander, Rotterdam	9,2	4
28	Prins Alexander, Rotterdam	Rotterdam Centrum, Rotterdam	17,3	2
29	Kralingen-Crooswijk, Rotterdam	Nieuwland, Schiedam	9,4	2
30	Kortenhoef, Wijdmeren	Kerkwerpe, Schouwen-Duiveland	149,0	4
31	Rokkeveen, Zoetermeer	Eethen, Altena	79,8	4
32	Erlecom, Berg en Dal	Industrieterreinen-Noord, Oss	39,1	4
33	Waalwijk, Waalwijk	Veghel, Meierijstad	38,3	4
34	Achtse Molen, Eindhoven	Veldhoven, Veldhoven	10,8	2
35	Mariahout, Laarbeek	Maarheeze, Cranendonck	31,2	4
36	Meerhoven, Eindhoven	Geldrop, Geldrop-Mierlo	14,7	4
37	Graft-De Rijp, Alkmaar	Mook, Mook en Middelaar	157,1	4
38	Meerhoven, Eindhoven	Wanssum, Venray	51,2	2
39	Achtse Molen, Eindhoven	Roggel, Leudal	56,2	4
40	Bloemendaal, Bloemendaal	Almere Buiten, Almere	61,4	4

Opmerkingen:

¹ Gemiddelde afstand over alle alternatieven over het LMS-netwerk in kilometers. De gemiddelde lengte over de gehele set aan alternatieven is 39,6 kilometer.

² Alt. geeft het aantal routealternatieven aan per herkomst-bestemmingspaar. Dat is tenminste 2 omdat voor alle routes een retourroute is beschouwd. Voor 20 routes zijn er ook een alternatieve route met retourversie beschouwd. In totaal beschouwen we dus 120 route alternatieven.

Terugvertaling geselecteerde routes naar floating car data segmenten

Voor het verwerken van de reistijden op de geselecteerde routes moeten vervolgens de binnen het LMS-netwerk gedefinieerde routes weer worden terugvertaald naar routes binnen de vier geanalyseerde floating car data netwerken. Om dat te doen, is het 'terugvertaalinstrument' dat is gebruikt in een eerder rapport (Koomen en Jacobs-Crisioni, 2021) nog verder verbeterd. In dat instrument wordt getracht de LMS-routes te reconstrueren in het floating car data netwerk op basis van geografische afstanden. Dit instrument heeft veel overeenkomsten met Markov Chain traceermethoden zoals beschikbaar zijn in bijvoorbeeld de Meili API van Valhalla. Een uitgebreide beschrijving van dit proces is opgenomen in Bijlage 1. Hier benoemen we kort de belangrijkste stappen in dat proces.

Binnen het terugvertaalinstrument worden de LMS links eerst beschreven als dynamische punten, waarbij er op elke link om de 25 meter een punt wordt gezet. Een selectie van die punten, zoals die horen bij een specifieke route (primaire of alternatieve, heen of weer), wordt daarop geïsoleerd om specifiek die route te traceren. Vervolgens wordt de afstand berekend vanaf het middelpunt van elk floating car data segment naar het dichtstbijzijnde punt binnen de geselecteerde LMS-puntenwolk. Doordat deze segmenten allemaal erg kort zijn (gemiddeld 40 meter), zijn middelpunten altijd een behoorlijk nauwkeurige beschrijving van de locatie van het floating car data segment.

Elk floating car data segment krijgt een score op basis van het kwadraat van de afstand naar een LMS-lijnenpunt. Vervolgens wordt een richtingsgevoelig kortste-pad algoritme gebruikt om van herkomst- naar bestemmingspunt te komen, waarbij de afstand naar LMS-lijnenpunten wordt geminimaliseerd. Het is belangrijk om te benadrukken dat, net als in het LMS-netwerk, floating car data segmenten een berijdbare richting beschrijven, zodat een ongescheiden weg met tweerichtingsverkeer wordt beschreven door twee identieke segmenten met tegenovergestelde richtingen. Daardoor is een richtingsgevoelig kortste-pad algoritme nodig.

Als precieze herkomst- en bestemmingspunten hebben we de weglocaties genomen waar, in de gekozen routes, het LMS-netwerk de zoneverbindingen aan het wegennetwerk koppelt. Doordat LMS-netwerk en floating car data netwerken allebei ruimtelijk zeer accuraat zijn, kan dit verbinden van route herkomst- en bestemmingspunten foutloos. Wel zijn een aantal fouten ontstaan doordat binnen de floating car data netwerken links zijn gecodeerd met verkeerde richtingen, of omdat de LMS-route een onwaarschijnlijke maar volgens de data kortere route volgt (bijvoorbeeld over aansluitende af- en opritten in de binnenbocht van een snelweg, als die dezelfde gecodeerde maximumsnelheid hebben). Om dit soort fouten op te lossen zijn alle getraceerde routes handmatig gecontroleerd, en waar nodig via codelijsten de fouten opgelost door segmenten softwarematig te draaien of ongewenste routeafwijkingen te bestraffen.



Figuur 2 Overzicht geselecteerde routes

4. Samenstellen reistijd database voor de vastgestelde routeset

Voor de routes, die in de vorige stap zijn gekozen, worden de reistijden per vertrekminuut bepaald tussen 6:00 en 19:00 uur op werkdagen. De hiervoor gevolgde werkwijze is uitgebreid beschreven en getest in een pilotstudie (Koomen en Jacobs-Crisioni, 2021). De basisstappen hierin zijn:

- Selectie van de wegsegmenten in de floating car data sets die deel uitmaken van de te analyseren routes. De relevante observaties worden voor elke dag in separate tekstbestanden (in .csv formaat) weggeschreven middels een zelfontwikkeld R-script.
- Gebruik van deze tekstbestanden als invoerdata in een eigen analyse-omgeving die door de VU is ontwikkeld met behulp van de open-source GeoDMS software.
- Optellen van de gemeten reistijden voor het rijden over de afzonderlijke segmenten waaruit de route bestaat. Doordat de tijd onderweg voortschrijdt, en reistijden verschillen per moment, zijn de beoogde reistijden van trajectoriën wezenlijk anders dan de optelsom van alle reistijden op alle betrokken segmenten op een specifiek moment. We hanteren hiervoor de eerder beschreven dynamische array methode (Koomen en Jacobs-Crisioni, pag. 16).

Deze methode creëert per route aparte bestanden voor de in de tijd opeenvolgende trips, waarin de verstreken reisduur voor de verschillende wegsegmenten wordt vastgelegd. Daarmee is het mogelijk het reistijdverloop over de dag en route te visualiseren. Bij wijze van voorbeeld toont Figuur 3 het verloop van achtereenvolgende trips tussen Leek en Het Hogeland (Route 1, retourversie) op 14 oktober 2019 van 6:00 tot 19:00 uur. Algemene resultaten voor vrijwel alle routes zijn ook te vinden in Tabel 5.



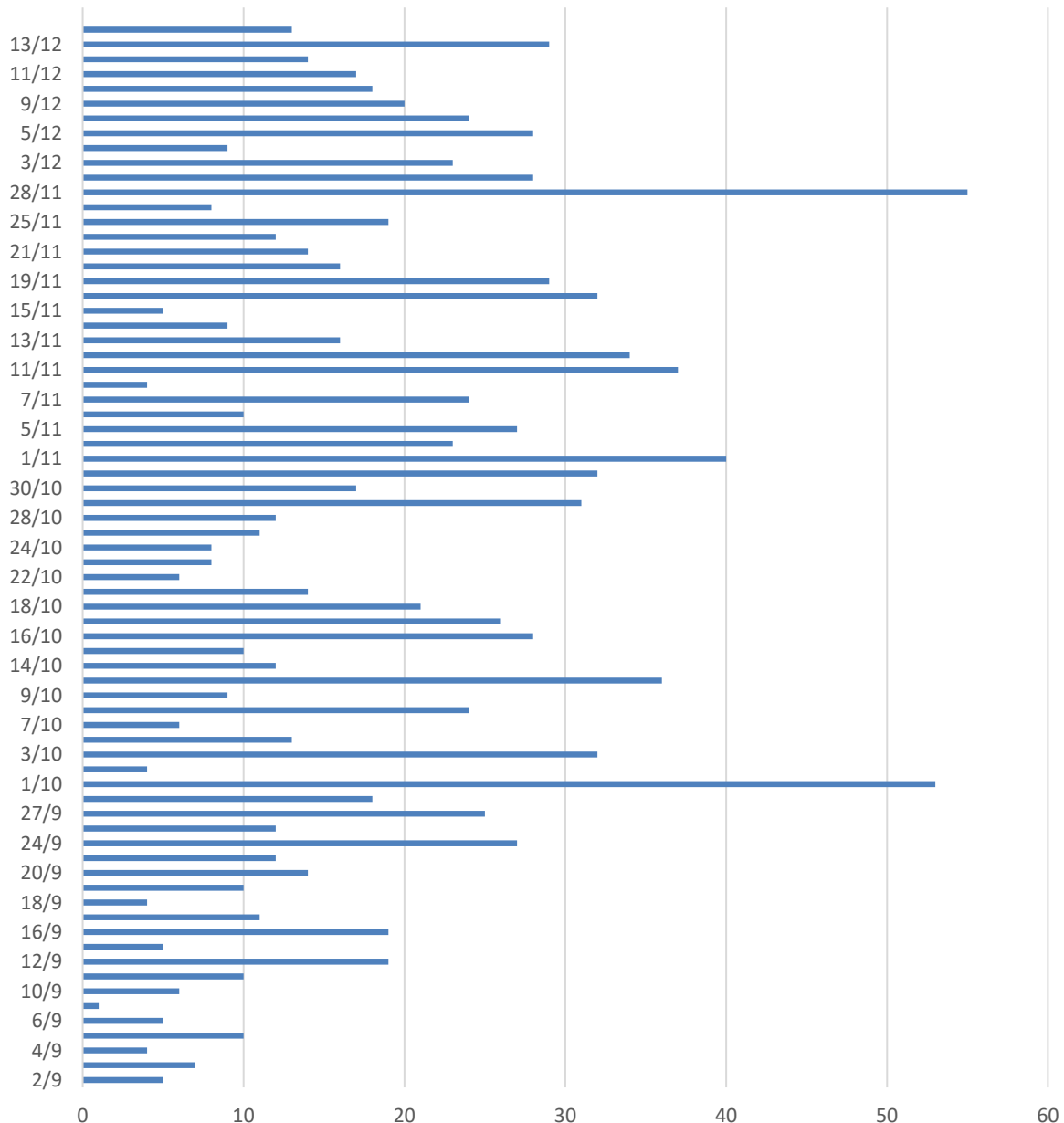
Figuur 3 Weergave van het reisverloop op de primaire route van Leek naar Het Hogeland gedurende de ochtendspits van 14 oktober 2019, bij steeds later vertrek in de ochtendspits vanaf 6:00 uur. De X-as geeft afgelegde afstand weer, de Y-as de tijd vanaf 6:00 tot 19:00 uur. De kleurtoon van de lijnen geeft de totale reistijd van een trip weer, waarbij donkerdere lijnen een relatief lange reistijd aangeven. Het witte vlak midden op de dag duidt op een incident waarbij het verkeer tijdelijk opgehouden werd.

Tabel 5 Berekening van de gemiddelde (Gem.) en freeflow (Free) reistijden in minuten op de geselecteerde routes, op basis van trajectoriën met floating car data snelheidsdata. Voor lege cellen zijn geen routes geselecteerd. De reistijden zijn bepaald over alle werkdagen in de studieperiode en alle beschikbare minuten (tussen 6.00 en 19.00 uur). Het volgende hoofdstuk gaat nader in op berekenen van de freeflow reistijd.

ID	Herkomst	Bestemming	Primair heen		Primair terug		Alternatief heen		Alternatief terug	
			Gem.	Free	Gem.	Free	Gem.	Free	Gem.	Free
1	Leek	Het Hogeland	37,6	32,5	37,5	32,2				
2	Groningen	Leeuwarden	48,4	39,8	46,7	39,2				
3	Groningen	Tweede Exloermond	42,6	35,9	40,7	34,9				
4	Staphorst	Sint Jansklooster	21,7	18,7	22,4	19,0				
5	Balkbrug	Steenwijk	32,8	27,8	31,7	26,9				
6	Arnhem	Velp	13,2	9,5	15,0	10,1	16,2	11,6	17,7	12,3
7	Wehl	Alverna	51,0	38,9	48,8	39,2	52,6	37,7	50,7	38,9
8	Warnsveld	Heteren	41,3	35,4	43,2	33,8				
9	Alverna	Rossum	40,7	31,6	41,9	32,0				
10	Nijkerk	Vuren	51,1	41,6	55,3	43,0				
11	Barneveld	Bilthoven Noord	32,9	24,8	35,9	26,1	34,0	26,5	40,1	29,2
12	Waardenburg	Lopik	29,5	21,1	26,3	21,4				
13	Vleuten-De Meern	Zeist-Noord	21,4	14,1	19,2	14,0				
14	Zaanstad	Amsterdam	11,7	8,9	10,2	8,1	13,6	9,4	12,9	9,0
15	Amstelveen	Amsterdam	24,7	15,4	23,6	14,8				
16	Buitenveldert, A'dam	Slotervaart, A'dam	10,7	6,3	11,4	6,8	10,5	6,1	11,3	6,5
17	Apollobuurt, A'dam	Sloterpolder, A'dam	13,1	6,4	8,9	5,5				
18	Apollobuurt, A'dam	Diemen	10,7	7,1	11,0	7,7				
19	Noord IJ-oever, A'dam	Schiphol	29,1	19,2	42,6	27,8	22,9	18,5	26,7	20,1
20	Heemstede	Leiden	31,0	24,0	33,3	23,0	34,9	27,4	34,8	27,6
21	Leiden	Oegstgeest	9,0	5,6	7,8	5,2				
22	Aalsmeer	Leimuiden	13,2	9,6	16,1	10,8	14,0	10,1	17,4	11,6
23	Den Haag	Rijswijk	14,1	9,8	15,2	9,7				
24	Wateringse Veld, DH	Voorburg	13,5	9,1	13,6	9,3	10,3	7,0	11,0	7,8
25	Amstelveen	Aarlanderveen	32,3	24,4	32,6	25,0				
26	Amsterdam	Zevenhoven	38,1	29,7	39,7	30,4	40,8	30,1	40,1	30,3
27	Schieveen, R'dam	Prins Alexander, R'dam	12,9	8,6	12,9	8,9	12,7	8,5	12,6	8,3
28	Prins Alexander, R'dm	Rotterdam Centrum	21,0	13,3	19,2	13,5				
29	Kralingen, R'dam	Schiedam	17,8	10,1	12,9	8,4				
30	Kortenhoef	Kerkwerve	115,1	91,6	116,5	91,7	127,0	98,7	131,6	98,7
31	Zoetermeer	Eethen	67,6	50,4	72,7	52,1	72,9	53,1	71,2	50,2
32	Erlecom	Oss	41,3	29,1	40,9	29,0	42,0	29,2	41,2	29,2
33	Waalwijk	Veghel	29,5	25,1	29,6	25,4	35,9	29,4	36,2	29,4
34	Eindhoven	Veldhoven	14,0	9,3	14,3	9,7				
35	Mariahout	Maarheeze	31,2	22,7	31,2	22,7	38,5	29,1	31,8	23,4
36	Meerhoven, Eindh.	Geldrop	17,3	11,8	17,1	11,7	17,0	13,6	16,4	13,4
37	Graft-De Rijp	Mook	114,9	90,7	109,4	89,3	115,4	92,0	118,7	93,2
38	Meerhoven, Eindh.	Wanssum	58,0	43,6	57,4	43,7				
39	Achtse Molen, Eindh.	Roggel	48,8	37,5	46,0	37,4	52,9	38,7	50,3	38,1
40	Bloemendaal	Almere Buiten	53,1	42,3	53,9	42,6	57,7	47,0	55,2	43,0

Per route resulteert een uitgebreide set aan uitkomsten waarin de totale reistijd per afzonderlijk dag en vertrekminuut is vastgelegd. In totaal bevat de routedatabase 15 waarnemingen per kwartier, 4 kwartieren in een uur, 13 uren op een dag, 5 dagen in de week en 15 weken voor 120 route-alternatieven. Dat zijn in theorie totaal 7.020.000 afzonderlijke reistijden. Drie dagen zijn echter uitgesloten, omdat de floating car data netwerken op die dagen wisselde. Dit betreft 25 september, 11 oktober en 26 november. Er is voor gekozen deze dagen volledig uit te sluiten omdat de

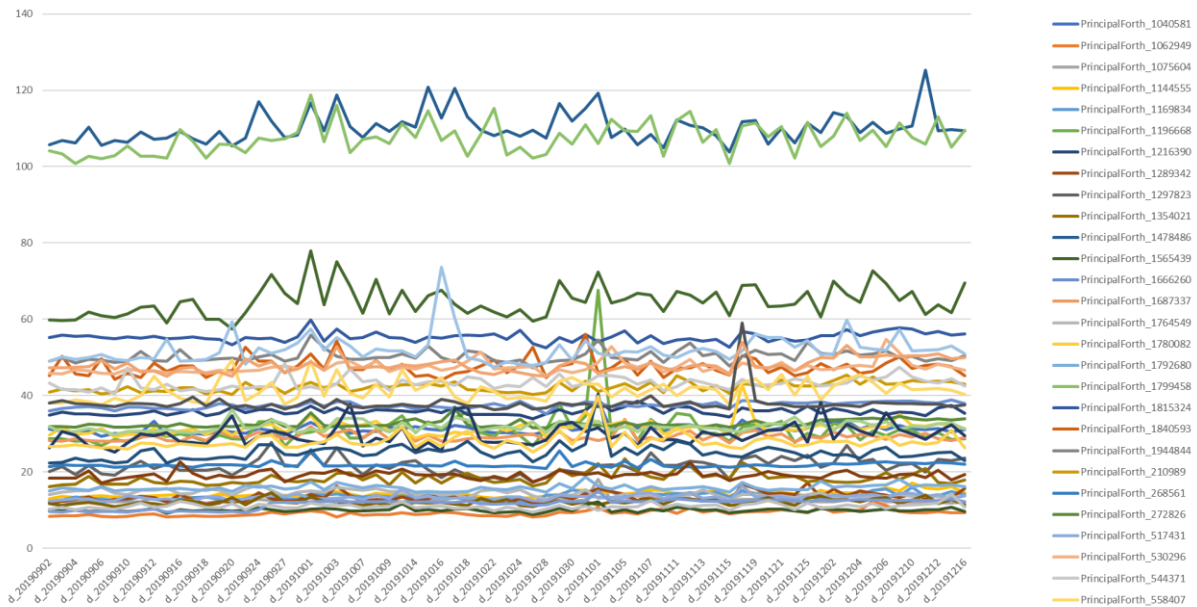
technische uitdaging van het doorzetten van trajectoriën met verschillende netwerken in onze inschatting niet opweegt tegen de meerwaarde van observaties op deze dagen. In de uiteindelijke analyse valt daarnaast nog een beperkt aantal waarnemingen weg. Allereerst omdat we in navolging van Kouwenhoven en Bel (2014, pag. 7) de extreme waarden weglaten. Deze worden gedefinieerd als reistijden die groter zijn dan anderhalf keer de gemiddelde reistijd, en die tevens meer dan drie keer de standaardafwijking hoger liggen dan de gemiddelde reistijd. In navolging van de eerdere studie sluiten wij de gehele dag uit waarop een extreme reistijd is geconstateerd. Het aantal routes met extreme reistijden verschilt sterk per dag. Op 9 september is maar één route-alternatief uitgesloten; maar op 28 november zijn dat er 55 en op 1 oktober 53 (van de 120)⁵. De variatie in uitsluitingen per datum is te vinden in Figuur 4.



Figuur 4 Aantal route-alternatieven dat per dag is uitgesloten wegens het voorkomen van extreme waarden.

⁵ 28 november was een dag met veel regen en ongelukken (<https://www.nhnieuws.nl/nieuws/257523/lange-files-door-regen-en-ongelukken-vooral-rond-amsterdam-erg-druk>), 1 oktober kende de langste ochtendspits ooit, dankzij massaal boerenprotest (<https://www.rtlnieuws.nl/editien/artikel/4868296/protest-den-haag-en-boeren-op-de-weg-dit-vindt-nederland-ervan>).

Figuur 5 geeft bij wijze van illustratie een overzicht van de gemiddelde, ongewogen reistijden per dag voor een selectie van routes. Op basis van de verzamelde reistijden worden diverse kenmerkende statistieken bepaald: freeflow reistijd, totale onbetrouwbaarheid (standaardafwijking van de hele reistijdverdeling), reguliere reistijd onbetrouwbaarheid (standaardafwijking van de reistijdverdeling zonder extreme gebeurtenissen) en verwachte reistijdvertraging (specifiek voor een werkdag en kwartier tijdvak). Statistieken met betrekking tot onbetrouwbaarheid, reistijdverwachting en vertraging zijn bepaald zoals voorgesteld door Kouwenhoven en Bel (2014), zoals nader wordt toegelicht in het volgende hoofdstuk.



Figuur 5 Gemiddelde reistijd per dag voor een selectie van 28 routes. Kleuren geven afzonderlijke routes aan.

5. Vergelijk met voorgaande studie

Op basis van de reistijdendatabase voor de set van representatieve routes is het verband tussen vertraging, betrouwbaarheid en andere invloedsfactoren geschat. In deze studie hanteren we zo veel mogelijk dezelfde uitgangspunten als in de eerdere studie (Kouwenhoven en Bel, 2014) die ten grondslag ligt aan de huidige rekenmodule LMS-BT (versie 2 t/m 4). We gebruiken dezelfde definitie van begrippen als vertraging en onbetrouwbaarheid en hetzelfde aggregatieniveau (werkdag-gemiddelde en tijdvakken als ochtend- en avondspits). Daarbij starten we onze analyses met dezelfde verklarende variabelen (wegtype) en hanteren we dezelfde criteria voor het uitsluiten van extreme waarden (meer dan 3 standaarddeviaties). Maar op enkele aspecten zijn er ook afwijkingen.

Het meest in het oog springende verschil is natuurlijk het gebruik van floating car data als bron voor reistijdeninformatie in deze studie. Deze data maakt het mogelijk om complete deur-tot-deur routes te onderscheiden, waar de oorspronkelijke LMS-BT studie reistijdbetrouwbaarheid separaat onderzocht voor trajecten op autosnelwegen en overige wegen. Dat onderscheid hing samen met de toen beschikbare data: tellusdata voor autosnelwegen en automatische kentekenherkennings-camera's op de overige wegen. Die wegindeling in deze eerdere studie wijkt ook iets af van het onderscheid in hoofdwegennet (HWN) en onderliggend wegennet (OWN) dat we in deze studie hanteren op basis van de classificatie in het LMS netwerk en onderliggend Nationaal Wegenbestand (NWB).

Ook het aantal onderzochte routes en hun lengtes verschilt zoals in Tabel 6 is weergegeven. De gemiddelde lengte van de routes is vrijwel gelijk, maar in de huidige studie is deze lengte verdeeld over een HWN en OWN deel, waar die in de vorige studie vrijwel alleen over autosnelwegen leidde. De aparte OWN routes waren in de vorige studie maar heel kort: gemiddeld krap 6 km tegenover 15 km in deze studie. Daarnaast heeft de floating car data een hogere ruimtelijke resolutie (een waarneming per wegsegment van circa 40 meter) en een hogere temporele resolutie (elke minuut een waarde) dan de data zoals die in de vorige studie is toegepast.

In dit hoofdstuk onderzoeken we in hoeverre de huidige studie afwijkt van de vorige. Daartoe gaan we allereerst dieper in op de basisbegrippen die in deze studie centraal staan. Daarna verklaren we variatie in reistijdbetrouwbaarheid op dezelfde wijze als de oorspronkelijke studie en vergelijken we de uitkomsten. Vervolgens gaan we dieper in op de mogelijke oorzaken voor de gevonden verschillen. In de volgende hoofdstukken breiden we de verklarende analyse van reistijdonbetrouwbaarheid verder uit en doen we een voorstel hoe LMS-BT aangepast kan worden.

Tabel 6 Overzicht opgenomen routes in deze studie (VU, 2022) en de voorgaande studie van Kouwenhoven en Bel (2014) die gebruikt is voor de huidige versie van LMS-BT.

Studie	Kenmerk route	Totaal	HWN/ASW deel	OWN deel	Gecombineerd
VU (2022)	aantal routes	120	120	120	120
	gemiddelde lengte (km)	39,6	24,3	15,4	39,6
LMS-BT (2014)	aantal routes	300	250	40	10
	gemiddelde lengte (km)	38,6	41,7	5,6	39,6

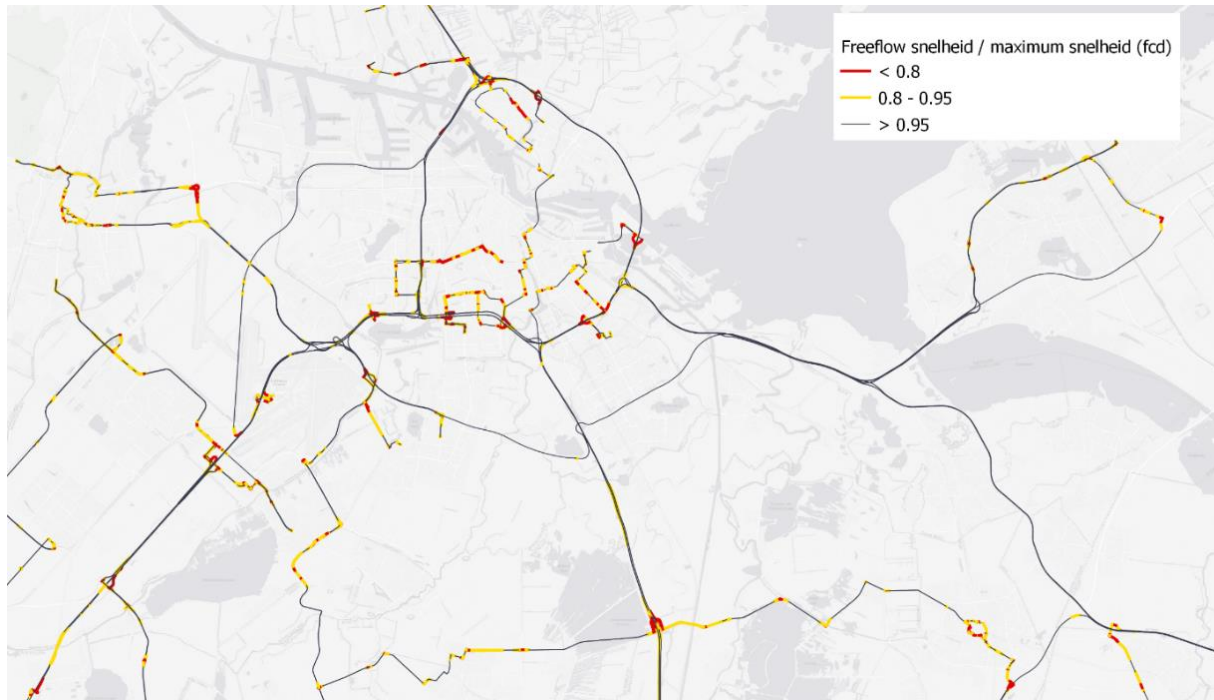
5.1. Reistijden, vertraging en onbetrouwbaarheid

Voordat we de resultaten van de verklarende analyse beschrijven, gaan we eerst nader in op de bepaling van reistijden, vertraging en onbetrouwbaarheid die in de analyse centraal staan.

Freeflow reistijd

Een eerste indicator van reistijd die berekend is, betreft de freeflow reistijd: de tijd die het afleggen van een route kost als het verkeer maximaal doorrijdt. De freeflow reistijd is gemeten als de som van

freeflow reistijden voor elk afzonderlijk segment dat deel uitmaakt van een route. Freeflow reistijden zijn bepaald op basis van het 95^e percentiel van snelheden die zijn geregistreerd op een segment in het geanalyseerde tijdvak. Waar geen registraties beschikbaar zijn, is de in de floating car data opgenomen maximumsnelheid aangehouden. Figuur 6 illustreert de gevonden freeflow snelheden voor de geselecteerde routes voor een gebied rond Amsterdam. De figuur maakt duidelijk dat vooral bij opritten en binnenstedelijke wegsegmenten de maximale snelheid niet wordt gehaald. De freeflow reistijden worden gebruikt om de vertraging per route te kunnen vaststellen.



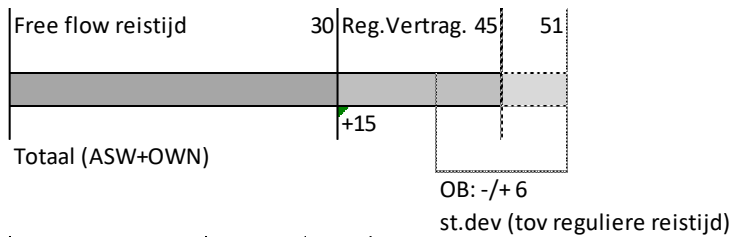
Figuur 6 Freeflow snelheden ten opzichte van de vigerende maximumsnelheid op de geselecteerde routes.

Reistijdverwachting en verwachte vertraging

De grondslag voor de analyse van betrouwbaarheid wordt gevormd door de beschrijving van verwachte reistijden. LMS-BT gaat uit van goed geïnformeerde reizigers die per werkdag en per kwartiertijdvak een bepaalde verwachte reistijd hebben. Deze verwachting wordt per werkdag en kwartier afgeleid uit de vier voorafgaande weken en de vier daaropvolgende weken. Anders gezegd, er wordt een voortschrijdend (lopend) gemiddelde van de reistijd bepaald waarin telkens negen weken beschouwd worden⁶. Een verschil met de eerdere berekeningen voor LMS-BT is dat wij onze reistijd per kwartier baseren op 15 reistijden (vertrek op elke individuele minuut) en niet op een enkele reistijd per kwartier. Op basis van de verwachte reistijd is het mogelijk de reguliere, te verwachten vertraging te bepalen. Deze definiëren we als de gemiddelde tijd die een reis voor een specifieke dag en kwartier langer duurt dan de freeflow reistijd. Vervolgens wordt voor elke reisminuut de berekende reistijd vergeleken met de voor dat kwartier verwachte reistijd. De standaardafwijking van deze verwachte reistijd noemen we onbetrouwbaarheid.

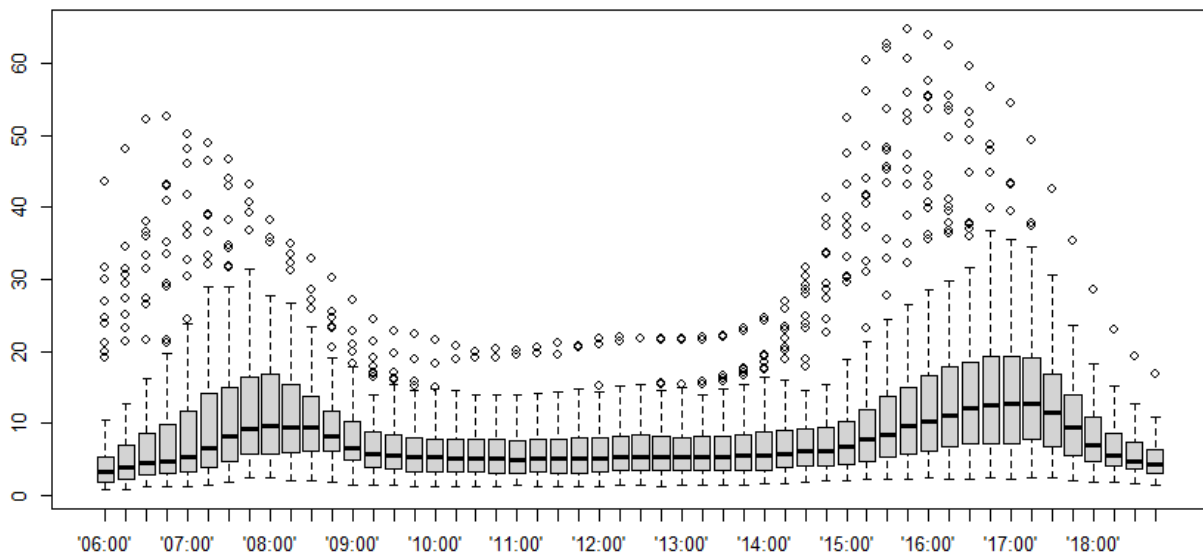
Figuur 7 geeft een visueel overzicht van deze begrippen. In deze studie bepalen we de onbetrouwbaarheid per route en kwartier. Deze is bepaald als standaarddeviatie ten opzichte van gemiddelde reguliere reistijden, die zijn vastgesteld per route, werkdag, en vertrekkwartier. Bij het aggregeren naar onbetrouwbaarheid maken we daarna geen onderscheid in dagen van de week, zodat onze onbetrouwbaarheidsschatting een generalisatie is van alle waarnemingen van alle weken en werkdagen binnen de studieperiode die betrekking hebben op het tijdvak 6:00-6:15 uur en een specifieke route.

⁶ Waarbij in navolging van Kouwenoven en Bel (2014, pag. 7) de extreme waarden zijn weggelaten.

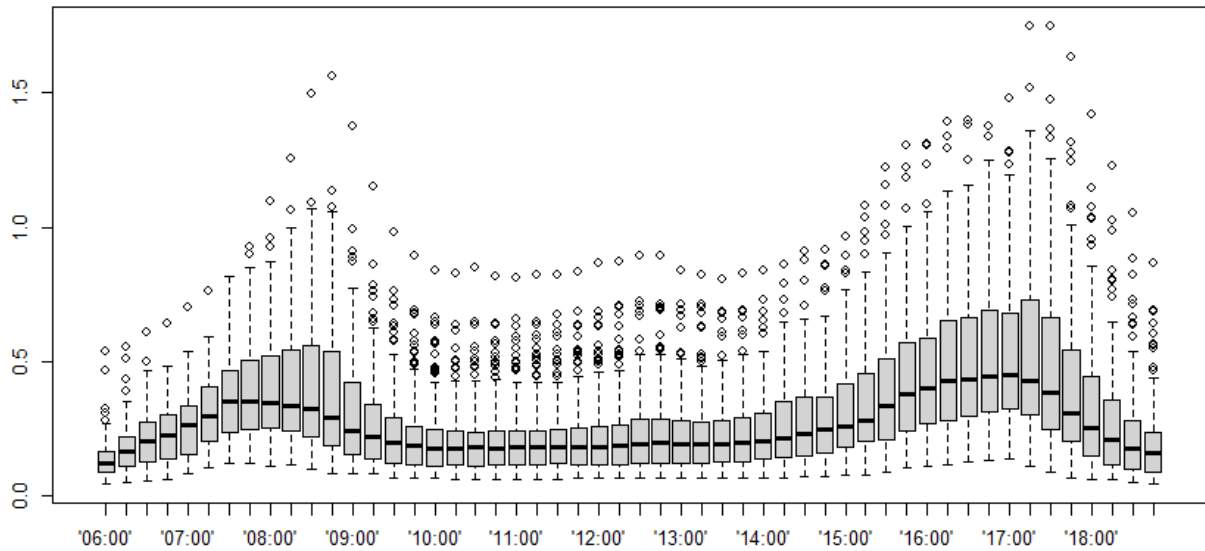


Figuur 7 De basisconcepten in deze analyse geïllustreerd met een getalsmatig voorbeeld voor een fictieve route. De freeflow reistijd (hier 30 minuten) en de reguliere vertraging (15 minuten) vormen samen de reguliere reistijd (45 minuten). De standaardafwijking van alle waargenomen reistijden ten opzichte van de reguliere reistijden noemen we de onbetrouwbaarheid (OB; in dit geval +/- 6 minuten).

Om meer gevoel te krijgen voor de variatie in waargenomen reistijden kijken we allereerst naar de verwachte vertragingen. Figuur 8 geeft aan hoe de verwachte reistijdvertraging fluctueert over de dag. Hierbij zijn de vertragingen over alle werkdagen gemiddeld. Deze figuur toont de kwartier-specifieke, werkdag-gemiddelde vertragingen voor elk van de 120 route-alternatieven die in deze studie onderzocht zijn. De figuur laat een zeer herkenbaar beeld zien. Niet alleen neemt de gemiddelde vertraging toe in de beide spitsperiodes, ook de spreiding rond dat gemiddelde wordt groter. De variatie in reguliere vertraging is nog geprononceerder als gecorrigeerd wordt voor het verschil in routelengte (Figuur 9). In deze weergave komt het asymmetrische karakter van de spits tot uiting: een langzaam opbouwende vertraging tot de piek van de spits en daarna snel afnemend.



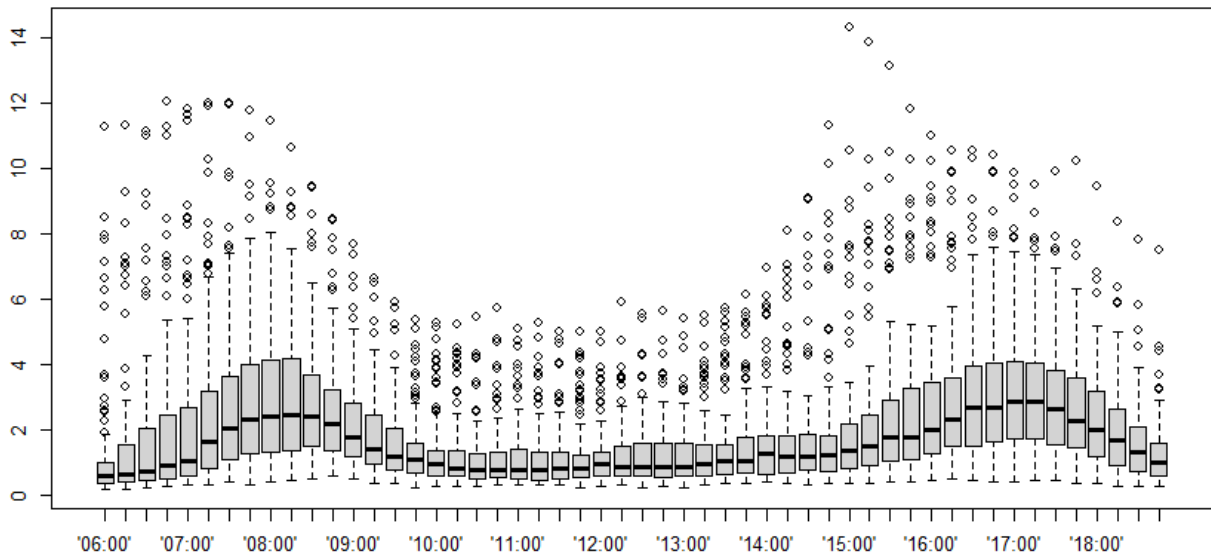
Figuur 8 De verdeling van verwachte vertragingen per vertrekkwartier en route in minuten. De horizontale lijn geeft de mediaanwaarde aan, de gesloten box het bereik van het tweede en derde kwartiel, de gestreepte lijn het eerste en vierde kwartiel, de daarbuiten gelegen extreme waarden (bolletjes) betreffen veelal relatief lange routes.



Figuur 9 Verwachte vertragingen per vertrekkwartier uitgedrukt in minuten per afgelegde kilometer.

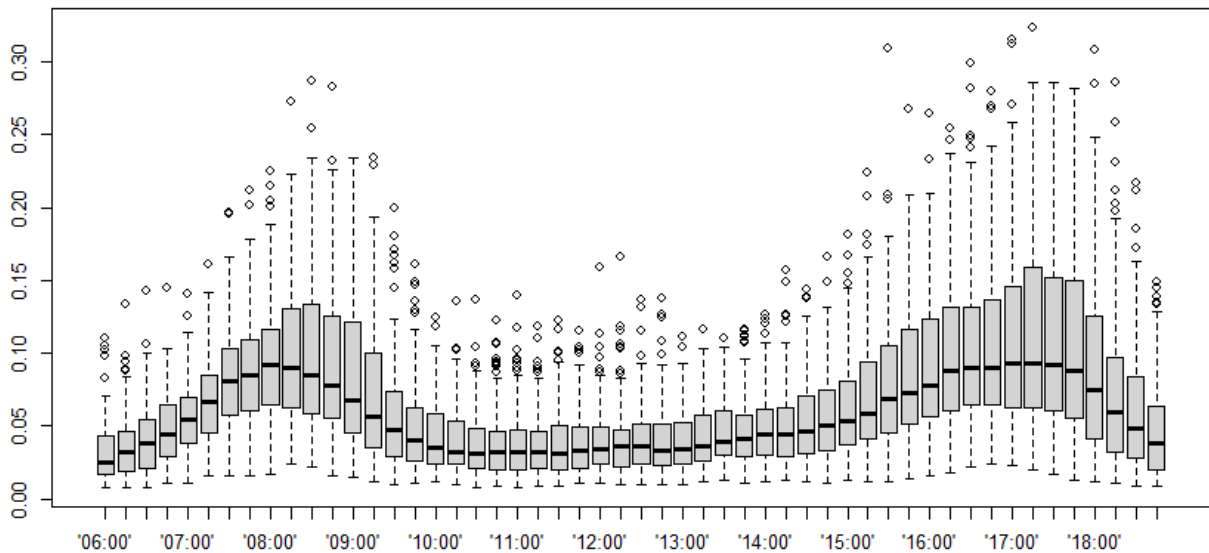
Onbetrouwbaarheid

Voor LMS-BT is het ons uiteindelijk te doen om de onbetrouwbaarheid van reistijden. Deze is bepaald als de standaardafwijking van de verschillen tussen de werkelijke reistijd per werkdag, kwartier en route en de verwachte reistijd voor die dag, dat tijdvak en die route⁷. Figuur 10 geeft aan hoe deze zogeheten reguliere onbetrouwbaarheid varieert over de dag voor alle routes en werkdagen. Per kwartier worden de resultaten uit alle route-alternatieven getoond. Ook hier komt een herkenbaar beeld naar voren. In beide spitsperiodes is de onbetrouwbaarheid gemiddeld groter en vertoont deze een grotere spreiding. De variatie in reguliere onbetrouwbaarheid is nog geprononceerder als gecorrigeerd wordt voor het verschil in routelengte (Figuur 11). In deze weergave komt wederom het asymmetrische karakter van de spits tot uiting.



Figuur 10 Reguliere onbetrouwbaarheid per vertrekkwartier voor alle routes en werkdagen uitgedrukt in minuten.

⁷ Wederom na uitsluiting van extreme waarden, dus conform vergelijking 5 in Kouwenhoven en Bel (2014).

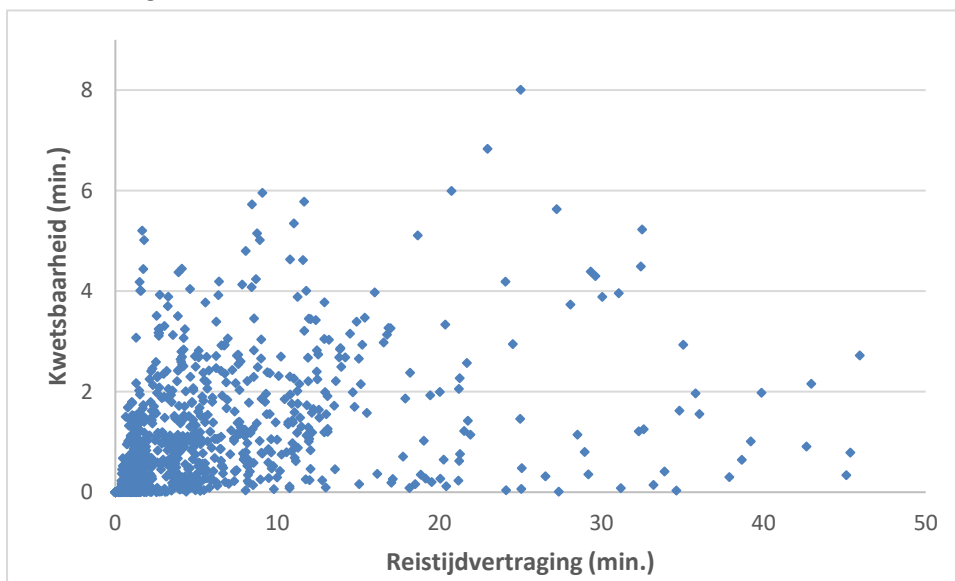


Figuur 11 Reguliere onbetrouwbaarheid per vertrekkwartier voor alle routes en werkdagen uitgedrukt in minuten per afgelegde kilometer.

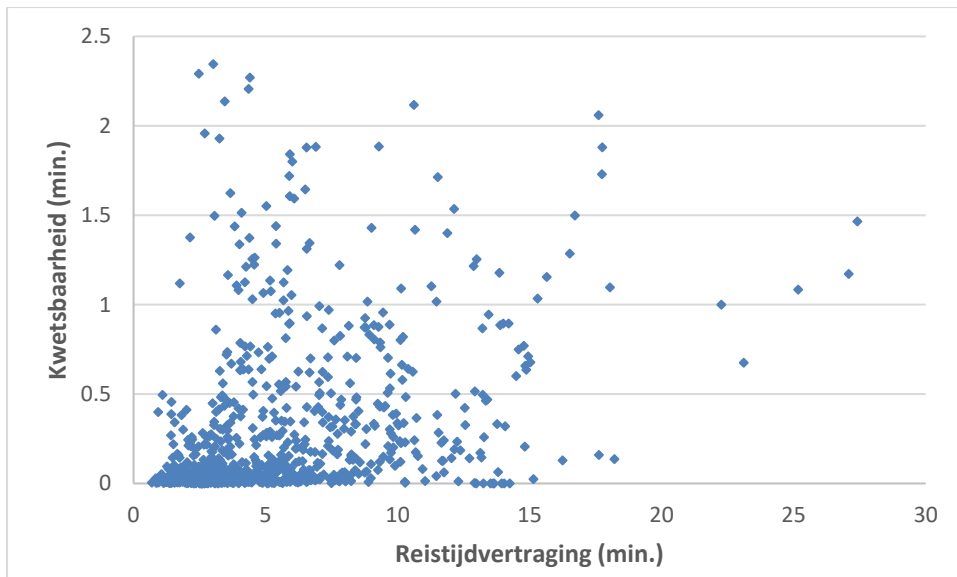
Alternatieve indicatoren van reistijd onbetrouwbaarheid

In andere studies worden alternatieve indicatoren gebruikt om reistijd onbetrouwbaarheid uit te drukken. Deze paragraaf gaat kort in op deze alternatieven en geeft een indruk hoe de in LMS-BT gehanteerde maatstaf voor onbetrouwbaarheid zich hiertoe verhoudt.

Kwetsbaarheid drukt uit hoe gevoelig reistijd is voor extreme gebeurtenissen. Kouwenhoven en Bel (2014) definiëren dit als het verschil tussen de totale onbetrouwbaarheid (de standaardafwijking van de hele reistijdverdeling) en de reguliere onbetrouwbaarheid (de standaardafwijking van de reistijdverdeling zonder extreme gebeurtenissen). In navolging van hun rapportage (pagina's xvi en xvii) vergelijken we de mate van kwetsbaarheid met reguliere reistijdvertraging (inclusief extreme gebeurtenissen). Figuur 12 toont de kwetsbaarheid op het hoofdwegennet tijdens de ochtendspits in relatie tot reistijdvertraging. Kwetsbaarheid loopt op tot 9 minuten, zonder een heel duidelijke relatie met vertraging te hebben. Zowel de waarden als het beeld in deze figuur stemmen overeen met de studie van Kouwenhoven en Bel. Figuur 13 toont dezelfde relatie voor het onderliggend wegennet. De kwetsbaarheidswaarden liggen hier, net als in de voorgaande studie, veel lager dan bij het hoofdwegennet.



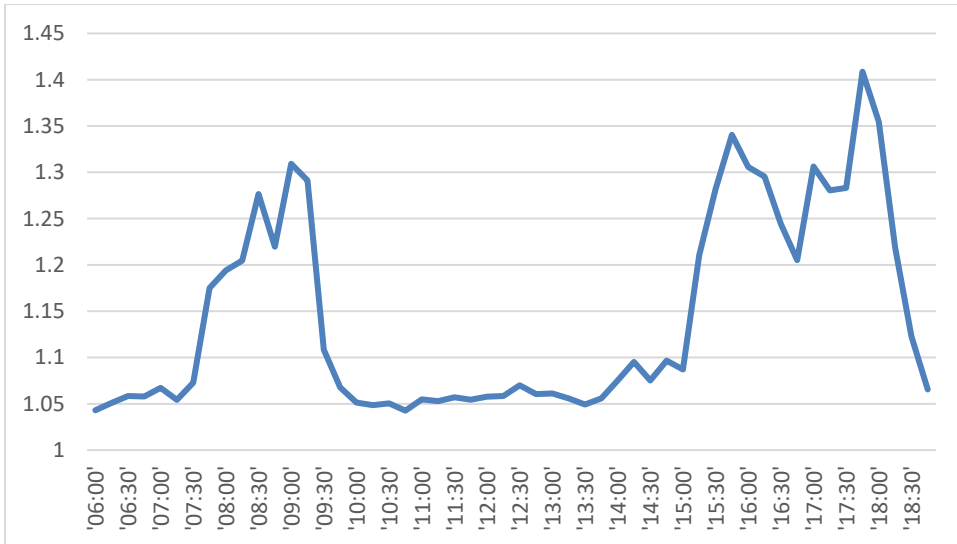
Figuur 12 Kwetsbaarheid als functie van de reistijdvertraging op het hoofdwegennet tijdens de ochtendspits.



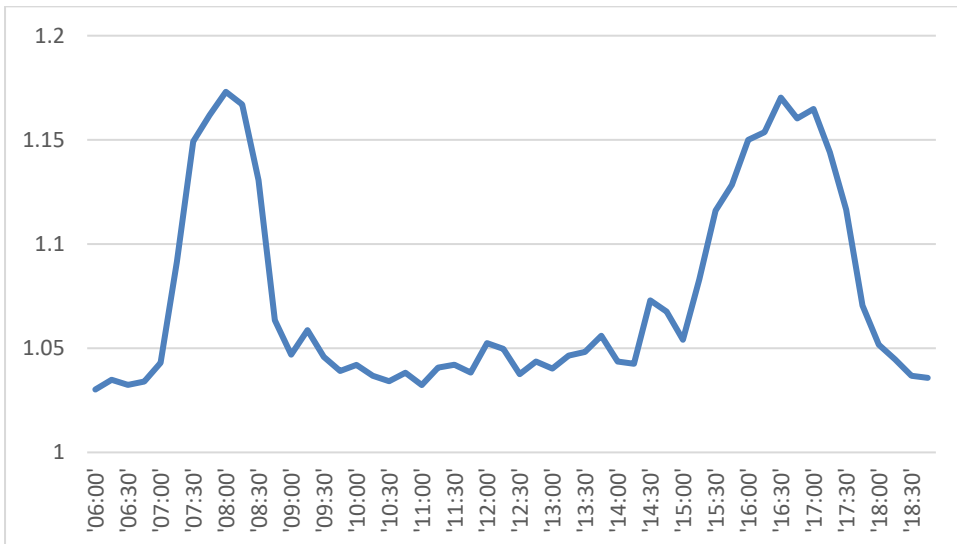
Figuur 13 Kwetsbaarheid als functie van de reistijdvertraging op overige wegen tijdens de ochtendspits.

De planningstijdindex (PTI) is ontwikkeld door LYNXX en NDW en beschrijft hoeveel tijd een reiziger moet inplannen ten opzichte van de gebruikelijke reistijd, om met een kans van 90% op tijd aan te komen op de eindbestemming (Turpijn en Visser, 2020). Deze alternatieve beschrijving van reistijd betrouwbaarheid wordt berekend door het 90^e percentiel van de reistijdverdeling te delen door de mediaan van die verdeling. Een waarde van 1,4 betekent dus dat een reiziger 40% extra reistijd moet rekenen om 90% zekerheid te hebben om op tijd aan te komen.

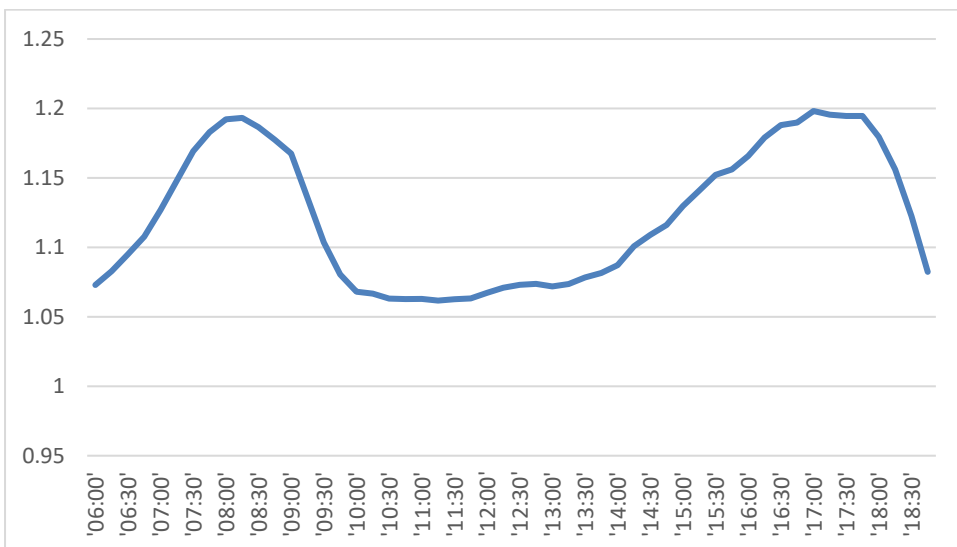
Op basis van de reistijdendatabase die we voor deze studie hebben samengesteld is het mogelijk de PTI te berekenen op het niveau van individuele routes en als gemiddelde over alle routes. Figuur 14 en 15 tonen het verloop van de PTI per vertrekkwartier tussen 6:00 's ochtends en 19:00 uur 's avonds op twee routes van uiteenlopende lengte. In beide gevallen blijkt duidelijk de sterke toename tijdens de ochtend- en avondspits. Op de korte route is dit beeld wat geprononcerder met maximale waarden van ongeveer 1,4, tegenover iets onder de 1,2 op de langere route. Het gemiddelde beeld over routes over alle geanalyseerde routes in Figuur 16 toont dit verloop op een geleidelijker manier. Het patroon lijkt sterk op het gemiddelde verloop van onbetrouwbaarheid over de dag (Figuur 10). De planningstijdindex is heel behulpzaam om de variatie in verwachte reistijd over de dag te illustreren, maar is niet direct toepasbaar in het LMS-BT instrument dat ingericht is op de indicator onbetrouwbaarheid zoals we die in de vorige sectie introduceerden. Deze indicator hanteren we in het vervolg van deze rapportage.



Figuur 14 De planningstijdindex voor route 14 (Banne Buiksloot, Amsterdam naar Nieuw West, Zaanstad, lengte 12.2 km).



Figuur 15 De planningstijdindex voor route 20 (Heemstede-Centrum naar Merenwijkdistrict, Leiden, lengte 33.9 km).



Figuur 16 De planningstijdindex gemiddeld over alle afzonderlijke routes.

5.2. Herhalen verklarende analyse huidige rekenmodule LMS-BT

Bij het opstellen van de data voor de verklarende analyse is dezelfde methode gevolgd als in Kouwenhoven en Bel (2014). In onze analyse zijn bijna 7 miljoen individuele tijdmetingen teruggebracht tot 15 observaties per waargenomen kwartier per route. De resulterende reistijdendatabase bestaat uit 120 routes * 13 uur * 4 kwartieren * 15 minuten = 93.600 waarnemingen. Hiervoor zijn alle waarnemingen van de werkdagen per route en individuele minuut gemiddeld. Deze databerg wordt in eerste instantie teruggebracht tot 6.240 kwartiertijdvakken waarvoor we de onbetrouwbaarheid (standaarddeviatie) van de reistijd bepalen ten opzichte van de binnen dat kwartier gemiddelde reistijd op basis van de individuele minuten in dat tijdvak.

In LMS-BT wordt nu een relatie gemodelleerd tussen twee variabelen, namelijk het effect van reguliere (verwachte) vertraging op de reguliere (verwachte) onbetrouwbaarheid. Met behulp van de gegenereerde data over onbetrouwbaarheid hebben we allereerst de oorspronkelijke LMS-BT schattingen herhaald. Deze hebben in basis de volgende vorm:

$$O_{r,t} = \beta_0 + \beta_1 V_{r,t} + \beta_2 \text{Log}(V_{r,t} + 1) + \beta_3 (V_{r,t} + 1) + L_r + \varepsilon_{r,t} \quad (2)$$

Waarbij in Kouwenhoven en Bel (2014) de reguliere onbetrouwbaarheid in minuten (O) van route r en tijdvak t , wordt verklaard uit reguliere vertraging (V) op die route en in dat tijdvak, de logaritme van diezelfde reguliere vertraging en de lengte (L) van de route. De specificatie wordt vervolgens toegepast op drie tijdvakken (ochtendspits, avondspits en rest van de dag)⁸ en twee typen wegen (autosnelwegen en overige wegen). In navolging van LMS-BT hanteren we daarvoor deze definities:

- Ochtendspits: vertrek vanaf 7:00 tot 9:00 uur;
- Avondspits: vertrek vanaf 16:00 tot 18:00 uur;
- Restdag: vertrek vanaf 10:00 tot 15:00 uur.

Op basis van de twee wegentypen en drie tijdvakken onderscheiden we zes aparte modellen. Tabel 7 geeft exact dezelfde specificatie als opgenomen in LMS-BT op basis van de studie van Kouwenhoven en Bel (Tabel 8). In enkele van hun modellen week de constante niet significant af van nul en zijn de schattingen uitgevoerd met de constante vastgezet op de waarde nul. Bij ons is de constante in de meeste gevallen wel significant afwijkend van nul en past de specificatie in Tabel 9 beter bij de huidige data. Voor het overig wegennet vinden we in onze analyse voor enkele tijdvakken significante relaties met $^{10}\log(\text{vertraging})$ en route lengte die niet aanwezig waren in de eerdere studie. Dat we meer significante relaties vinden is waarschijnlijk te danken aan het feit dat wij meer routes hebben voor dit type wegen. Maar ook in onze studie vinden we minder significante relaties voor het overige wegennet dan voor de hoofdwegen. Over het algemeen komen de nieuwe resultaten sterk overeen met de eerdere analyses uit 2014. **De orde van grootte van de coëfficiënten en verklaringskracht (R^2) is zeer vergelijkbaar. Een verschil met de eerdere studie is dat wij een minder sterke respons van onbetrouwbaarheid op vertraging vinden.** De coëfficiënten in onze studie zijn voor de meeste tijdvakken ongeveer een factor drie lager. In de volgende paragraaf gaan we nadere in op mogelijke oorzaken voor dat verschil.

⁸ De overige vier uren van de dag waarvoor we wel reistijdinformatie hebben (vanaf 6:00 tot 7:00 uur, van 9:00 tot 10:00 uur, van 15:00 tot 16:00 uur en van 18:00 tot 19:00), betrekken we in verdiepende analyses verderop in dit rapport.

Tabel 7 Schattingsresultaten analyse reistijdonderhoudbaarheid in minuten op basis van floating car data uit 2019, met aparte modellen voor hoofdwegenet en overig wegennet en voor ochtendspits (OS), avondspits (AS) en rest dag (RD). Omwille van de vergelijkbaarheid is deze specificatie exact dezelfde als in de studie van Kouwenhoven en Bel (2014), inclusief het op 0 zetten van de constante, wat in onze dataset niet gepast is.

	Hoofdwegenet (HWN)			Overig wegennet (OWN)		
	OS	RD	AS	OS	RD	AS
Reguliere vertraging HWN of OWN	0,146** (0,008)	0,303** (0,007)	0,059** (0,005)	0,208** (0,002)	0,118** (0,002)	0,186** (0,002)
¹⁰ log(reg.vertraging HWN/OWN + 1)	2,051** (0,112)		2,529** (0,095)			
Lengte HWN	0,015** (0,001)	0,007** (0,001)	0,022** (0,001)			
Constante	-0,301** (0,049)	0	-0,448** (0,050)	0	0,187** (0,010)	0
Waarnemingen	960	2.400	960	960	2.400	960
Adjusted R ²	0,928	0,891	0,929	0,910	0,584	0,902
Residuele standaardfout	0,613	0,519	0,579	0,432	0,226	0,441
F Statistic	4.116**	9.768**	4.191**	9.669**	3.373**	8.869**

Significantie codes coëfficiënten: * p < 0,1 ** p < 0,01.

De R² voor regressie met op 0 geforceerde constante is niet vergelijkbaar met R² voor reguliere regressie.

Tabel 8 Schattingsresultaten Kouwenhoven en Bel (2014; tabellen 5.2 en 5.5) voor autosnelwegen en overige wegen voor ochtendspits (OS), avondspits (AS) en rest dag (RD).

	Autosnelwegen (ASW)			Overige wegen (OW)		
	OS	RD	AS	OS	RD	AS
Reistijdvertraging ASW of OW	0,476** (0,026)	0,764** (0,030)	0,268** (0,017)	0,499** (0,052)	0,346** (0,030)	0,626** (0,042)
¹⁰ log(reistijdvertraging +1)	4,538** (0,415)	-	5,555** (0,351)			
Lengte	-0,009** (0,003)	-0,010** (0,003)	0,011** (0,003)			
Constante	-0,540** (0,186)	0	-0,901** (0,172)	0	-0,153** (0,076)	0
Adjusted R ²	0,956	0,926	0,960	0,688	0,777	0,859

Significantie codes coëfficiënten: * p < 0,1 ** p < 0,01.

NB In de specificaties waar de waarde van de constante niet significant afweek van 0, is de schatting herhaald met de constante vastgezet op de waarde 0. De R² voor dergelijke regressies is niet vergelijkbaar met R² voor reguliere regressie.

Tabel 9 Schattingsresultaten analyse reistijdonderbetrouwbaarheid in minuten op basis van floating car data uit 2019, apart voor hoofdwegennet en overig wegennet en voor ochtendspits (OS), avondspits (AS) en rest dag (RD). In deze specificatie is de constante alleen gelijkgesteld aan 0 voor de variant waar deze niet significant van 0 afwijkt.

	Hoofdwegennet (HWN)			Overig wegennet (OWN)		
	OS	RD	AS	OS	RD	AS
Reguliere vertraging HWN of OWN	0,146** (0,008)	0,212** (0,007)	0,059** (0,005)	0,199** (0,004)	0,071** (0,004)	0,102** (0,012)
¹⁰ log(reg.vertraging HWN/OWN + 1)	2,051** (0,112)	1,898** (0,087)	2,529** (0,095)		0,497** (0,027)	0,900** (0,211)
Lengte HWN	0,015** (0,001)	-0,001* (0,001)	0,022** (0,001)		0,003** (0,001)	
Constante	-0,301** (0,049)	-0,270** (0,023)	-0,448** (0,050)	0,072** (0,026)		-0,135** (0,103)
Waarnemingen	960	2.400	960	960	2.400	960
Adjusted R ²	0,928	0,864	0,929	0,727	0,916	0,630
Residuele standaardfout	0,613	0,460	0,579	0,430	0,224	0,414
F Statistic	4.116**	5.068**	4.191**	2.557**	8.727**	817**

Significantie codes coëfficiënten: * p < 0,1** p < 0,01.

NB In de specificaties waar de waarde van de constante niet significant afweek van 0, is de schatting herhaald met de constante vastgezet op de waarde 0. De R² voor dergelijke regressies is niet vergelijkbaar met R² voor reguliere regressie.

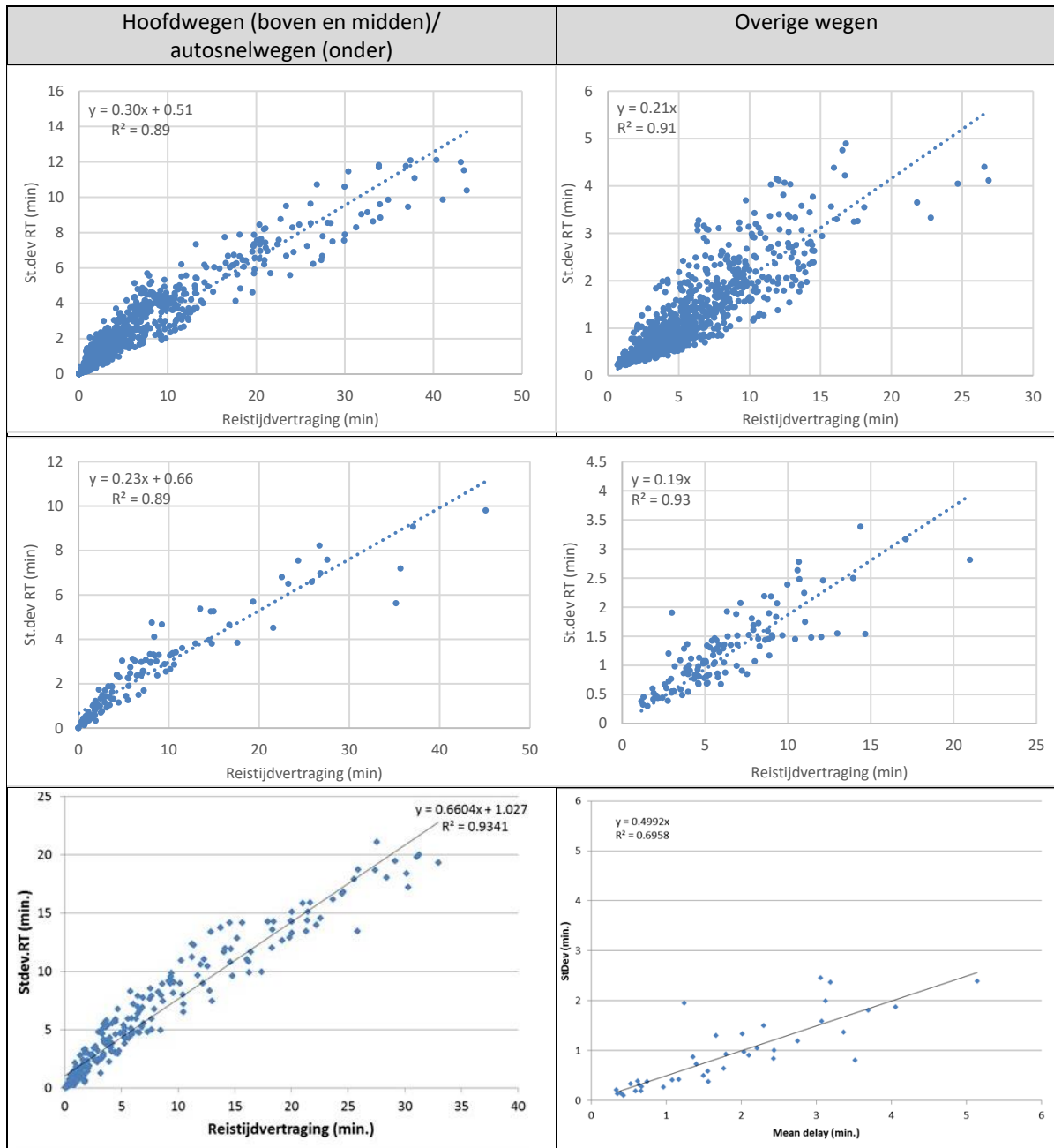
5.3. Mogelijke oorzaken voor verschil in uitkomsten ten opzichte van de eerdere analyse

Zoals we in de inleiding van dit hoofdstuk al aangaven zijn er enkele verschillen tussen de huidige studie en die uit 2014. Allereerst verschilt natuurlijk de bron van de data (floating car data tegenover tellus of kentekenregistratiedata). Maar ook de lengte en locatie van de gekozen routes en de onderzoeksperiode wijken af. Voor verschil in lengte trachten we te corrigeren door dit als verklarende variabele op te nemen. Verschillen die samenhangen met een ander verkeersbeeld op de geselecteerde routes (bijvoorbeeld in intensiteit of spreiding over de dag) zullen grotendeels opgevangen worden door het gebruik van reistijdvertraging als verklarende variabele. Deze variabele helpt ook in het opvangen een eventueel effect van toenemende verkeersintensiteiten tussen beide perioden. Verschillen in rijgedrag tussen de onderzoeksperioden (bijvoorbeeld op basis van betere, real-time reisinformatie, of toenemend gebruik van hulpmiddelen zoals cruise-control) zijn lastiger te onderzoeken en vragen een eigen onderzoeksopzet. In deze paragraaf gaan we dieper in op verschillen die samenhangen met de basisdata waarop beide studies gebaseerd zijn. Wij kijken allereerst naar de gevonden waarden voor onbetrouwbaarheid en vertraging in beide studies en de manier waarop deze basisdata verwerkt is.

Figuur 17 vergelijkt de verdeling van de waargenomen onbetrouwbaarheid en reistijdvertraging in deze studie (bovenste twee rijen, nadere uitleg in volgende paragraaf) met zijn voorganger (onderste rij). De figuur visualiseert wat ook al uit de regressieanalyse bleek: de gevonden relaties zijn zeer vergelijkbaar: de onbetrouwbaarheid neemt gelijkmatig toe bij oplopende reistijdvertraging. Die relatie vinden we zowel op hoofdwegen als onderliggende wegen. Met name voor laatstgenoemde wegtypen is dat een belangrijke constatering, want in de eerdere studie was dit gebaseerd op een relatief klein aantal trajecten. In de huidige studie kunnen we de gevonden relaties bevestigen op basis van een veel groter aantal trajectendelen over overige wegen. Dat grotere aantal waarnemingen vertaalt zich ook in een grotere verklaringskracht.

Duidelijk is verder dat de range aan gevonden onbetrouwbaarheidswaarden voor de hoofdwegen in deze studie beperkter is dan in de vorige studie. De range in vertragingwaarden is ongeveer gelijk, met als resultaat dat bij gelijke vertraging deze studie minder onbetrouwbaarheid vindt. Voor de overige wegen vinden we in deze studie een iets grotere range aan onbetrouwbaarheidswaarden bij

een beduidend grotere range aan reistijdvertragingen, wat wederom leidt tot een geringere respons van onbetrouwbaarheid op vertraging. De verschillen in waarde ranges zullen deels samenhangen met de verschillen in lengte van de gekozen trajecten. Onze trajectdelen over de hoofdwegen zijn gemiddeld genomen korter, terwijl de delen over de overige wegen juist langer zijn. Te verwachten is dat langere routes in absolute zin ook langere gemiddelde vertragingen kennen.



Figuur 17 Waardeverdeling reguliere onbetrouwbaarheden (op y-as) tegenover reguliere vertraging (op x-as), gemiddeld per werkdagkwartier per route op basis van floating car data (boven; N = 960), gemiddeld per route op basis van floating car data (midden; N= 120), en gemiddeld per route volgens tellus (ASW) of camera (OWN) metingen (onder, uit Kouwenhoven en Bel, 2014; N = 250 voor autosnelwegen en N= 40 voor overige wegen). Alle relaties betreffen de ochtendspits.

Verskil in temporele resolutie

Om na te gaan wat de geringere respons van onbetrouwbaarheid op vertraging kan verklaren, hebben we allereerst onderzocht wat het effect is van verschillen in temporele resolutie van de gebruikte data. In onze studie aggregeren we de waarnemingen naar gemiddelden per werkdag-kwartier per route (dus bijvoorbeeld het gemiddelde tussen 6:00 en 6:15 over alle werkdagen voor een individueel route-alternatief), terwijl in de vorige studie slechts een gemeten waarde per kwartier per route gebruikt werd. Figuur 17 toont het verschil tussen een uitgebreide set aan waarnemingen die per werkdagkwartier per route een waarde voor onbetrouwbaarheid en vertraging geeft (bovenste twee grafieken) en een beperkter set waarin deze waarden per route gemiddeld zijn (middelste grafieken) zoals dat ook in de eerdere studie van Kouwenhoven en Bel het geval was (onderste grafieken). Deze extra aggregatiestap reduceert de waargenomen variantie, maar maakt geen verschil voor de gevonden relaties tussen onbetrouwbaarheid en vertraging: de verklaringskracht neemt iets toe, de coëfficiënten blijven vrijwel gelijk.

Een ander verschil in temporele resolutie is dat wij een trajectorie meting per minuut vertrektijd gebruiken om een kwartiergemiddelde waarde per individuele dag, tijdvak en route te bepalen, terwijl de oude studie slechts één trajectorie meting per kwartier beschikbaar had. Om het mogelijke effect van dit verschil te beschrijven, hebben wij de aggregatie naar kwartiergemiddelden herhaald op basis van één sample per kwartier in plaats van alle 15 minuten in een kwartier. Het oorspronkelijke model (uit paragraaf 5.2) is vervolgens opnieuw geschat voor de set met alle waarnemingen (het 15/15 sample) en de set met één waarneming per kwartier (het 1/15 sample). Bijlage 2 geeft een grafische weergave van deze verschillende selectiemethoden. De resultaten van de hierop gebaseerde regressie-analyses zijn te zien in Tabel 10. Hieruit blijkt dat de regressie uitkomsten per periode bijzonder sterk op elkaar lijken. **Een minder frequente sampling van trajectorie metingen maakt dus geen verschil voor de gevonden relaties tussen onbetrouwbaarheid en vertraging.**

Tabel 10. Regressie-analyses van onbetrouwbaarheid op hoofdwegennet op basis van 15 samples per kwartiertijdvak (15/15) of 1 sample per kwartiertijdvak (1/15), voor ochtendspits (OS), avondspits (AS) en rest dag (RD). Vanaf deze analyses hanteren we het gangbaarder natuurlijk logaritme (ln), in plaats van de ¹⁰log.

Periode Sampling	Onbetrouwbaarheid op hoofdwegen					
	OS 15/15	RD 15/15	AS 15/15	OS 1/15	RD 1/15	AS 1/15
Reguliere vertraging HWN	0,146** (0,008)	0,212** (0,007)	0,059** (0,005)	0,154** (0,008)	0,215** (0,008)	0,062** (0,005)
ln(reg.vertraging HWN + 1)	0,891** (0,049)	0,824** (0,038)	1,098** (0,041)	0,846** (0,049)	0,815** (0,040)	1,066** (0,041)
Lengte HWN	0,015** (0,001)	-0,001** (0,001)	0,022** (0,001)	0,015** (0,001)	-0,001* (0,001)	0,022** (0,001)
Constante	-0,301** (0,049)	-0,270** (0,023)	-0,448** (0,050)	-0,315** (0,049)	-0,282** (0,024)	-0,457** (0,050)
Waarnemingen	960	2.400	960	960	2.400	960
Adjusted R ²	0,928	0,864	0,929	0,930	0,849	0,930
Residuele standaardfout	0,613	0,460	0,579	0,621	0,475	0,587
F Statistic	4.116**	5.069**	4.191**	4.228**	4.488**	4.224**

Significantie codes coëfficiënten: * p < 0,1** p < 0,01.

Ontbrekende waarnemingen

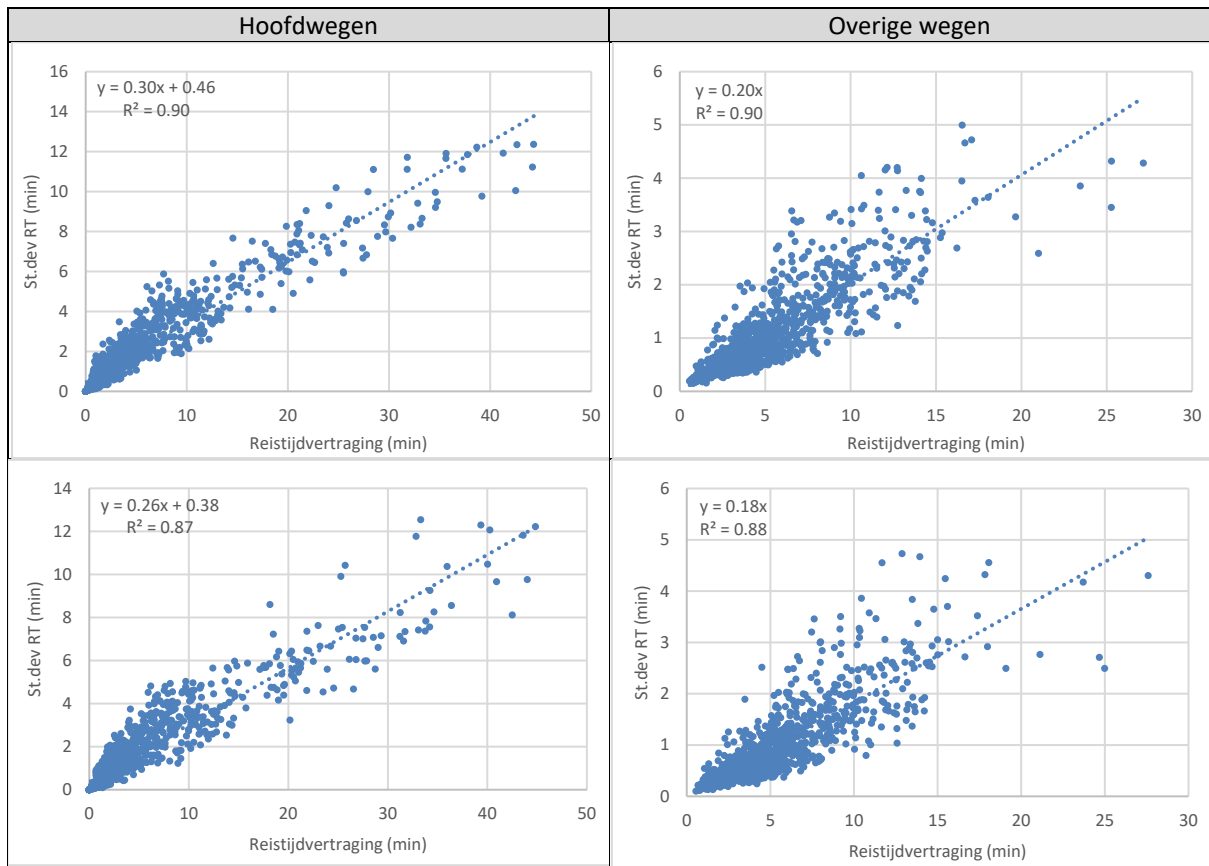
De floating car data heeft niet voor elk wegvak voor elke minuut een waarneming. Dat komt deels doordat er niet op elk moment van de dag op elk wegvak verkeer is, maar is ook gevolg van het feit dat de floating car data afkomstig is van een selectie van het totaal aantal weggebruikers. Als er minder dan 10 observaties in een minuut beschikbaar zijn, worden observaties uit voorgaande

minuten gebruikt. Na een half uur zonder observaties wordt het segment behandeld alsof er geen observaties aanwezig zijn.

Informatie over de initiële inferentie (gebruik van voorliggende minuten) door de dataleverancier is maar beperkt voorhanden in de aan NDW geleverde data. Deze beperkte informatie is niet overgenomen in de bestanden waarmee trajectoriën worden doorgerekend om de bestandsgroottes in te perken en reistijd te optimaliseren. De mate van initiële inferentie is dus niet direct per vertrekmoment te bepalen. Om toch enig gevoel te krijgen voor het effect van de initiële inferentie is gekeken wat het effect is van het slechts beperkt beschikbaar hebben van waarnemingen. Om dit te benaderen is per half uur een reistijd geselecteerd. Daarmee bootsen we de extreme situatie na dat slechts eens per 30 minuten een snelheidswaarneming beschikbaar is. Met deze selectie zijn we er tevens zeker van dat de per vertrekmoment gehanteerde snelheden onafhankelijk van elkaar bepaald zijn. Er worden door Be-Mobile immers tot 30 minuten teruggekeken om snelheden uit de voorliggende minuten over te nemen. De zo geselecteerde reistijden zijn dus bepaald zonder eventuele overlap in de hiervoor gebruikte snelheidswaarnemingen. We kiezen ervoor om op de oneven dagen de reistijden van de hele en halve uren te kiezen (:00 en :30) en op de even dagen de overige volle kwartieren (:15 en :45). Bijlage 2 illustreert deze selectiemethode.

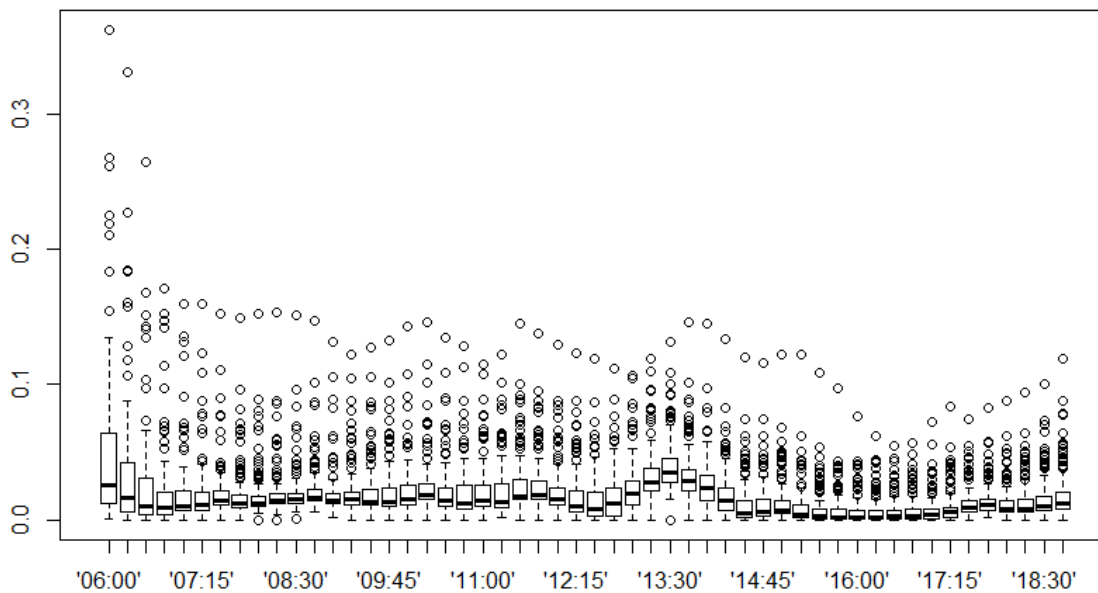
De hier beschreven selectie levert een dataset die gebaseerd is op een beperkt sample van effectief 1 uit 30 waarnemingen. Hieruit zijn vervolgens onbetrouwbaarheid en reistijdvertraging per route afgeleid voor de complete set van routealternatieven (Figuur 18). Deze analyse karakteriseert de fictieve situatie dat steeds 30 minuten dezelfde waarde gehanteerd zou worden. **Zoals op basis van het kleinere aantal waarnemingen en daarmee samenhangende beperkter variantie in uitkomsten verwacht mag worden, wordt de onbetrouwbaarheidsrespons *kleiner* bij een verdere verlaging van de sample dichtheid.** Het effect op zowel de geschatte coëfficiënten als verklaringskracht is overigens gering, zoals blijkt uit de vergelijkingen die in de figuur zijn opgenomen. Al met al suggereert dit, dat de invloed van het gebruik van waarnemingen uit voorliggende minuten in de floating car data geen belangrijke factor zal zijn in de relatief lage onbetrouwbaarheden.

Het effect van de ontbrekende snelheden (als over een half uur waarnemingen ontbreken) is te benaderen doordat wij hebben bijgehouden welk percentage van een reistijd is bepaald op basis van zelf ingevulde snelheden. Als de snelheid op een bepaalde segment-tijd combinatie niet voorhanden is, is steeds de freeflow snelheid genomen. De totale hoeveelheid tijd die middels deze inferentie is geschat, is als variabele per vertrektijd en route beschikbaar in de variabele T_{inf} . Het effect van ontbrekende snelheidsdata is geschat middels de variabele $P_{inf} = T_{inf} / T_{min}$, waarbij T_{min} die minimale tijd betreft waarbinnen de route (op basis van freeflow snelheid) afgelegd kan worden. De variabele beschrijft het deel van de route waarvoor inferentie is gedaan. In het hypothetische geval dat de reistijd voor de hele route zo geschat is heeft P_{inf} de waarde 1. In werkelijkheid wordt slechts een klein deel (gemiddeld 1,95%) van de reistijden per kwartier op deze manier ingevuld.

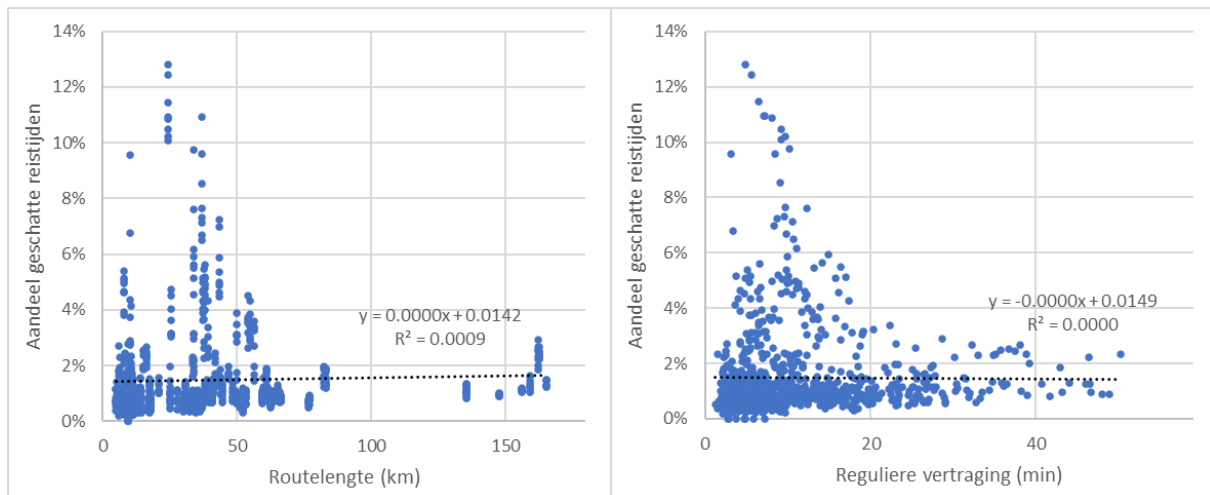


Figuur 18 Reguliere onbetrouwbaarheid en reguliere reistijdvertraging per kwartiertijdvak en route tijdens de ochtendspits (N = 960) op basis van floating car data en 1/15 sampling (boven) en 1/30 sampling (onder).

Het aandeel geschatte reistijden is hoger dan gemiddeld en kent enkele uitschieters in de vroege ochtend (Figuur 19), wat logischerwijs samenvalt met het moment dat er weinig verkeer aanwezig is. Daarnaast is er een kleine piek rond de lunchtijd. Daarnaast vinden we ook geen structureel verband tussen aandeel geschatte reistijden en routelengte of vertraging (Figuur 20), zodat deze karakteristieken als (controle) variabelen gebruikt kunnen worden.



Figuur 19 Aandeel reistijd (ten opzichte van minimale reistijd) dat geschat is wegens ontbreken van snelheidsinformatie. De aandelen zijn gemiddeld per kwartiertijdvak van vertrek over alle route-alternatieven.



Figuur 20 Relatie geschatte reistijd en route lengte (links), en reguliere vertraging (rechts) gemiddeld per kwartier per route alternatief, alleen voor de ochtendspits (N=960).

Het effect van het toevoegen van geschatte snelheidsinformatie voor tijdstippen waarop deze ontbrak is geschat in een vergelijkbaar regressiemodel als beschreven in de vorige paragraaf. De resultaten zijn te zien in Tabel 11. Daaruit blijkt dat P_{inf} in de ochtendspits en rest van de dag het verwachte effect heeft op onbetrouwbaarheid: meer inferenties leiden hier tot minder onbetrouwbaarheid. In de avondspits is het effect niet significant. Belangrijker is nog dat de relatie tussen onbetrouwbaarheid en vertraging vrijwel gelijk blijft als je corrigeert voor toevoegen van geschatte snelheden. **Daarmee lijken de eerdere resultaten niet heel sterk beïnvloed door de missende data en verklaren ze ook niet de beperkte respons van onbetrouwbaarheid op vertraging in de floating car data.**

Tabel 11. Resultaten regressie onbetrouwbaarheid (hele traject) versus vertraging (hele traject), inclusief effect van inferentie van snelheden.

	Onbetrouwbaarheid op hele traject		
	Ochtendspits	Rest dag	Avondspits
Reguliere vertraging	0,163** (0,009)	0,245** (0,008)	0,079** (0,007)
ln(reg.vertraging + 1)	0,339** (0,101)	-0,702** (0,067)	0,944** (0,102)
Lengte	0,017** (0,001)	0,010** (0,001)	0,020** (0,001)
P_{inf}	-1,013 (1,045)	-1,435* (0,523)	-2,215 (1,820)
Constante	-0,481** (0,157)	0,686** (0,081)	-1,221** (0,184)
Waarnemingen	960	2.400	960
Adjusted R ²	0,902	0,799	0,883
Residuele standaardfout	0,695	0,539	0,688
F Statistic	2.214**	2.385**	1.802**

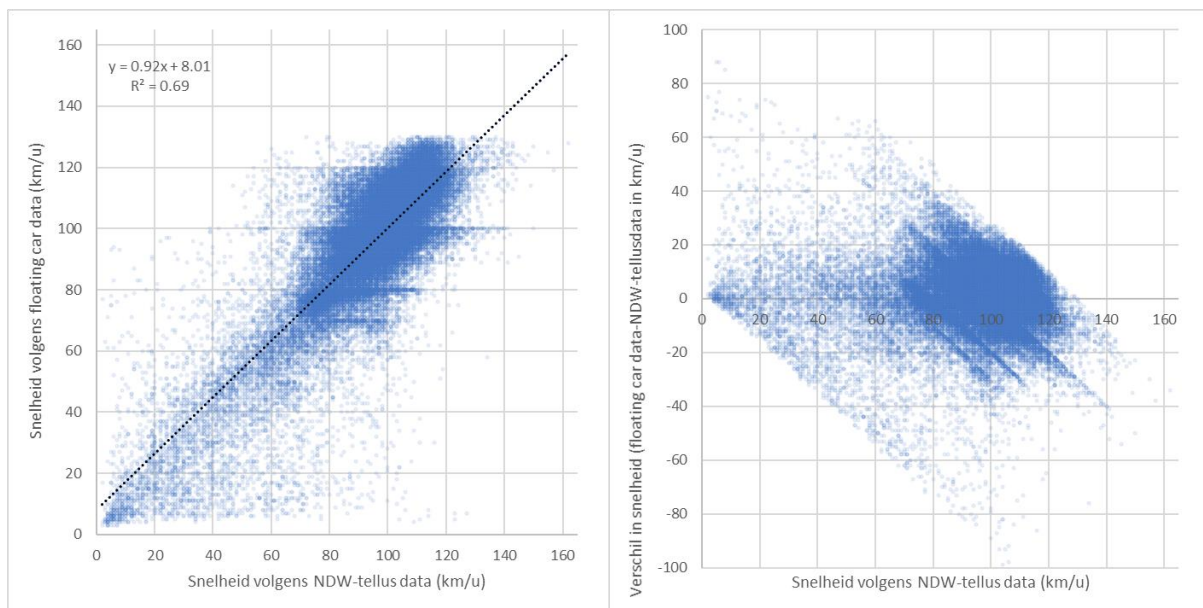
Significantie codes coëfficiënten: * $p < 0,1$ ** $p < 0,01$.

Verskil in gerapporteerde snelheden

De snelheden in de floating car data kunnen om verschillende redenen afwijken van de snelheden uit andere informatiebronnen. Allereerst heeft de leverancier, voorafgaand aan uitlevering van deze gegevens, alle waargenomen snelheden boven de maximumsnelheid afgekapt tot de maximumwaarde. Daarnaast zijn de floating car data afkomstig van een selectievere set aan voertuigen dan bijvoorbeeld de tellusdata uit de vorige studie die in basis alle voertuigen bevat. Het is aannemelijk dat floating car data chauffeurs (onder meer gebruikers van de Flitsmeister app en met internet verbonden auto's, zie Knoope, 2020) waar mogelijk iets harder rijden dan de gemiddelde weggebruiker. Zo registreren de tellussen ook vrachtverkeer dat de gemiddelde snelheid zal verlagen op wegvakken met een hoge maximumsnelheid. Om dit verschil in beeld te brengen hebben we voor

een selectie aan overeenkomstige locaties de gerapporteerde floating car data snelheden vergeleken met snelheden uit de tellusdata van het NDW.

Figuur 21 vergelijkt de gemiddelde snelheid per kwartiertijdvak voor 14 werkdagen (binnen de periode oktober-november 2019), 10 kwartiervakken (van 7:00 tot en met 8:00 en van 12:00 tot en met 13:00) op ruim 800 locaties die zowel in de floating car data als tellusdata van NDW voorkomen. In totaal betreft dit ruim 109.000 datapunten⁹ die betrekking hebben op de studieperiode en locaties die we ook in deze studie onderzoeken. De snelheden in de floating car data zijn afgekapt op de maximumsnelheid die per wegvak gold. Dat is te zien in de figuur aan het ontbreken van waarden boven 130 km/uur, maar bijvoorbeeld ook aan de concentraties op bijvoorbeeld 100 km/u en 80 km/u. Voor het overgrote deel komen de snelheden overeen, al zijn de niet afgekapte waarden in de floating car data gemiddeld genomen iets hoger (+0,58 km/u) dan uit de tellusdata naar voren komt.



Figuur 21 Vergelijk gemiddelde snelheid per kwartier op basis van floating car data en NDW-tellusdata voor een selectie van overeenkomstige locaties.

Die verschillen in waargenomen snelheden zijn niet constant over het snelheidsbereik, zoals blijkt uit Tabel 12. Het beeld dat uit deze tabel en Figuur 21 naar voren komt, lijkt sterk op de vergelijkbare analyse die Willigers en Kouwenhoven (2020) uitvoerden. Het aandeel snelheidswaarnemingen met een verschil kleiner of gelijk aan 5 km/u ligt in ons vergelijk met 48% lager (ten opzichte van 73% in de vorige studie). Dat kan een gevolg zijn van de striktere selectie van meetpunten in de vorige studie, waar alle waarnemingen werden uitgesloten van meetpunten die over de onderzoeksperiode een gemiddeld verschil van meer dan 10 km/u vertoonden. Dat hebben wij in onze korte verkenning van verschillen niet gedaan.

Het is niet direct evident wat de verschillen in gerapporteerde snelheden tussen de floating car data en de tellusdata betekenen voor de door ons bepaalde onbetrouwbaarheid en vertraging. Afkappen van in werkelijkheid hogere snelheden leidt mogelijk tot onderschatting van onbetrouwbaarheid, want de variantie naar boven toe is artificieel verwijderd. Het leidt mogelijk ook tot overschatting van vertraging, omdat snelheidsuitschieters boven de maximumsnelheid niet meetellen voor de schatting van gemiddelden. Aan de andere kant tellen ze ook niet mee voor de bepaling van de

⁹ In de figuur zijn bijna 29.000 datapunten (26% van het totaal) weggelaten waarvoor geen snelheid beschikbaar was in de floating car data. De missende snelheden zijn in verdere analyses vervangen door de freeflow snelheid. De geselecteerde NDW telluslocaties zijn in beheer bij Rijkswaterstaat en betreffen locaties op het hoofdwegenet.

freeflow reistijden. Deze reistijden, die als referentie gelden voor de te bepalen vertraging, zullen op basis van de afgekapte floating car data mogelijk iets langer zijn en daarmee tot iets lagere waarden voor reistijdvertraging leiden. Dat laatst roept overigens een meer fundamentele vraag op: in hoeverre is het gewenst vertraging te bepalen ten opzichte van niet toegestane snelheden. Deze vraag is voor vervolgonderzoek helemaal relevant omdat de maximumsnelheid sinds maart 2020 overdag nog slechts 100 km/uur is.

Tabel 12. Vergelijk snelheden floating car data (FCD) en een selectie van ruim 800 door RWS beheerde tellusdata van (verkregen via NDW).

Snelheidsbereik (NDW)	Aantal waarnemingen	Gemiddeld verschil (FCD-NDW)
0-20 km/u	704	7,7
20-40 km/u	1.231	1,2
40-60 km/u	1.886	-0,7
60-80 km/u	6.180	1,6
80-100 km/u	34.233	2,1
100-120 km/u	34.728	-0,7
120-165 km/u	1.514	-11,6
Alle waarnemingen	80.476	0,59

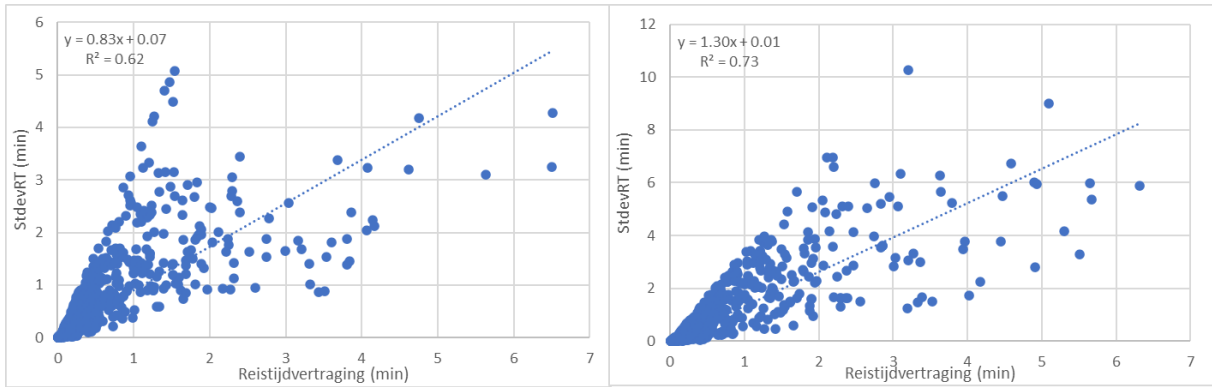
Ruimtelijke resolutie

Als laatste mogelijke verklaring voor de geringere respons van onbetrouwbaarheid op vertraging hebben we proberen te benaderen wat het effect van de veel beperkter ruimtelijke resolutie van de tellusdata is. Een belangrijk verschil tussen beide datasets is namelijk dat de floating car data waarnemingen heeft voor elk afzonderlijk wegvak van gemiddeld 40 meter lengte, terwijl bijvoorbeeld de tellusdata beschikbaar is voor veel minder locaties. Dit betekent dat in het geval van tellusdata moet worden aangenomen dat de snelheden gemeten op een lus representatief zijn voor een veel langere wegsectie dan bij gebruik van floating car data.

Om dit verschil tussen beide datasets te reproduceren, hebben we ruimtelijke selecties gemaakt van de datapunten waarop we reistijdvertraging en onbetrouwbaarheid bepalen. Omwille van de rekentijd doen we dit voor een beperkte set van telluslocaties die overlappen met tien in dit rapport gegenereerde routes in het floating car data netwerk. Op de routenetwerken is vervolgens steeds bepaald over welk deel van de route de telluslocatie een indicatie van de snelheid zou kunnen geven. Dit is steeds het routedeel tot de volgende telluslocatie op de route. Vervolgens zijn voor elke route alleen telluslocaties geselecteerd met een wegvak van minimaal 1 en maximaal 5 kilometer lengte; de laatste tellus op een route is steeds buiten beschouwing gelaten. In totaal zijn zo 188 telluslocaties geselecteerd met corresponderende wegvakken.

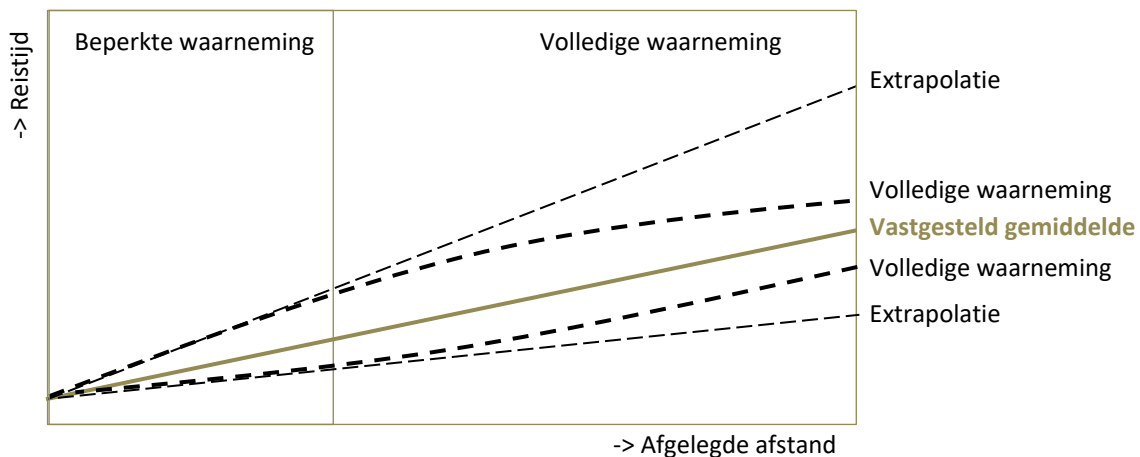
Voor deze wegvakken is de reistijd bepaald op basis van de trajectoriën die zijn gemaakt voor deze studie met floating car data. De reistijd is ook geschat op basis van de floating car data trajectorie snelheden, geobserveerd op alleen de eerste vijf segmenten van een veel langer wegvak. Vervolgens zijn gemiddelde reistijden en onbetrouwbaarheid berekend middels de methode Kouwenhoven en Bel voor reistijden op basis van alle waarnemingen en voor reistijden op basis van een beperkte observatie.

Figuur 22 geeft het resultaat van deze analyse weer: links de uitkomsten wanneer we alle beschikbare waarnemingslocaties gebruiken en rechts de versie wanneer we slechts de eerste vijf floating car wegvakken gebruiken als indicatie van de snelheid op een veel langere sectie.



Figuur 22 Reistijdvertraging en onbetrouwbaarheid voor een selectie van 10 route-alternatieven op basis van alle waarnemingen (links) en een beperkte sequentie van waarnemingen (rechts).

De analyse laat zien dat reistijden gebaseerd op een beperktere sequentie van observaties meer onbetrouwbaarheid laten zien dan reistijden die gebaseerd zijn op een veel groter aantal ruimtelijk nabije datapunten. Het lijkt erop dat een beperkt aantal, ruimtelijk verder gespreide waarnemingen de variantie in reistijd en daarmee hun onbetrouwbaarheid accentueert. Figuur 23 geeft een schematische weergave van het vermoede verschil, waarbij in het geval van extrapolatie van snelheden de variatie in snelheden op een langer traject waarschijnlijk wordt overschat. Naar alle waarschijnlijk is dit ook het geval bij de tellusdata die in de eerdere studie is gebruikt. Dit suggereert dat de floating car data een mogelijk betrouwbaarder inschatting geeft van onbetrouwbaarheid.



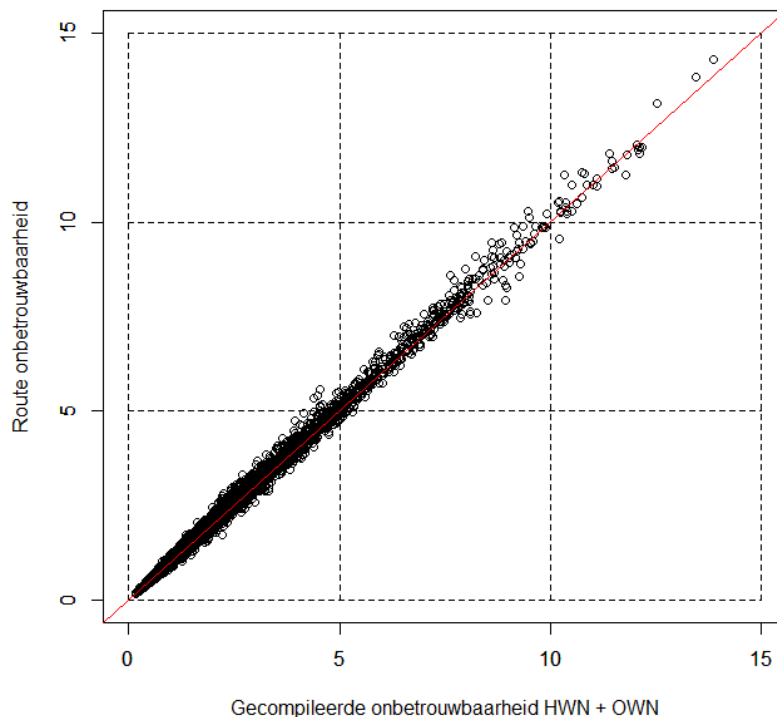
Figuur 23 Schematische weergave van de mogelijke oorzaak van verschillen in onbetrouwbaarheid op basis van snelheden geëxtrapolerd van snelheidsmetingen met een relatief lage ruimtelijke resolutie (beperkte waarnemingen) versus gemeten op een relatief hoge ruimtelijke resolutie (volledige waarnemingen).

5.4. Totale onbetrouwbaarheid per route

De oorspronkelijke LMS-BT schattingen beschouwen de onbetrouwbaarheid op autosnelwegen en overige wegen apart omdat die op andere routes en databronnen zijn gebaseerd. De separaat geschatte onbetrouwbaarheden werden vervolgens opgeteld onder de aanname dat deze niet gecorreleerd zijn (Kouwenhoven en Bel, 2014; pag.40). Die aanname is destijds onderbouwd door de correlatie tussen beiden onbetrouwbaarheden te vergelijken voor 10 trajecten waarvoor onbetrouwbaarheden geïntegreerd waren bepaald voor autosnelwegen en overige wegen. De correlatie tussen onbetrouwbaarheden op de twee wegsoorten bleek voor de beschikbare trajecten te variëren tussen -0,021 en +0,211. De onderzoekers gaven daarbij aan dat het niet mogelijk was sterke conclusies te trekken op basis van de beperkte set trajecten. Om aan de voorzichtige kant te blijven is er daarom destijds voor gekozen om de beide onbetrouwbaarheden als niet gecorreleerd

te beschouwen. Met de nu verzamelde reistijdendatabase over 120 routealternatieven is het mogelijk hier in meer detail naar te kijken.

Als eerste stap daarin kijken we in meer detail naar de relatie tussen totale (reguliere) reistijd-onbetrouwbaarheid per route en de som van afzonderlijk bepaalde onbetrouwbaarheden¹⁰ voor autosnelwegen en overige wegen. Zoals Figuur 24 laat zien, leidt het apart bepalen van deze onbetrouwbaarheden over beide weggedelen en het vervolgens optellen daarvan tot een samengestelde maat voor onbetrouwbaarheid (zoals in LMS-BT gebeurt) tot kleine onnauwkeurigheden in het bepalen van de onbetrouwbaarheid over de gehele route: voor de meeste routes is er een beperkt, maar zichtbaar verschil tussen beide benaderingen. De regressie-analyse resultaten in Tabel 13 geven diezelfde relatie weer per periode van de dag. Hier blijkt nog duidelijker dat de samengestelde maat voor totale onbetrouwbaarheid stelselmatig lagere waarden geeft dan de totale onbetrouwbaarheid zoals die over de hele route wordt bepaald. **Dit geeft aan dat de onbetrouwbaarheid op beide routedelen in enige mate gecorreleerd is en het beter is deze over de hele route te bepalen, dan uit de som der delen.**



Figuur 24 Vergelijking van absolute route-specifieke onbetrouwbaarheid en samengestelde onbetrouwbaarheid uit absolute onbetrouwbaarheden bepaald over de HWN en OVN delen afzonderlijk, voor alle routes en gehele dag.

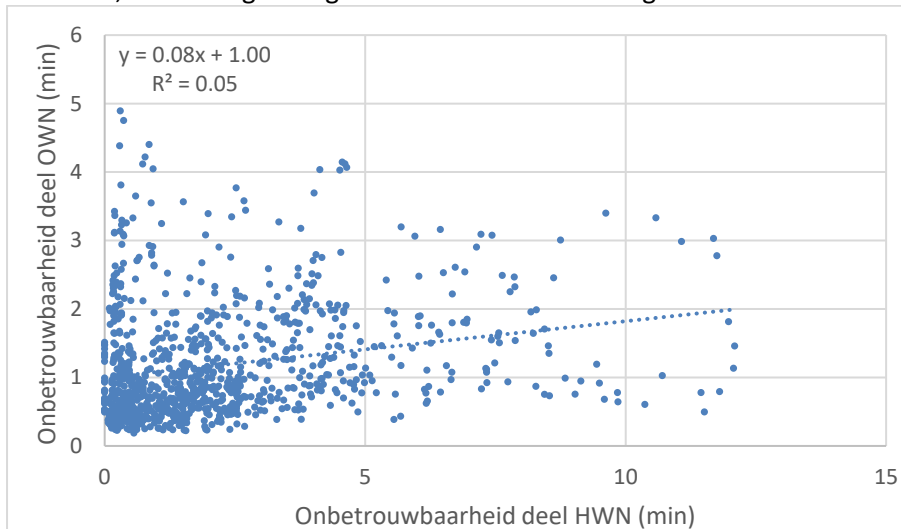
Tabel 13 Resultaten regressie verklaren totale onbetrouwbaarheid (hele traject) uit samengestelde onbetrouwbaarheid.

	Ochtendspits	Avondspits	Rest dag
Onbetrouwbaarheid (compilatie HWN-OWN)	1,031**	1,026**	1,054**
Constant	-0,002	0,002	-0,040**
Waarnemingen	960	960	2.400
Adjusted R ²	0,919	0,913	0,867

Significantie codes coëfficiënten: * $p < 0,1$ ** $p < 0,01$. Standaardfouten tussen haakjes.

¹⁰ Om deze onbetrouwbaarheden op te tellen wordt de wortel genomen uit de som van de varianties (kwadraten van de standaarddeviaties die als onbetrouwbaarheid geïnterpreteerd worden).

Figuur 25 geeft aan dat beide onbetrouwbaarheden inderdaad enigszins gecorreleerd zijn. Uit de puntenwolk blijkt de correlatie echter niet heel sterk, maar hier speelt mee dat naar absolute waarden van onbetrouwbaarheid gekeken wordt en de beide wegdelen van een route soms ongelijk in lengte zijn. Aangezien de totale duur van vertraging en dus ook onbetrouwbaarheid afhankelijk kan zijn van de lengte van de route, is deze weergave wellicht wat misleidend. In Tabel 14 kijken we daarom ook naar de onbetrouwbaarheid per afgelegde kilometer. Daaruit blijkt dat de correlatie sterker is, wanneer gecorrigeerd wordt voor routelengte.



Figuur 25 Correlatie onbetrouwbaarheid op de HWN en OWN delen van een route in de ochtendspits.

Tabel 14 Correlatie absolute onbetrouwbaarheid in minuten en relatieve onbetrouwbaarheid (in minuten per kilometer) op de HWN en OWN delen van een route.

	Absolute onbetrouwbaarheid op HWN	Onbetrouwbaarheid per km op HWN
Onbetrouwbaarheid per km op OWN		0,428** (0,014)
Absolute onbetrouwbaarheid op OWN	0,965** (0,038)	
Constante	0,687** (0,040)	0,055** (0,002)
Waarnemingen	6.240	6.084
Adjusted R ²	0,094	0,130
Residuele standaardfout	1,789	0,083
F Statistic	645**	913**

Significantie codes coëfficiënten: * p < 0,1 ** p < 0,01. Standaardfouten tussen haakjes.

6. Verdiepende analyse

In aanvulling op het herhalen van de schattingen zoals die voor de huidige rekenmodule LMS-BT zijn gedaan, is een verdiepende analyse opgezet om te verkennen hoe de huidige module verbeterd kan worden. Dit hoofdstuk doet kort verslag van de zoektocht naar een verbeterde specificatie en toont een selectie van de verschillende statistische modellen die hiervoor geschat zijn. Het voornaamste doel hiervan is te laten zien hoe elke opeenvolgende stap van invloed is op de gevonden relaties tussen vertraging en onbetrouwbaarheid.

Figuur 26 geeft een overzicht van de stappen die we gevolgd hebben om tot een verbeterde specificatie te komen. Na het herhalen van de eerdere schattingen hebben we allereerst onderzocht of het mogelijk is onbetrouwbaarheid te beschouwen over de gehele route, dus zonder deze op te splitsen in hoofdwegen en andere wegen zoals dit nu gebeurt. Aansluitend is gekeken of de schattingen gevoelig zijn voor de dominante rijrichting en is onderzocht of het zinvol is de set verklarende variabelen verder uit te breiden met extra factoren zoals die zijn beschreven in de literatuurstudie (Hoofdstuk 2). Daarna is geanalyseerd wat de toegevoegde waarde is van het verklaren van de natuurlijk logaritme van onbetrouwbaarheid in plaats van de absolute waarde hiervan. Tot aan deze stap hanteren we omwille van de vergelijkbaarheid een standaard (ordinary least squares, ofwel OLS) benadering zoals die ook in de voorgaande studie is gehanteerd. De schattingen in deze modellen gebruiken geclusterde standaardfouten omdat da data herhaalde observaties van dezelfde route betreffen. In paragraaf 6.5 hanteren we vervolgens alternatieve econometrische specificaties die meer recht doen aan de structuur van de gebruikte data. Op basis daarvan komen we tot een voorkeursspecificatie.

	Onderwerp	Methode
Vergelijkend	Herhaling Significance modellen (5.2) Waarom leiden FCD schattingen tot andere resultaten? (5.3) Heeft integratie HWN en OWN modellen meerwaarde? (5.4)	OLS modellen
Verkendend	Welk effect heeft integratie HWN en OWN delen op schatting (6.1) Welk effect heeft rijrichting op schatting (6.2) Verkenning route-specifieke factoren (6.3) <i>o.a. stroken, snelheid, weefvakken, knooppunten, trajectcontrole</i> Verkenning tijd-specifieke factoren (6.3) <i>Weersomstandigheden, wegwerkzaamheden</i>	OLS modellen Fixed-effects panel model
Specificatie	Meerwaarde loglineaire transformaties? (6.4) Endogeniteit in o.a. effect aantal rijstroken, trajectcontrole? (6.5) Random effects model met panel structuur, tests? (6.5)	OLS modellen Two-stage Least Squares Random effects model
Aanbeveling	Voorgesteld model voor schatten LMS-BT (7) Gewogen tijdvak-specifieke dummy's (7.1) Aanbeveling bijdrages HWN en OWN (7.2)	

Figuur 26 Positie verdiepende analyse in deze studie en stappenplan van haar afzonderlijke onderdelen.

6.1. Geïntegreerde benadering onbetrouwbaarheid

Als eerste stap in aanpassing van het basismodel hebben we een geïntegreerde benadering opgesteld waarin we de onbetrouwbaarheid over de complete deur-tot-deur reis als één geheel beschouwen. We combineren hierin de delen over zowel het hoofdwegennet als het onderliggend

wegennet om tot een directe inschatting totale onbetrouwbaarheid te komen (Tabel 15). Dit was in de eerdere studie niet mogelijk omdat dat de data hiervoor ontbrak, zoals aan het begin van het vorige hoofdstuk is beschreven. Ten opzichte van de apart geschatte modellen voor hoofdwegen en overige wegen (Tabel 9) blijkt deze benadering zeer vergelijkbare resultaten te leveren. De coëfficiënten hebben een vergelijkbare omvang, zijn even significant en de analyse heeft een vergelijkbare verklaringskracht. Het voordeel van deze benadering is dat een enkele set aan samenhangende coëfficiënten ontstaat die een compleet beeld van onbetrouwbaarheid geeft. Hiermee is het niet langer nodig routes in afzonderlijke HWN- en OWN-delen te verdelen.

Tabel 15 Verklaring van reistijd onbetrouwbaarheid in een geïntegreerd model op basis van trajectdelen over HWN en OWN.

	Ochtendspits	Rest dag	Avondspits
Reguliere vertraging HWN	0,149** (0,017)	0,257** (0,016)	0,073** (0,013)
ln(reg.vertraging HWN + 1)	0,572** (0,124)	0,416* (0,163)	0,717** (0,090)
Reguliere vertraging (OWN)	0,153** (0,039)	-0,009 (0,051)	0,076* (0,030)
ln(reg.vertraging OWN + 1)	0,006 (0,271)	0,502* (0,232)	0,280 (0,205)
Lengte HWN	0,015** (0,003)	-0,002 (0,003)	0,021** (0,003)
Lengte OWN	0,005 (0,005)	0,004 (0,003)	0,005 (0,006)
Constante	-0,205 (0,135)	-0,415* (0,188)	-0,365 (0,242)
Waarnemingen	960	2.400	960
Adjusted R ²	0,922	0,865	0,914
Residuele standaardfout	0,622	0,442	0,590
F Statistic	1.880**	2.558**	1.692**

Significantie codes coëfficiënten: * $p < 0,1$ ** $p < 0,01$. Standaardfouten tussen haakjes. Omdat het waarschijnlijk is dat de observaties per kwartier op dezelfde route met elkaar gecorreleerd zijn, zijn de standaardfouten geclusterd per route.

6.2. Effect rijrichting

De routeset betreft 40 herkomst-bestemmingsparen die middels primaire en secundaire routes aan elkaar verbonden zijn. Voor al deze routes beschouwen we reistijden over de hele dag voor zowel de heenrichting van herkomst naar bestemming, als de terugrichting. Als heenrichting kiezen we de verbinding die het best scoorde op de route criteria zoals omschreven in Hoofdstuk 3. Een van die criteria betrof de maximale stroomdichtheid van de route in de ochtend- of avondspits. Aangezien de avondspits vaak net iets drukker is dan de ochtendspits¹¹, zal in een aantal gevallen de drukte op dat moment doorslaggevend zijn geweest (waarmee heenrichting een wat misleidende term is). Maar voor veel routes is dit onderscheid niet zo duidelijk te maken. De korte routes, bijvoorbeeld, weerspiegelen de trips die in de directe woonomgeving worden gemaakt (bijvoorbeeld om kinderen naar school te brengen of inkopen te doen) en die op elk moment van de dag kunnen plaatsvinden. De lange routes relateren meer aan zakelijk verkeer en zullen ook niet per se aan een specifieke spitsperiode gebonden zijn.

Om te onderzoeken of de keuze van rijrichting van invloed is op de gevonden relatie tussen onbetrouwbaarheid en vertraging herhalen we de analyse uit de vorige sectie per rijrichting. Deze analyse is te beschouwen als een gevoeligheidsanalyse waarin we kijken hoe doorslaggevend de selectie van routes is voor de gevonden relaties. Tabel 16 geeft de resultaten van deze analyse weer, waarbij afzonderlijke regressies gedaan zijn voor de groep routes die we als representatief voor de heenrichting beschouwen en de groep die we tot de terugrichting rekenen. Beide groepen zijn uiteraard even groot en geven zoals gehoopt zeer vergelijkbare resultaten in termen van gevonden coëfficiënten, significantie en verklaringskracht. Reguliere vertraging blijft een goede voorspeller voor onbetrouwbaarheid. Voor wat betreft de bijdrage van het hoofdwegennet zijn de gevonden relaties voor beide groepen vrijwel gelijk. De bijdrage van het onderliggend wegennet is iets minder

¹¹ <https://www.cbs.nl/nl-nl/nieuws/2016/41/tussen-vier-en-zes-onderweg-voor-werk-of-boodschappen>

robuust, wat mogelijk aangeeft dat de doorstroming daar minder goed voorspelbaar is. De analyseresultaten blijken dus niet heel gevoelig voor de veronderstelde rijrichting. Door beide richtingen te combineren in één analyse en over de gehele dag te beschouwen geven we een zo mogelijk compleet beeld van de dagelijkse verkeerssituatie.

Tabel 16 Verklaring van reistijdontrouwbaarheid uitgesplitst voor de heen- en terugrichting en voor ochtendspits (OS), avondspits (AS) en rest dag (RD).

	Heenrichting			Terugrichting		
	OS	RD	AS	OS	RD	AS
Reguliere vertraging HWN	0,126** (0,020)	0,242** (0,035)	0,071** (0,024)	0,183** (0,030)	0,266** (0,013)	0,065** (0,015)
ln(reg.vertraging HWN + 1)	0,556** (0,135)	0,567* (0,271)	0,827** (0,154)	0,525** (0,179)	0,299* (0,179)	0,634** (0,125)
Reguliere vertraging OWN	0,133** (0,037)	0,025 (0,075)	0,0164 (0,052)	0,196** (0,065)	-0,037 (0,071)	0,099** (0,037)
ln(reg.vertraging OWN + 1)	0,330 (0,258)	0,400** (0,353)	0,681* (0,313)	-0,376 (0,485)	0,584* (0,307)	0,127 (0,268)
Lengte HWN	0,023** (0,005)	-0,005 (0,005)	0,017** (0,003)	0,009* (0,005)	0,001 (0,003)	0,028** (0,003)
Lengte OWN	-0,003 (0,007)	0,002 (0,005)	0,007 (0,008)	0,010 (0,009)	0,005 (0,004)	0,003 (0,008)
Constante	-0,536* (0,269)	-0,398 (0,302)	-0,872* (0,353)	0,182 (0,462)	-0,423* (0,231)	-0,136 (0,329)
Waarnemingen	480	1.200	480	480	1.200	480
Adjusted R ²	0,940	0,831	0,895	0,901	0,895	0,931
Residuele standaardfout	0,588	0,457	0,585	0,638	0,415	0,575
F Statistic	1.241**	986**	683**	730**	1.713**	1.079**

Significantie codes coëfficiënten: * $p < 0,1$ ** $p < 0,01$. Standaardfouten tussen haakjes. Omdat het waarschijnlijk is dat de observaties per kwartier op dezelfde route met elkaar gecorreleerd zijn, zijn de standaardfouten geclusterd per route.

6.3. Aanvullende verklarende variabelen

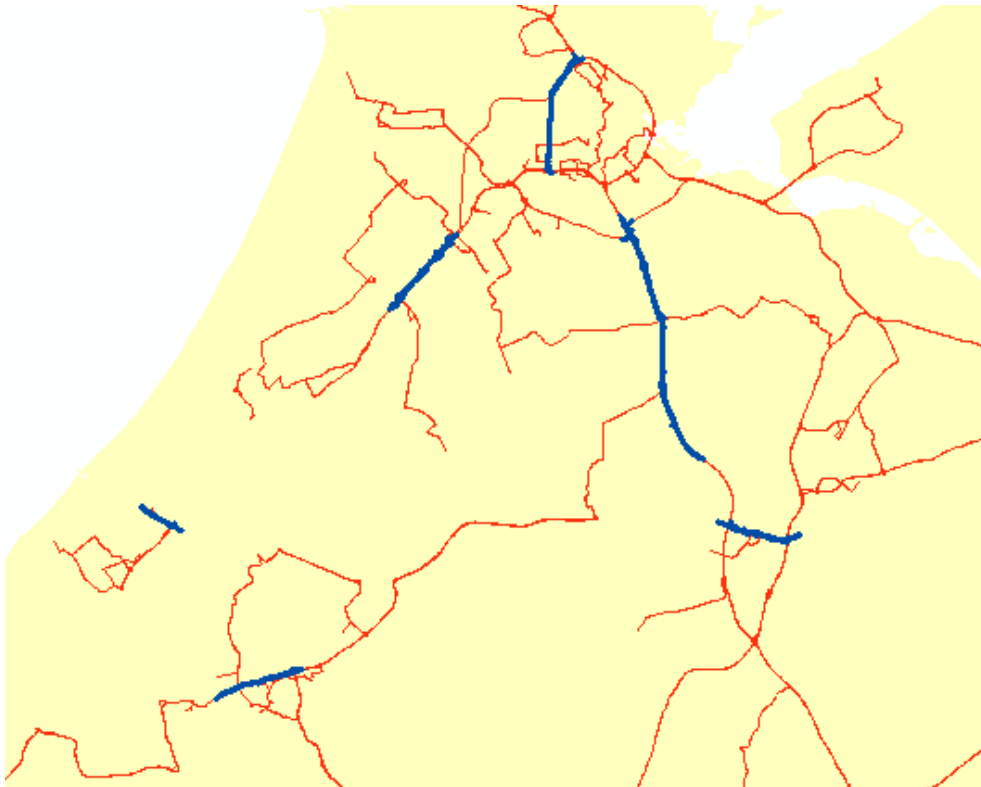
Bij de specificatie van de aanvullende verklarende variabelen is getracht deze zo op te stellen dat aansluiting op het LMS en het Nederlands Regionaal Model (NRM) mogelijk is. De eventuele toepassingen van LMS-BT (zoals effecten van klimaatverandering en frequentie extreme omstandigheden) staan daarbij voorop, maar in enkele gevallen is ook gekeken naar aspecten die vanuit wegbeheerperspectief relevant zijn (denk aan werkzaamheden).

We hebben een groot aantal variabelen beschouwd in deze aanvullende analyse en gekeken naar het effect van:

- lokale verkeersmaatregelen zoals trajectcontrole en spitsstroken;
- routekenmerken (aantallen afritten, lengte van de route, aantal rijstroken, weefvakken);
- gemiddelde hoeveelheid verkeer op route (in aanvulling op waargenomen vertraging);
- variatie in weersomstandigheden (zoals regen en mist);
- wegwerkzaamheden;
- regionale variatie (door middel van een dummy variabele per aankomstprovincie);
- vertraging in voorgaande tijdvakken (zogenaamde *lagged delays*).

Uiteindelijk bleek slechts een beperkt aantal variabelen tot robuuste, significante effect te leiden. Een belangrijk aspect daarbij zal zijn dat sommige kenmerken relatief zeldzaam zijn op de nationale schaal van de analyse. Neem bij wijze van voorbeeld trajectcontroles. Trajectcontroles komen slechts op een beperkt deel van de geselecteerde routes voor, en zijn bovendien geconcentreerd in een deel van de Randstad (Figuur 27). Het is mogelijk dat het effect van deze lokale maatregelen in

de huidige analyse lastig te onderscheiden is van de andere factoren waarop deze routes verschillen van de overige routes. Daarnaast hebben trajectcontroles waarschijnlijk invloed op zowel reistijd onbetrouwbaarheid als vertragingen. De aanwezigheid van trajectcontroles heeft volgens verschillende geprobeerde regressie specificaties steeds een monotoon verlagende invloed op onbetrouwbaarheid, die echter niet significant is in het uiteindelijk gekozen model. Wellicht zou een andere analyseopzet beter geschikt zijn het effect van dergelijke maatregelen te onderzoeken. Bijvoorbeeld door een gerichte analyse te doen van specifieke trajecten en daarbij te focussen op de situatie voor en na invoering van een specifieke maatregel. Of mogelijk zijn veel meer routes nodig om genuanceerde effecten op te vangen, zoals bij trajectcontroles waarschijnlijk het geval is.



Figuur 27 Delen van de geselecteerde routeset (in rood) waar gedurende de onderzoeksperiode trajectcontrole gold (in blauw).

Routespecifieke kenmerken

Enkele algemene routekenmerken bleken wel van toegevoegde waarde in uitgevoerde tests. Deze routekenmerken nemen we in onze voorkeurspecificatie op zoals in sectie 6.5 beschreven wordt. Dit betreft allereerst de gemiddelde freeflow snelheid zoals die over de gehele route gereden kan worden, berekend uit totale lengte van de route (in km) gedeeld door de freeflow reistijd over deze route (in uren). **Deze variabele onderscheid relatief snelle routes van de langzamere. In de testspecificaties bleek dit een belangrijker gegeven dan bijvoorbeeld gemodelleerd tijdsverlies of aantal kruisingen op een route.**

Een andere aanvullende variabele die een significant effect had op het verklaren van betrouwbaarheid, betrof het aandeel van de route dat minder dan 3 rijstroken heeft. Deze variabele geeft een indicatie van de capaciteit van de weg en gaf consistentere resultaten dan alternatieven zoals percentage van de route dat een enkele rijstrook had. Het verder uitsplitsen van deze variabele naar precies aantal rijstroken was problematisch omdat dat vaak sterk gecorreleerd is met bijvoorbeeld freeflow snelheden en de verhouding tussen de HWN en OWN delen van de route. Op routes die grotendeels uit één rijstrook bestaan zal veel van de route op het OWN liggen, en zijn de freeflow snelheden waarschijnlijk relatief laag; op routes met meer dan twee rijstroken bestaat waarschijnlijk

een groot deel uit HWN, en liggen de freeflow snelheden relatief hoog. In grote lijnen geldt dat wegen met meer dan twee rijstroken onbetrouwbaardere reistijden kennen.

Verdere tests laten zien dat het aantal rijstroken sterk is gecorreleerd met de gemiddelde gemodelleerde verkeershoeveelheid op een route. **Dit geeft aan dat onbetrouwbaarheid structureel toeneemt als gemodelleerde verkeersdrukke toeneemt.** De geschatte structurele effecten van meerdere rijstroken vervalt volledig als de gemiddelde verkeershoeveelheid als verklarende variabele wordt meegenomen. Rijstroken functioneren dus waarschijnlijk als een proxy voor vervoersvraag; alhoewel uitspraken over de richting van causaliteit tussen wegcapaciteit en weggebruik ver buiten de doelstellingen van deze studie liggen. De wenselijkheid van het opnemen van gemodelleerde verkeershoeveelheid is discutabel. Doordat het hier gemodelleerde schattingen betreft, introduceert deze variabele extra onzekerheid in het model. Aan de andere kant is het gebruik van deze variabele wel praktisch nuttig, gezien de bedoelde toepassing in LMS-BT. Om de afweging te faciliteren zijn voorkeurspecificaties opgenomen op basis van rijstroken of op basis van gemiddelde verkeershoeveelheid.

Tijdspecifieke kenmerken

Een belangrijk deel van de onderzochte aanvullende variabelen heeft als kenmerk dat zij variabel zijn in tijd en ruimte. Ze komen slechts op bepaalde momenten en bepaalde locaties voor en kunnen dus niet onderzocht worden met de benadering zoals we die tot dusver hebben toegepast. Daarin worden immers alle waarnemingen over een specifiek kwartiertijdvak over alle werkdagen van de studieperiode gemiddeld; temporeel variërende aspecten worden daarin tot een enkele, gemiddelde waarde voor de hele periode teruggebracht. Om toch een analyse te kunnen doen van deze tijdstip-specifieke effecten is een andere type analyse opgezet waarin alle afzonderlijke trips als aparte waarnemingen zijn opgenomen. Elke reis op elke vertrekminuut op elke route is hierin een aparte observatie. Dit zijn in theorie bijna 7 miljoen waarnemingen maar een deel hiervan valt weg omdat we ze classificeren als extreme reistijd; of omdat snelheidswaarnemingen voor het reismoment ontbreken; of omdat ze vallen op een van de 3 dagen waarop het onderliggende floating car data netwerk wisselt.

Met deze uitgebreide dataset is vervolgens een *fixed effects* verklarend model opgesteld waarin elke observatie is behandeld binnen een zogenaamde panelstructuur. Als verklarende variabele is gekozen voor het onverwachte deel in de reistijdcomponent; die definiëren we als $t_{r,t}^{onverwacht} = t_{r,t} - \overline{RT}_{r,k}$, waarbij $t_{r,t}$ de op basis van floating car datatrajectoriën geschatte reistijd is op een bepaalde minuut, minus \overline{RT} , de volgens Kouwenhoven en Bel (2014) gedefinieerde verwachte reistijd, welke is bepaald op basis van de gemiddelde reistijd per kwartier per werkdag in voorgaande en opvolgende weken. Het voornaamste verschil tussen de hier gebruikte onverwachte reistijd en de in de rest van de studie gebruikte onbetrouwbaarheid is dat in de laatste maat alle waarden boven nul liggen. Onze inschatting is dat de hier gebruikte maat meer informatie geeft over hoe weer en werkzaamheden leiden tot geobserveerde afwijkingen in reistijd. Voor directe toepassing in LMS-BT zijn de geschatte modellen niet geschikt.

Om rekening te houden met reguliere variatie in onverwachte reistijd hanteren we voor elke route een *fixed effect*, en daarnaast dummy variabelen waarmee algemene effecten van het moment van de dag, de specifieke werkdag en de week van de reis worden. Deze dummies zijn aantrekkelijk omdat zij het structurele karakter van de onverwachte reistijd beschrijven. Tabel 21 in bijlage 3 geeft een overzicht van deze resultaten die goed aansluiten bij de verwachtingen over onzekerheden in reistijden. **Ten opzichte van reizen op de maandag is de onverwachte reistijd het grootst op dinsdagen en donderdagen.** De variatie over de weken lijkt niet een heel structureel patroon te kennen. Binnen de dag daarentegen komen de patronen terug die we ook zagen in Figuur 10: **de onverwachte reistijd loopt sterk op in de ochtendspits tot 8:30 uur, is vrij constant (en licht**

negatief) gedurende de dag en neemt weer toe in de avondspits, tot iets lagere niveaus dan de ochtendspits. In de uiteindelijk voorgestelde specificatie betrekken we de variatie over de dag, maar laten we de variatie tussen de dagen buiten beschouwing omdat die niet in te passen is in de huidige opzet van LMS-BT.

Tijd- en ruimtespecifieke kenmerken

Met de hiervoor geïntroduceerde opzet voor het analyseren van variatie in reistijden kunnen we ook het effect van omstandigheden analyseren die variabel zijn in tijd en ruimte. Dit is van belang voor aspecten die vooral voorkomen op bepaalde momenten van de dag (bijvoorbeeld mist in de ochtend), tijdens bepaalde delen van het jaar (sneeuw in de winter), in bepaalde delen van het land (ijsvorming in Oost-Nederland), of gedurende een beperkte periode op een deel van de routes (werkzaamheden). De combinatie van tijdspecifieke dummies en route-specifieke fixed effects zorgt ervoor dat we een betrouwbaarder inschatting kunnen maken van het toegevoegde effect van dergelijke condities. Hiermee vonden we significante effecten van weersomstandigheden en wegwerkzaamheden op onbetrouwbaarheid. Daar gaan we in deze paragraaf nader op in.

Voor elk vertrekmoment is bepaald welk deel van de route te maken had met specifieke weersomstandigheden of werkzaamheden. De weergegevens komen van KNMI¹² en geven per uur aan of een specifiek weertype optrad in het voor elk wegvak dichtstbijzijnde weerstation. Vervolgens is voor elk weertype berekend over welk deel van de routelengte dat weertype voorkwam tijdens het vertrekkuur. In navolging van de literatuurstudie onderscheiden we een aantal weerscondities: mist, regen, sneeuw en ijsvorming. Die laatste variabele beschrijft of de temperatuur op het moment van waarneming of het voorgaande uur beneden de nul graden lag. Op dergelijke momenten kan gladheid optreden, maar dat hoeft niet (overall) het geval te zijn. Onder droge omstandigheden is gladheid minder een probleem en daarbij wordt in ieder geval op de hoofdwegen preventief gestrooid. Overigens was de potentieel koudste periode in deze analyse (december 2019), zeer zacht en relatief droog¹³.

De gegevens over werkzaamheden komen van Rijkswaterstaat (2020) en geven per dag en wegvak op het hoofdwegennet aan of er aan de weg gewerkt werd. Gegevens met betrekking tot werkzaamheden op het onderliggend wegennet zijn niet landsdekkend beschikbaar en daarom buiten beschouwing gelaten. Met betrekking tot wegwerkzaamheden zijn twee variabelen opgesteld. Een variabele geeft aan op welk deel van de route wegwerkzaamheden gebeurd zijn tijdens de periode van de observatie. De andere variabele geeft aan op welk deel van de route er werkzaamheden zijn geweest tijdens werkdagen. Het is te verwachten dat beide variabelen de onverwachte reistijd opstuwten. Ook als werkzaamheden alleen worden uitgevoerd in het weekend is het te verwachten dat reistijden op het traject wat hoger liggen tijdens werkdagen, bijvoorbeeld door veranderingen in belijning of tijdelijke vernauwing van rijstroken. Gedurende de studieperiode werd op 45 locaties aan de weg gewerkt. In veruit de meeste gevallen (36) duurden de werkzaamheden minder dan een week, zodat geen effect verwacht wordt op de verwacht reistijd voor een route die over een achtweeks periode bepaald wordt.

Tabel 17 toont de resultaten van deze analyse. In aanvulling op de gebruikte dummy's vinden we een significant positief (dus extra vertragend) effect van het op het uur van vertrek voorkomen van regen of sneeuw. **Sneeuw is het zeldzaamste fenomeen dat we onderzochten¹⁴, maar levert per**

¹² <https://www.daggegevens.knmi.nl/klimatologie/uurgegevens>

¹³ <https://www.knmi.nl/nederland-nu/klimatologie/maand-en-seizoensoverzichten/2019/december>

¹⁴ Dit valt af te lezen uit de gemiddelde waarde voor de variabele: met grofweg 5 miljoen beschikbare observaties betekent een waarde van 0,0004 dat sneeuw voorkwam op in totaal circa 2000 route en vertrekminuut combinaties. Dit betrof ongeveer 40 routes op twee afzonderlijk dagen: 12 november tussen

individueel voorkomen de belangrijkste bijdrage aan de onverwachte reistijd: gemiddeld 2,7 minuten wanneer sneeuwval is vastgesteld tijdens het vertrek van de trip bij het dichtstbijzijnde weerstation. Dat sneeuw het grootste effect heeft is in overeenstemming met bestaande studies (Cools e.a., 2010; Maze e.a., 2006). In onze studie is het effect op onbetrouwbaarheid relatief beperkt, maar bedacht moet worden dat geen onderscheid gemaakt is in duur of intensiteit van de weerscondities. Het effect is dus een gemiddelde van korte buien binnen hetzelfde uur in de nabijheid van de route, langdurige zware regenval tijdens de gehele route en alle gradaties daartussen. Waarbij ook nog geldt dat de extreme vertragingen buiten beschouwing zijn gelaten. Overigens zijn de effecten van de verschillende dummies en variabelen hier additief, zodat sneeuw op een maandag leidt tot 2,7 minuten, en sneeuw op een dinsdag voor $0,304 + 2,7 =$ ongeveer 3 minuten.

Verrassend genoeg verminderen mist en ijsvorming de onbetrouwbaarheid. Dit zou te maken kunnen hebben met het feit dat reizigers hun gedrag aanpassen aan dergelijke omstandigheden en bijvoorbeeld een ander tijdstip kiezen of de reis afzeggen zoals in de literatuur beschreven wordt (Cools e.a., 2010b). Zeker bij relatief zeldzame omstandigheden, die bij vertrek van huis in de morgen al bekend zijn zoals (nacht)vorst of mist kan dit een belangrijke rol spelen. Regen is in Nederland lastig te vermijden en de incidentele (natte) sneeuwval in deze periode viel gelijk met regen waardoor gedragsaanpassing voor deze omstandigheden mogelijk minder van belang is.

Werkzaamheden hebben zoals verwacht een substantieel opdrijvend effect op de onbetrouwbaarheid. We zien daarbij zowel een effect van het aanwezig zijn van waarschijnlijke tijdelijke wegaanpassingen voor de werkzaamheden, als van het op de betreffende werkdag gaande zijn van de werkzaamheden. Kennelijk zijn de voorbereide maar nog niet actieve aanpassingen ook al van invloed. Dat effect is bijna even groot als de extra bijdrage van het aanwezig zijn van actieve werkzaamheden. **Bij elkaar zorgt het aanwezig en actief zijn van de werkzaamheden gemiddeld genomen voor een onverwachte extra reistijd van $0,739 + 1,081 =$ bijna 2 minuten.**

De verklaringskracht van dit model is, gezien de adjusted R^2 , bijzonder laag, alhoewel de verklaringskracht lastig te beoordelen is omdat deze schattingsmethode alleen de aanvullende verklarende bijdrage weergeeft ten opzichte van het model met fixed effects dat de structurele variatie in onverwachte reistijden per route verklaart. Daarnaast is het model gebaseerd op minuut-specifieke deviaties ten opzichte van de verwachte reistijd, waarbij de laatste al veel van de reguliere temporele variatie in reistijden beschrijft. Wat overblijft is een stochastisch proces waarin onvoorspelbare gebeurtenissen (zoals aanrijdingen, pechgevallen, variatie in verkeersaanbod) een grote rol spelen. De lage verklarende variantie is daarmee niet heel verrassend.

Hoewel deze resultaten een interessant inzicht bieden in het effect weerseffecten en werkzaamheden, hanteren we deze aspecten niet in de het uiteindelijke model omdat dergelijke tijd-specifieke variabelen geen rol spelen in LMS-BT toepassingen. In het reguliere wegbeheer kunnen de hier gevonden relaties wel relevant zijn en daarmee illustreren deze resultaten alternatieve toepassingsmogelijkheden van de floating car data.

9:00 and 9:45 en 5 december tussen 12:00 and 12:45. Zeker die tweede datum zal geen regulier verkeersbeeld hebben gehad. Voor een betrouwbaarder inschatting van dit soort zeldzame fenomenen is een langere tijdreeks gewenst.

Tabel 17. Effect van weerseffecten en wegwerkzaamheden op reistijd onbetrouwbaarheid. Resultaten van een regressieanalyse waarin vaste effecten zijn geschat voor de gemiddelde variatie per kwartiertijdvak, afzonderlijke werkdag en week ten opzichte van een verwachte reistijd per route en kwartiertijdvak. De complete schattingsresultaten zijn te vinden in Bijlage 3.

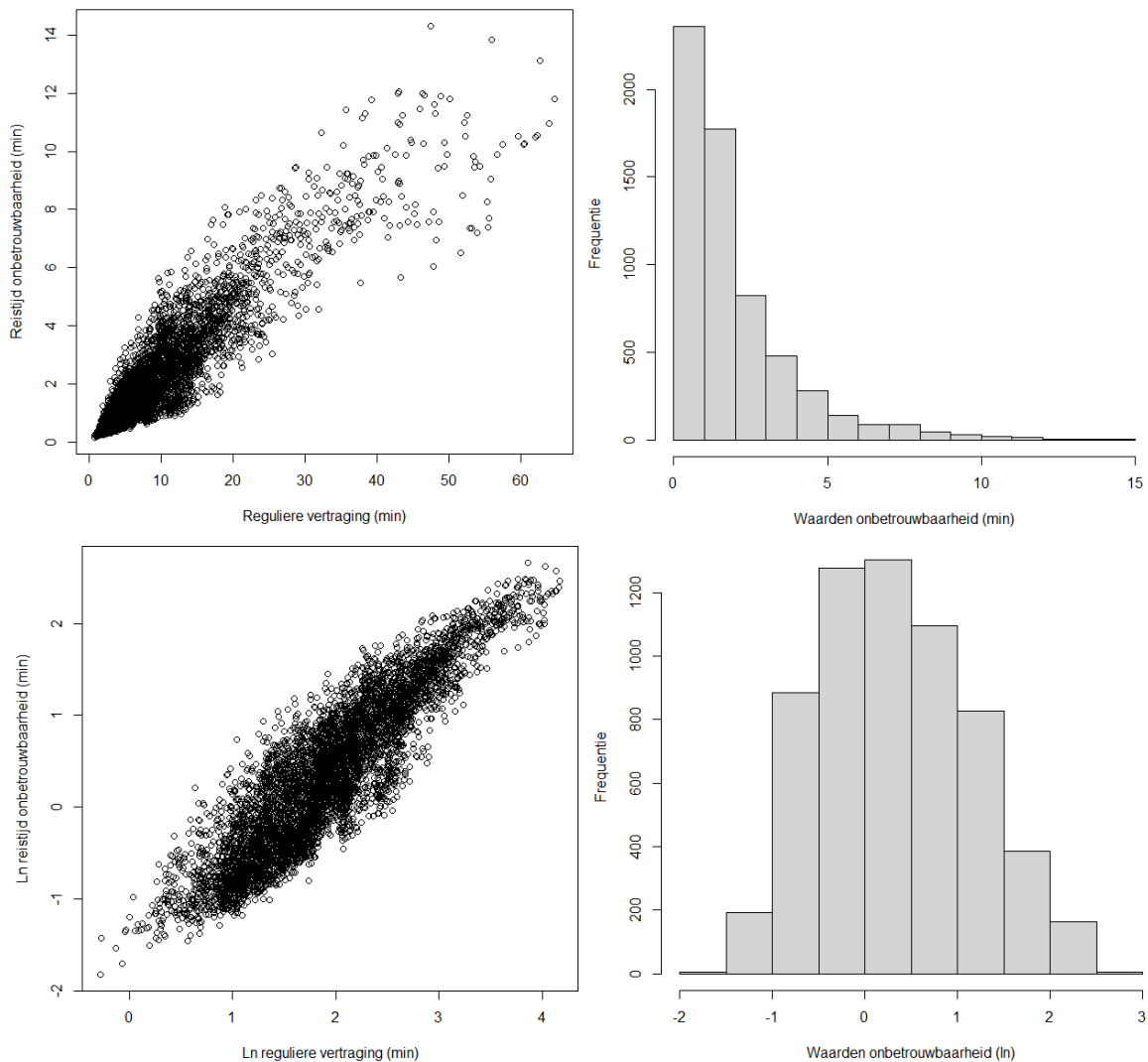
Verklarende variabele	Gemiddelde waarde in dataset	Effect op reistijdafwijking (in minuten)
Mist	0,0246	-0,031** (0,012)
Regen	0,3018	0,963** (0,004)
Sneeuw	0,0004	2,769** (0,082)
IJsvorming	0,0060	-0,175** (0,024)
Wegwerkzaamheden in periode	0,0030	0,739** (0,062)
Wegwerkzaamheden op werkdagen	0,0008	1,081** (0,137)
Kwartier dummy's (52x)	-	Ja
Weekdag dummy's (5x)	-	Ja
Week dummy's (13x)	-	Ja
Waarnemingen		5.745.480
R ²		0,035
Adjusted R ²		0,035
F Statistic		2.710**

Significantie codes coëfficiënten: * $p < 0,1$, ** $p < 0,01$. De gemiddelde waarde voor dummy (0 of 1) variabelen is te interpreteren als de fractie van de dataset met de waarde 1. De waarde 0,3 voor regen geeft dus aan dat in 30% van de observaties regen voorkwam.

6.4. Verklaren natuurlijk logaritme onbetrouwbaarheid

In voorgaande analyses is de relatie tussen onbetrouwbaarheid en vertraging beschreven met (log)lineaire relaties waarin onbetrouwbaarheid in absolute vorm (minuten) beschouwd wordt en vertraging in absolute of logaritmisch getransformeerde vorm. Net als vertraging heeft onbetrouwbaarheid echter ook een tamelijk scheve verdeling van waarden: er zijn relatief veel lage en weinig hoge waarden. Bij dergelijke waarde-verdelingen is het aan te bevelen een logaritmische transformatie toe te passen. Figuur 28 illustreert dit door de verhouding tussen beide variabelen te plotten met zowel de absolute waarden van beide variabelen (bovenste figuur) als met de natuurlijke logaritmes van die waarden. Een duidelijk voordeel van het nemen van natuurlijke logaritmes is dat de observaties zich beter spreiden en ogenschijnlijk minder risico leveren op heteroscedasticiteit in de schattingen.

De toegevoegde waarde van het verklaren van de natuurlijk logaritme van onbetrouwbaarheid in plaats van de absolute waarde hiervan is te vinden in Tabel 18. Het model behoudt zijn verklarende kracht en ook de significantie van afzonderlijke variabelen blijft in stand. Hiermee is een belangrijk econometrisch probleem in de schattingen gedeeltelijk verholpen. Er resteren echter nog een aantal andere aspecten die we in de volgende sectie bespreken.



Figuur 28 Vergelijk reguliere onbetrouwbaarheid in lineaire vorm (boven) en na het nemen van natuurlijke logaritmes (onder). De puntenwolkenrelaties (links) vergelijken de relaties tussen onbetrouwbaarheid en reguliere vertraging (op de X-as). De histogrammen (rechts) geven de frequentieverdeling van de waarden. De log-getransformeerde waarden geven een gecomprimeerder puntenwolk en normaler verdeling van waarden.

Tabel 18 Schattingsresultaten analyse natuurlijk logaritme van reistijd onbetrouwbaarheid.

	Ochtendspits	Rest dag	Avondspits
Reguliere vertraging HWN	0,003 (0,004)	0,023** (0,004)	-0,001 (0,002)
ln(reg.vertraging HWN + 1)	0,391** (0,032)	0,378** (0,051)	0,362** (0,026)
Reguliere vertraging OWN	0,027** (0,010)	-0,007 (0,022)	0,020* (0,011)
ln(reg.vertraging OWN + 1)	0,139* (0,078)	0,288** (0,104)	0,099 (0,077)
Lengte HWN	0,002** (0,0007)	0,0001 (0,0007)	0,003** (0,0004)
Lengte OWN	0,003* (0,002)	0,001 (0,001)	0,003* (0,0021)
Constante	0,091 (0,073)	-0,095 (0,077)	0,234* (0,091)
Waarnemingen	960	2.400	960
Adjusted R ²	0,895	0,840	0,885
Residuele standaardfout	0,170	0,155	0,154
F Statistic	1.359**	2.098**	1.237**

Significantie codes coëfficiënten: * p < 0,1 ** p < 0,01. 6.136 waarnemingen in alle specificaties.

Standaardfouten tussen haakjes. Omdat het waarschijnlijk is dat de observaties per kwartier op dezelfde route met elkaar gecorreleerd zijn, zijn de standaardfouten geclusterd per route.

6.5. Alternatieve specificatie om aan te sluiten bij de datastructuur

De data die we hanteren heeft een aantal kenmerken die om een specifieke benadering vragen. Een daarvan betreft heteroscedasticiteit: een ongelijke spreiding van de variantie van de te verklaren (afhankelijke) variabele over waarden van de onafhankelijke waarden zoals we in de vorige sectie bespraken. Dat is grotendeels ondervangen door de logaritmische transformatie van zowel de afhankelijke als de onafhankelijke variabele. Niet alle heteroscedasticiteit is echter tenietgedaan door logtransformaties. Voor volledige tenietdoening zouden transformaties nodig zijn die vervolgens de praktische toepassing van het model zo goed als onmogelijk maken, zodat heteroscedasticiteit hier niet volledig op te lossen is. Het risico van het bestaan van heteroscedasticiteit in schattingen is dat de daadwerkelijke numerieke effecten licht kunnen afwijken van de geschatte effecten. **Herhaling van hetzelfde model met verschillende schattingsmethodieken en verschillende subsets van de data leidt niet tot sterk verschillende uitkomsten, zodat naar onze inschatting de gebruikte coëfficiënten betrouwbaar zijn.**

Een ander relevante eigenschap van de data is ruimtelijke heterogeniteit. Bij een onvolledig gespecificeerd model varieert de grootte van de geschatte effecten afhankelijk van de locatie van de route. Het bestaan van ruimtelijke heterogeniteit in de schattingen is al gevonden in de voorgaande studie (Koomen & Jacobs-Crisioni, 2021). Om ruimtelijke heterogeniteit te meten, hebben we in aanvullende regressies getest of dummy's van de herkomst- of bestemmingsprovincie een significant effect hebben op onbetrouwbaarheid. Het opnemen van dit soort dummy's is gebruikelijk in de ruimtelijke interactie modellen literatuur (Fotheringham & O'Kelly, 1989). Als dummy's een significant effect hebben, is dat een teken dat de andere variabelen onvoldoende in staat zijn om de ruimtelijke variaties in onbetrouwbaarheid op te vangen. In deze aangepaste schatting introduceren we twee variabelen die nodig zijn voor het tenietdoen van ruimtelijke heterogeniteit. Dit betreft de natuurlijk logaritme van de gemiddelde freeflow snelheid zoals die over de gehele route gereden kan worden, en de natuurlijk logaritme van het fractie van de route dat minder dan 3 rijkstroken heeft.

Maar er speelt nog een belangrijk aspect. De hiervoor gerapporteerde schattingen zijn in navolging van de eerdere studie van Kouwenhoven en Bel (2014) uitgevoerd met de vaak gebruikte ordinary least squares methode. Deze methodiek vereist dat observaties niet onderling afhankelijk zijn. De data die wij analyseren beschrijft echter een beperkte set routes over een langere periode, waarbij herhaaldelijke metingen zijn gedaan. Dat betekent dat de waarnemingen niet onafhankelijk van elkaar zijn, we beschouwen immers dezelfde route meerdere keren achter elkaar. De data is beter te beschouwen als een panelstructuur: een set over de tijd herhaalde observaties van dezelfde individuen. Dit vereist gebruik van schattingsmethoden die expliciet voor de onderlinge afhankelijkheden in zo'n panel kunnen corrigeren.

Bij modellen met een panelstructuur moet meestal worden gecorrigeerd voor kenmerken van individuen. Dat kan door middel van individu-specifieke schatters (in een zogenaamde fixed-effect benadering) of door middel van een individu-specifieke fout-term (in een zogenaamde random-effect benadering). Fixed-effects benaderingen zijn beperkter en hebben de voorkeur als zogenaamde Hausman tests niet aantonen dat een random effect benadering acceptabel is. Zo'n fixed-effects benadering leent zich echter slecht voor toepassing op routes die niet zijn gebruikt in de modelschatting. Aangezien het aantal potentiële routes in LMS-BT vele malen groter is dan het aantal hier geanalyseerde routes, is er veel energie gestoken in het zoeken naar een model-specificatie waarbij het volgens Hausman tests is toegestaan om een random-effects model te schatten.

Een aanvullende aanpassing betreft dat we geen aparte modellen schatten voor de verschillende periode van de dag, maar kwartierspecifieke effecten in aparte dummy's opnemen. Er is immers geen reden om aan te nemen dat de respons van onbetrouwbaarheid op vertraging structureel

anders zou zijn op verschillende momenten van de dag. Deze kwartierspecifieke dummy's zijn een vereiste voor het slagen van de Hausman test. Dit heeft consequenties voor de toepassing die we verder toelichten in het volgende hoofdstuk. Tabel 19 geeft een beknopt overzicht van de resultaten, terwijl de tabellen in Bijlage 3 de complete specificatie geeft met de waarden voor de kwartier dummy's.

De kwartier dummy's werken als proxies voor een tijdsdynamiek die waarschijnlijk vooral te maken heeft met veranderingen in vertraging. Daarbij laten de verderop geanalyseerde kwartierdummy's zien dat de onbetrouwbaarheid toeneemt aan het eind van de spitsen, als de vertraging begint af te nemen. Om dit te illustreren is er hier een derde model toegevoegd, waarbij de kwartierdummy's zijn vervangen voor dummy's per dagdeel en de relatieve verandering in vertraging. De laatste variabele is gemeten als $\ln \frac{V_{i,t}}{V_{i,t-1}}$, gegeven vertraging (V), route *i*, het kwartier in *t* en het voorgaande kwartier in *t-1*. Daarbij moet worden benadrukt dat deze specificatie niet praktisch gebruikt kan worden in LMS-BT, dus hier vooral dient ter illustratie.

Tabel 19 Random-effects regressies ter verklaring van de natuurlijk logaritme van reistijd/onbetrouwbaarheid.

	Model met rijstroken	Model met verkeershoeveelheid	Model met gradiënt van vertraging
Route kenmerken			
Ln lengte (km)	-0,014 (0,028)	0,011 (0,027)	-0,049 (0,026)
Ln freeflow snelheid (km/u)	0,619** (0,127)	0,406** (0,125)	0,617** (0,122)
Ln fractie < 3 rijstroken	-0,237** (0,044)		
Ln verkeershoeveelheid		0,288** (0,027)	0,282** (0,027)
Vertragingskenmerken			
Ln vertraging	0,925** (0,016)	0,916** (0,016)	0,991** (0,013)
Ln fractie vertraging OWN	-0,110** (0,019)	-0,103** (0,019)	-0,048** (0,018)
Ln gradiënt vertraging (t/ t-1)			-0,251** (0,019)
Constante/ Controles			
Constante	-4,244** (0,478)	-5,455** (0,458)	-6,276** (0,453)
Dummy's	Per kwartier	Per kwartier	Per dagdeel
Waarnemingen	6.240	6.240	6.240
Adjusted R ²	0,791	0,794	0,777
F-Statistic	23.692	24.155	21.690

Significantie codes coëfficiënten: * p < 0,1 ** p < 0,01

Uit de resultaten blijkt dat lengte van de route een niet significant effect heeft op onbetrouwbaarheid. Bij hogere freeflow snelheden is de onbetrouwbaarheid veel hoger. Op wegen met een of twee rijstroken is de onbetrouwbaarheid lager; als de rijstroken variabele wordt vervangen voor gemiddelde verkeershoeveelheid, leidt dit tot een structureel hogere onbetrouwbaarheid bij meer verkeer. Meer reguliere vertraging zorgt voor meer onbetrouwbaarheid in reistijden. Een groter aandeel reguliere vertraging op het OWN-deel van een reis zorgt dan weer voor minder onbetrouwbaarheid. **Ook als we corrigeren voor de meestal lagere freeflow snelheden op het OWN, blijven reistijden op het OWN betrouwbaarder volgens de gebruikte data.** In het model waarbij kwartierdummy's zijn vervangen voor een variabele die de relatieve verandering in vertraging beschrijft, blijkt dat deze gradiënt in vertraging een negatief effect heeft op onbetrouwbaarheid. Vroeg in de spits wordt de onbetrouwbaarheid kleiner ten opzichte van de vertraging, en aan het eind van de spits groter.

Een laatste kenmerk van de data dat vermelding verdient, is dat observaties op verschillende routes ook onderling aan elkaar gerelateerd kunnen zijn, waardoor vertragingen op een route mogelijk bijdragen aan verhoogde onbetrouwbaarheid op een andere route. Relaties tussen routes kunnen bestaan in nauwe zin, bijvoorbeeld doordat routes een traject delen. Maar die relaties bestaan

sowieso in brede zin, doordat alle routes deel uitmaken van hetzelfde netwerk, en daardoor elkaar per definitie beïnvloeden. Om dit in te schatten zou een model moeten worden geschat dat een scheiding maakt tussen de intrinsieke effecten van de variabelen op onbetrouwbaarheid, en de ruimtelijke golf-effecten van een reistijdafwijking op alle andere routes. Met zogenaamde ruimtelijk-econometrische modellen zou dit mogelijk moeten zijn, alhoewel de combinatie van ruimtelijke en temporele dimensies kunnen zorgen voor een bijzonder uitdagende uitwerking. Gezien de complexiteit van dit fenomeen voor zowel schatting als uitvoering in LMS-BT is besloten dit probleem te laten voor verder onderzoek.

7. Voorstel aanpassing LMS-BT

Dit rapport beschrijft een uitgebreide zoektocht naar een passende specificatie van een model dat reistijd-onbetrouwbaarheid beschrijft op basis van floating car data, en dat toegepast kan worden binnen LMS-BT. Onze aanbeveling is het model te gebruiken als voorgesteld in de voorgaande sectie (Tabel 19). Dat model kan zo worden toegepast:

$$Obt = e^{-4,244+K} * (V)^{0,925} * (V_{OWN})^{-0,11} * (S_{freeflow})^{0,619} * (< 3 \text{ stroken})^{-0,237} \quad (3)$$

Of, als de specificatie met verkeershoeveelheden de voorkeur geniet:

$$Obt = e^{-5,455+K} * (V)^{0,916} * (V_{OWN})^{-0,103} * (S_{freeflow})^{0,406} * (Verkeer)^{0,288} \quad (4)$$

Daarbij is:

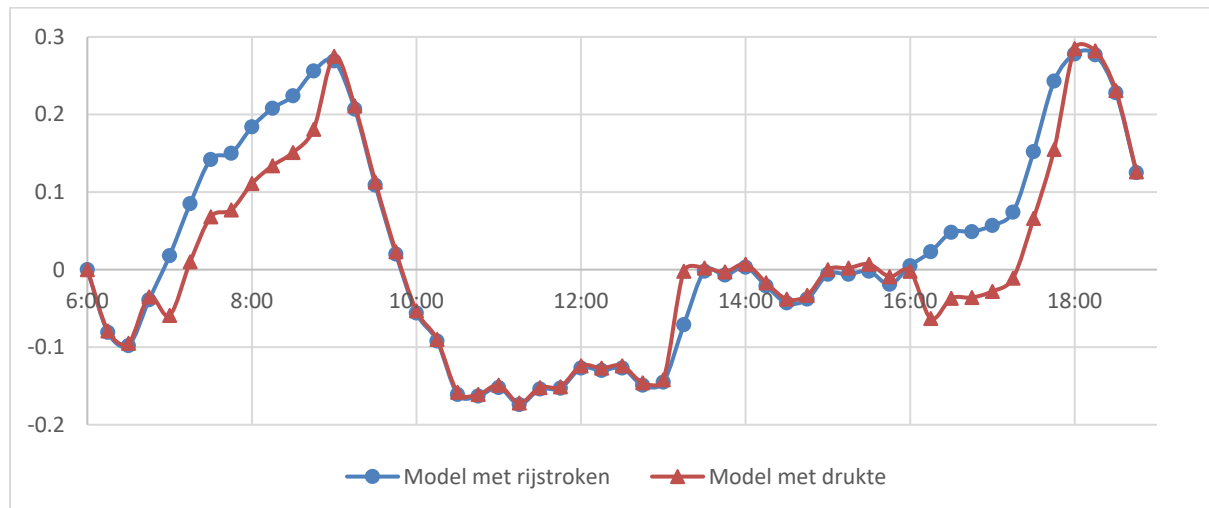
- Obt de reistijd-onbetrouwbaarheid (in minuten) voor een route, in een bepaalde periode.
- K een dagdeel-specifiek gemiddeld effect, dat hieronder wordt toegelicht.
- V de reguliere vertraging op de gehele route (zie Kouwenhoven en Bel, 2014).
- V_{OWN} de fractie van vertraging op het OWN deel van de route ten opzichte van de totale vertraging. Deze heeft een waarde tussen 0 en 1. Vertraging is gedefinieerd zoals in Kouwenhoven en Bel (2014). Als een route in totaal 10 minuten vertraging kent, waarvan 3 minuten op het OWN, dan heeft deze variabele een waarde van 0,3.
- $S_{freeflow}$ is de totale afstand van de route, gedeeld door de freeflow reistijd om de route af te leggen, volgens het 95^{ste} percentiel van de fcd snelheidsobservaties.
- < 3 stroken beschrijft de fractie van de route die minder dan 3 rijstroken telt. Deze heeft een waarde tussen 0 en 1. Als een route van 100 kilometer lang voor 75 kilometer bestaat uit een of twee rijstroken, dan heeft deze variabele een waarde van 0,75.
- Verkeer is de lengte-gewogen gemiddelde hoeveelheid verkeer op een route, berekend als de som van instroomdruktes op elke LMS link op de route op het relevante dagdeel, vermenigvuldigd met de lengte van de betreffende LMS link, en uiteindelijk gedeeld door de totale lengte van de route.

Deze schatting is op basis van een random-effects panel model waarin de effecten van zowel route-specifieke als tijd-specifieke variabelen kunnen worden geschat. Om de verklaarde variabele beter te verdelen, en problemen met heteroskedasticiteit te verminderen, is een lognormale transformatie toegepast op de verklaarde en de meeste verklarende variabelen. Het voorgestelde model geeft een geïntegreerde schatting van onbetrouwbaarheid over een gehele route, zonder dat deze wordt verdeeld over separate modellen voor de HWN en OWN delen van een route. De floating car data maken zo'n geïntegreerde benadering mogelijk. Ons geïntegreerde model heeft wel consequenties voor het gewenste rapporteren van de onbetrouwbaarheid op de HWN en OWN delen van het netwerk. Zo'n onderverdeling ligt niet voor de hand, maar kan worden geschat; we zullen hier later in deze sectie op terug komen. Door toevoeging van drie variabelen, namelijk de freeflow snelheid op de route, het percentage van de route met een of twee rijstroken, en het percentage vertraging op het OWN deel van de route, is een model geschat waarin geen significante residuele effecten zijn van herkomst- of bestemmingsprovincie. Dit model is zo betrouwbaar te gebruiken voor schattingen in het hele land. Alle voorgestelde variabelen zijn naar onze inschatting te herleiden uit het LMS wegennetwerk.

7.1. Dagdeel-specifieke dummy's

Een nieuw aspect in de voorgestelde specificatie betreft de opname van vertrekkwartier-specifieke dummy's. Deze geven de structurele bijdrage van het vertrekmoment weer op de reistijd-onbetrouwbaarheid. Het voordeel van deze benadering is dat zij op een flexibeler manier omgaat

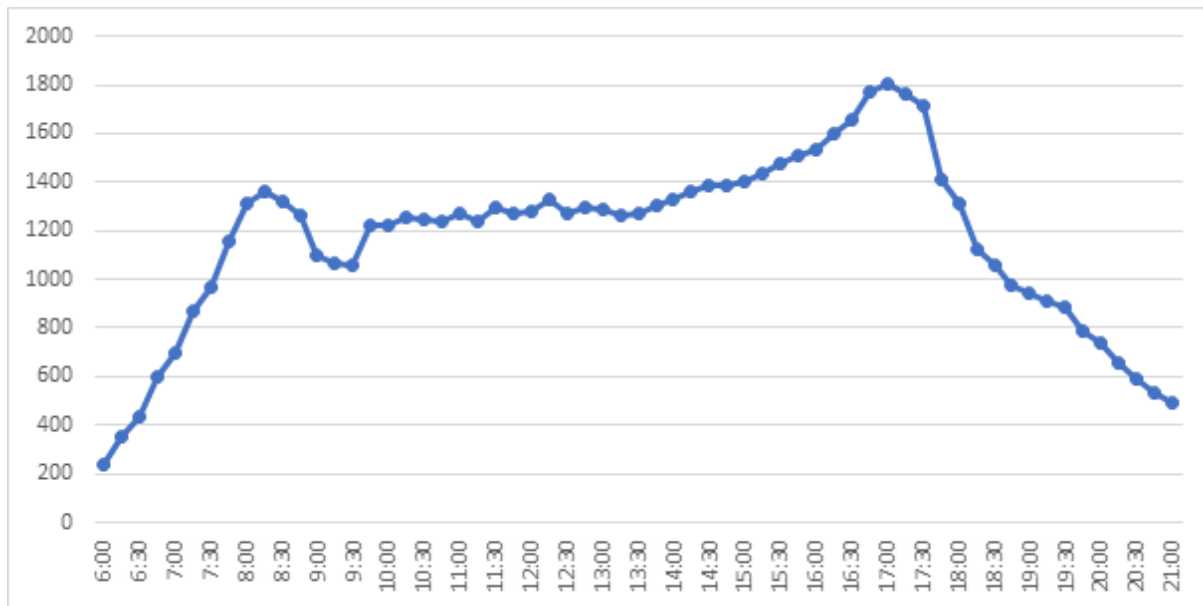
met deze temporele variatie dan in de hiervoor gebruikte opdeling in vaste tijdvakken (zoals ochtendspits, avondspits). Deze benadering was nodig om te slagen voor de Hausman test, zonder welke een in de praktijk veel minder aantrekkelijke fixed-effects specificatie nodig was. Figuur 29 geeft het verloop van deze bijdrage aan de onbetrouwbaarheid over de dag weer, voor de modellen op basis van rijstroken en verkeersshoeveelheid. **Hieruit blijkt dat onbetrouwbaarheid over het algemeen het hoogst is in de eindfases van de ochtend en avondspits.** Op die vertrekmomenten worden de reguliere vertragingen minder, maar kent de reistijd nog onzekere uitschieters.



Figuur 29 Structurele variatie in reistijd onbetrouwbaarheid over de dag op basis van het geschatte effect per vertrekkwartier.

Voor LMS-BT is het noodzakelijk de gevonden variatie in kwartier effecten om te rekenen naar dagdelen. Ons voorstel is om dit te doen met een gemiddeld dagdeel-specifiek effect, de variabele K . Als alternatief zou een range aan waarden binnen een tijdvak kunnen worden geschat, mits dat zou passen binnen de toepassingspraktijk van LMS-BT. We stellen voor om dummy's te middelen naar het relatief aantal vertrekken per kwartier binnen een dagdeel, zodat er maar een dummy hoeft te worden gebruikt per dagdeel in LMS.

Om vertrek-gewogen gemiddelden te verkrijgen zijn gewogen ODIN-data gebruikt zoals we ook in Hoofdstuk 3 toepasten. Uit ODIN zijn gerapporteerde vertrektijden voor alle autotrips genomen en geaggregeerd naar kwartieren. Het betreft hier antwoorden van respondenten die klaarblijkelijk geneigd zijn op hele of halve uren afgeronde vertrektijdstippen te rapporteren. Om een plausibeler verdeling van vertrekmomenten te verkrijgen is daarom een zwevend gemiddelde genomen met alle vertrekken in een kwartier, alsmede alle vertrekken in het half uur voor en na dat kwartier (Figuur 30). Vervolgens zijn de kwartierdummy's gewogen naar het relatieve aantal vertrekken in dat kwartier, ten opzichte van alle andere kwartieren in hetzelfde tijdvak.



Figuur 30 Zwevend gemiddelde van het aantal trips per vertrekkwartier op basis van ODiN-data.

Als alternatief is ook de gemiddelde relatieve verkeersdruk per NDW-telpunt gebruikt. De verkeersdruk per kwartier volgens die data geeft een beeld dat sterk vergelijkbaar is met ODiN-vertrekken, wat bemoedigend is voor de betrouwbaarheid van de ODiN-data. Het gebruik van verkeerstoppunten geeft echter mogelijk een vertekend beeld van het aantal vertrekken, doordat het precieze vertrekmoment van het getelde verkeer niet bekend is. Daarom gaat onze voorkeur uit naar het gebruik van tijdvak-dummy's gewogen naar ODiN-vertrekken. De waarden per tijdvak, gewogen op basis van ODiN vertrekken, NDW-tellingen, en een ongewogen gemiddelde, zijn aangegeven in Tabel 20.

Tabel 20 Voorgestelde dagdeel-specifieke dummywaardes in te vullen voor het tijdvak-gemiddelde effect K in formule (3) of (4). De dagdelen zijn gedefinieerd als in Kouwenhoven en Bel (2014) en beschreven in sectie 5.2.

Periode	Gewogen naar ODiN vertrekken	Gewogen naar NDW telpunt tellingen	Ongewogen
<i>Model op basis van rijstroken</i>			
Ochtendspits	0,172269	0,158254	0,158375
Avondspits	0,07898	0,080597	0,081375
Rest dag	-0,09708	-0,09564	-0,0981
Overig	0,094838	0,075818	0,078875
<i>Model op basis van verkeershoeveelheid</i>			
Ochtendspits	0.098222	0.08403	0.084125
Avondspits	0.002429	0.004563	0.0055
Rest dag	-0.09068	-0.08917	-0.09165
Overig	0.100362	0.080654	0.0835625

7.2. Splitsing van HWN en OWN bijdragen voor rapportages

Met de voorgaande LMS-BT modellen (Kouwenhoven en Bel, 2014) is het mogelijk om de contributies van het HWN en OWN aan totale route-onbetrouwbaarheid separaat te schatten. Dit werd binnen de LMS praktijk gebruikt voor rapportages. Met het voorgestelde, geïntegreerde model in vergelijking (3) is het helaas niet eenvoudig om de contributies van onderliggende roudedelen precies te scheiden. Voor het geval het voor rapportage toch nodig is om HWN en OWN bijdragen te scheiden, adviseren wij om vergelijking (3) of (4) te herhalen voor de gescheiden HWN en OWN delen; en vervolgens de verhouding tussen die separate HWN en OWN-delen toe te passen op de geschatte totale onbetrouwbaarheid.

Dit leidt tot de volgende formules, gegeven het model met verkeershoeveelheid:

$$\begin{aligned} \hat{A}^{OWN} &= e^{-5,455+K} * (V^{OWN})^{0,916} * (\%V_{OWN})^{-0,103} * (S_{freeflow}^{OWN})^{0,406} * (Verkeer^{OWN})^{0,288} \\ \hat{A}^{HWN} &= e^{-5,455+K} * (V^{HWN})^{0,916} * (\%V_{OWN})^{-0,103} * (S_{freeflow}^{HWN})^{0,406} * (Verkeer^{HWN})^{0,288} \\ \hat{O}bt^{OWN} &= Obt * (\hat{A}^{OWN} / [\hat{A}^{OWN} + \hat{A}^{HWN}]) \\ \hat{O}bt^{HWN} &= Obt - \hat{O}bt^{OWN} \end{aligned}$$

Daarbij is het advies om de fractie vertraging op het onderliggend wegennet gelijk te wegen, aangezien deze variabele niet zinnig te scheiden is voor de separate OWN en HWN delen van de route.

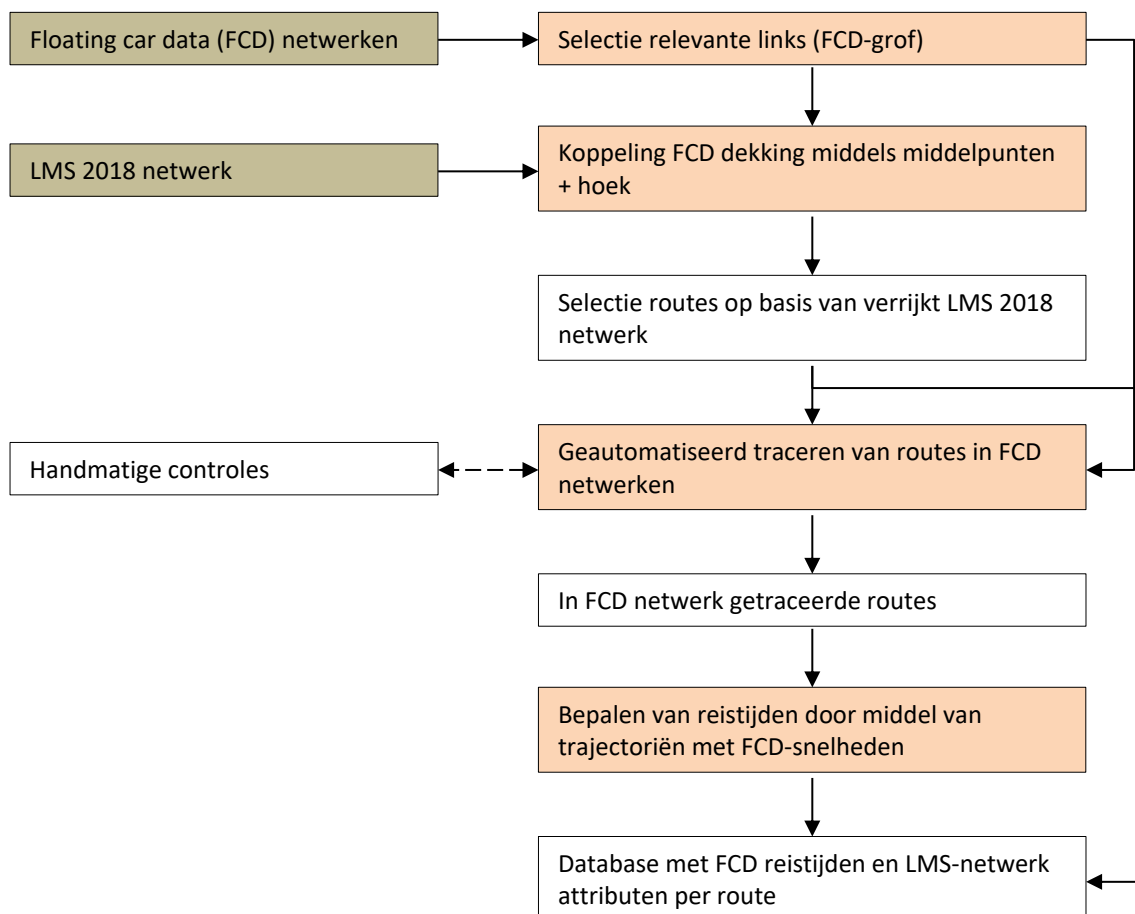
Referenties

- CBS (2020). Onderweg in Nederland (ODiN) 2019; Onderzoeksbeschrijving. Centraal Bureau voor de Statistiek (CBS). https://www.cbs.nl/-/media/_pdf/2020/37/onderzoeksbeschrijvingodin-2019v10.pdf, data beschikbaar via DANS: <https://doi.org/10.17026/dans-xpv-mwpg>
- Brockfeld, E., Lorkowski, S., Mieth, P., Wagner, P. (2007). Benefits and limits of recent floating car data technology - An evaluation study. 11th World conference on transport research: 1-13.
- Caceres, H., Ha, H., Qing, H. (2016). Estimating freeway route travel time distributions with consideration to time-of-day, inclement weather, and traffic incidents. *Journal of Advanced Transportation*: 967-987.
- Carrion, C., Levinson, D. (2013). Valuation of travel time reliability from a GPS-based experimental design. *Transportation Research Part C: Emerging Technologies*: 305-323.
- Cools, M., Moons, E., Wets, G. (2010). Assessing the Impact of Weather on Traffic Intensity. *American Meteorological Society*, 60-68.
- Cools, M., Moons, E., Creemers, L., Wets, G. (2010b). Changes in Travel Behavior in Response to Weather Conditions Do Type of Weather and Trip Purpose Matter? *Transportation Research Record*, 22-28.
- Fotheringham A.S., O'Kelly M.E. (1989). *Spatial interaction models: formulations and applications*. Kluwer, Dordrecht
- Javid, R.J., Javid, R.J. (2018). A framework for travel time variability analysis using urban traffic incident data. *IATSS Research*: 30-38.
- Kerner, B.S., Demir, C., Herrtwich, R.G., Klenov, S.L., Rehborn, H., Aleksic, M., Haug, A. (2005). Traffic state detection with floating car data in road networks. *Proceedings of the 8th international IEEE Conference on Intelligent Transportation Systems*: 700-705.
- Knoope, M. (2020). Floating car data: geschikt om reistijdverlies in kaart te brengen? Kennisinstituut voor Mobiliteitsbeleid. Ministerie van Infrastructuur en Waterstaat, Den Haag. 28 pp.
- Koomen, E., Andr e, B.P.J., Jacobs-Crisioni, C. (2018). Verkenning van de toepassingsmogelijkheid van floating car data voor het beschrijven van verplaatsingsgedrag. Rapport in opdracht van Rijkswaterstaat-CIV. Vrije Universiteit, Amsterdam. 30 pp.
- Koomen, E., Jacobs-Crisioni, C. (2021). Deur-tot-deur reistijdbetrouwbaarheid op basis van floating car data. Rapport in opdracht van Rijkswaterstaat-CIV. Vrije Universiteit, Amsterdam. 71 pp.
- Kouwenhoven, M., Bel, N. (2014). Raming omvang betrouwbaarheid Nederlandse wegennet. Eindrapport (versie 5). Rapport voor Rijkswaterstaat Water, Verkeer en Leefomgeving. Significance. 124 pp.
- Li, R., Rose, G., Sarvi, M. (2006). Using automatic vehicle identification data to gain insight into travel time variability and its causes. *Transportation Research Records*: 24-32.
- Turpijn, B., Visser, E. (2020). Nieuwe maat voor betrouwbaarheid: de planningstijdindex. *NM magazine* 15 (2): 32-33. https://www.nm-magazine.nl/pdf/NM_Magazine_2020-2.pdf
- Maze, T., Agarwal, M., Burchett, G. (2006). Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow. *Transportation Research Record*, 170-176.
- Rijkswaterstaat (2020). Rapportage Rijkswegennet 3e periode 2019, 1 september-31 december, RWS Informatie, 10 maart 2020.
- Rijkswaterstaat (2021). Plan van aanpak: Betrouwbaarheid reistijden. Project: Betrouwbaarheid reistijden. Rijkswaterstaat. Juli 2021. 6 pp.
- Willigers, J., Kouwenhoven, M. (2020). Bepaling congestie met Floating Car Data. Ontwikkeling bereikbaarheid OWN 2017-2018. Technische Rapportage. Den Haag: Significance.

Bijlage 1 Koppeling LMS en floating car data netwerken

Aan de basis van dit project staan twee netwerken die niet direct verenigbaar zijn. Het LMS-netwerk is een versimpelde, relatief grofmazige aggregatie van het nationaal wegenbestand (NWB), dat vervolgens verrijkt is met veel attributen die relevant zijn voor verkeersmodellering. De floating car data (FCD) zijn gekoppeld aan bijzonder fijnmazige netwerken die afkomstig lijken te zijn van het OpenStreetMap (OSM) netwerk. Kenmerken die beide netwerken delen zijn topologische correctheid en geografisch detail.

Voor dit project was het nodig gegevens te koppelen uit beide netwerken voor de geanalyseerde routes. Het koppelen van netwerken is notoir lastig. Er is hier een iteratieve techniek toegepast waarin veel is geautomatiseerd, en pas op een laat moment handmatige controles nodig waren. Een aanvullende complicatie is dat het floating car data netwerk ongeveer maandelijks wordt bijgewerkt, waardoor er voor de studie vier netwerken moesten worden gekoppeld aan het LMS-netwerk. De procedure wordt in de rest van deze bijlage stapsgewijs uitgelegd; en is schematisch weergegeven in Figuur 31.



Figuur 31 Schematische weergave totstandkoming reistijden-database. Groen gemarkeerde velden geven de input data weer die gekoppeld moesten worden, de oranje velden de belangrijkste stappen in het genereren van de reistijden database zoals die in deze bijlage beschreven worden.

Selectie relevante links in floating car data netwerken

Als eerste stap is een selectie gemaakt van links in de FCD-netwerken, welke door hun plek in de weghiërarchie of karakter zeer waarschijnlijk ook in het LMS-netwerk voorkomen. Dit is nodig om latere koppelingen beter uitvoerbaar te maken (doordat niet-relevante kandidaat koppels buiten

beschouwing worden gelaten) en het versnelt de verwerkingstijd. Om dit te bereiken zijn de aanknooppunten van alle LMS-subzones in Nederland gekoppeld aan FCD links; en vervolgens zijn alle FCD links geselecteerd waarover laagste-weerstand verbindingen lopen naar de dichtstbijzijnde 15% van LMS-subzone-aanknooppunten. Er is voor gekozen om, vanaf elk aanknooppunt, slechts naar een beperkt deel van de aanknooppunten te routeren om de rekenvraag beheersbaar te houden. Bij het bepalen van routes zijn gewogen netwerkweerstand gebruikt op basis van reistijden, gebaseerd op maximumsnelheden volgens FCD; en een gewicht dat wegen met een lokale functie (FCD road class > 3) twee keer zo zwaar weegt als doorstromingswegen, om het algoritme een preferentie te geven voor doorstromingswegen. Deze actie geeft een netwerk van FCD links dat qua grofmazigheid ongeveer gelijk is aan het LMS-netwerk. We noemen deze netwerken FCD-grof.

Koppeling van FCD dekkingsgraad aan LMS netwerk

Het LMS 2018 netwerk is gebruikt als basis om de routeset vast te stellen. Daarvoor moest eerst een FCD-grof specifiek attribuut, de dekkingsgraad per link, gekoppeld worden aan LMS-links. Voor die koppeling is het LMS-netwerk beschreven met zogenaamde dynamische punten. Die dynamische punten beschrijven LMS-links met intervallen van 25 meter. In het geval dat een weg twee richtingen kent wordt die in zowel het FCD-grof als LMS-netwerk beschreven door twee verder identieke links met tegenovergestelde richtingen. Voor de LMS dynamische punten is bepaald of de punten deel zijn van twee identieke links, en in die gevallen is het aantal dynamische punten gereduceerd tot een verzameling unieke punten voor beide richtingen. Een markering geeft aan of de punten een of twee richtingen beschrijven.

De koppeling is vervolgens gebeurd op basis van de nabijheid van LMS dynamische punten ten opzichte van het middelpunt van FCD-grof links. In het geval van LMS links met twee-richting verkeer, zijn twee FCD-grof links voor tegenovergestelde richtingen zo vaak gekoppeld aan hetzelfde dynamische punt. Een aanvullende berekening bepaalt daarom de hoek van de FCD-grof links en beide LMS-richtingen, en koppelt de FCD-grof links dan aan de LMS-richting waarmee het hoekverschil het kleinst is. Het uiteindelijke resultaat is een FCD-grof attribuut dat aangeeft met welke LMS-link elke FCD-grof link is gekoppeld. Als de afstand tussen FCD-grof en LMS groter is dan 500 meter, of het hoekverschil te groot is, wordt dit attribuut overigens op NULL gezet, zodat de bewuste FCD-grof link aan geen enkele LMS-link wordt gekoppeld.

Na totstandkoming van de koppeling, is de FCD-dekkingsgraad gemiddeld voor elke gekoppelde LMS-link. Daarbij is ook een inschatting gemaakt van de kwaliteit van de koppeling, op basis van de ratio tussen de totale lengte van de gekoppelde FCD-grof links gedeeld door de lengte van de LMS-link. Deze actie is voor alle vier de FCD-grof netwerken herhaald, waarna de gemiddelde dekkingsgraad per LMS-netwerk is berekend. De dekkingsgraad is in kaart gebracht in Figuur 32.

Traceren van routes

In het rapport is beschreven hoe routes zijn geselecteerd op basis van het verrijkte LMS-netwerk. Het eindproduct van deze routeselectie is een shapefile die elke route beschrijft als LMS-links, de zogenaamde LMS-routes. Vervolgens is het nodig om die routes te beschrijven als FCD-grof links, zodat passende trajectoriën kunnen worden gegenereerd voor elke route. Daarbij moeten die routes in FCD-grof topologisch correct zijn, omdat de links voor het bepalen van de trajectoriën in de juiste sequentie worden gezet op basis van hun positie in de route.

Om dit te bewerkstelligen zijn routes bepaald op de FCD-grof netwerken tussen de in de eerste stap aan het FCD-netwerk gekoppelde LMS zone-aanknooppunten, welke de begin- en eindpunten vormen van een LMS-route. Voor elke geselecteerde route is vervolgens eerst een buffer gemaakt van 1000 meter, waarna alleen de links in FCD-grof zijn geselecteerd die binnen die buffer vallen. Daaropvolgend hebben alle links in FCD-grof een attribuut gekregen dat de gekwadrateerde afstand

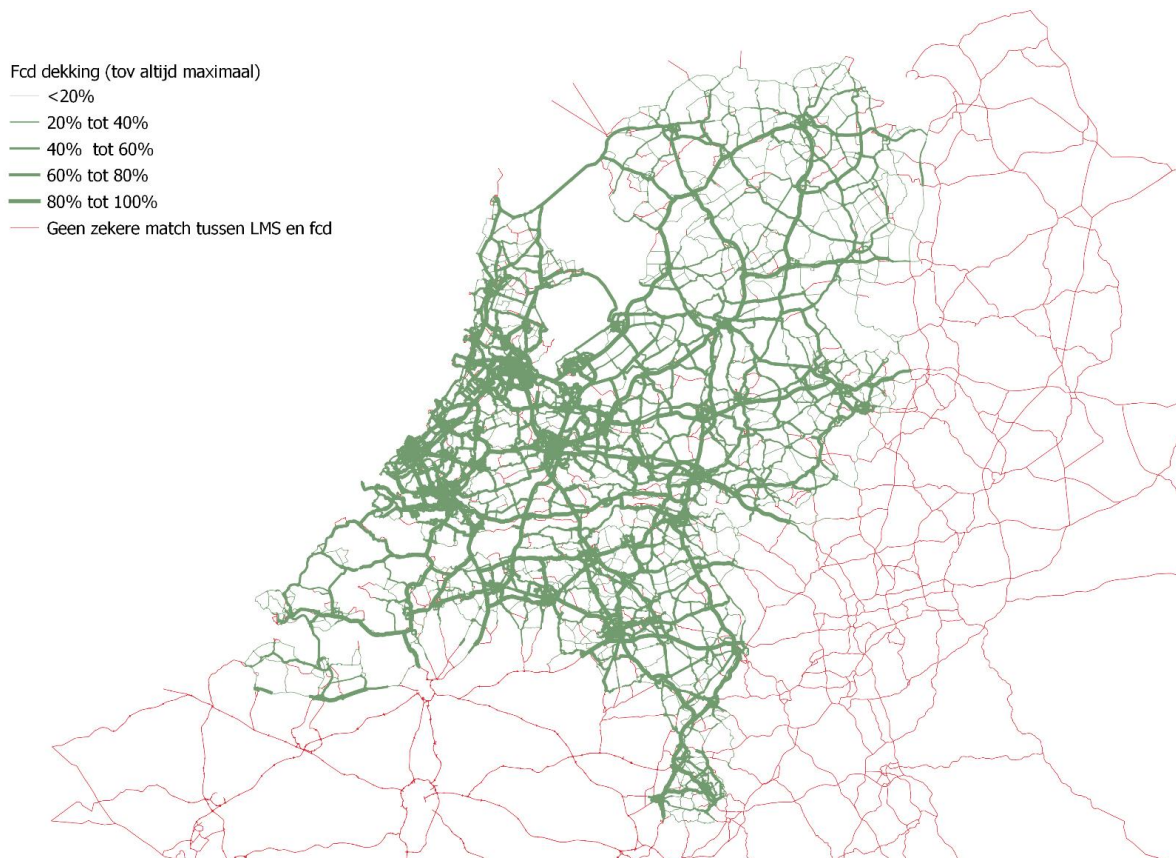
tussen dynamische punten van de LMS-route en het middelpunt van FCD-grof links beschrijft. Deze gekwadraterde afstand is gebruikt als weerstand. Daarmee is het pad van de minste weerstand bepaald tussen begin- en eindpunt. Deze methodiek geeft een topologisch correct pad dat zo veel mogelijk op een LMS-route lijkt.

In voorkomende gevallen leidt het zo traceren van routes toch tot fouten. Alle getraceerde FCD-routes, voor alle vier de FCD-netwerken, zijn handmatig nagelopen om eventuele fouten te traceren en te corrigeren. Hierbij zijn bijvoorbeeld foutjes in de richtingen van het FCD-netwerk gevonden of gevallen ontdekt waarbij een LMS route een bocht afsneed via een afrit en oprit. Waar nodig zijn daarom FCD-links van richting veranderd, of hebben FCD-links een kunstmatige weerstandsboete gekregen, om de tracés de juiste kant op te begeleiden. Na een paar controleslagen waren de FCD-routes klaar voor de volgende stap.

Bepalen van trajectoriën en koppelen van LMS-route informatie

Op basis van de getraceerde FCD-routes en de FCD-snelheden database zijn trajectoriën bepaald. Dit is gedaan met behulp van de dynamische array methode die is voorgesteld door Koomen en Jacobs-Crisioni (2021). Het resultaat van elke trajectorie is een schatting van de totale reistijd op een route voor elke minuut binnen het geanalyseerde tijdframe. Dit leidt tot een database met bijna 7 miljoen reistijden voor alle 120 routes.

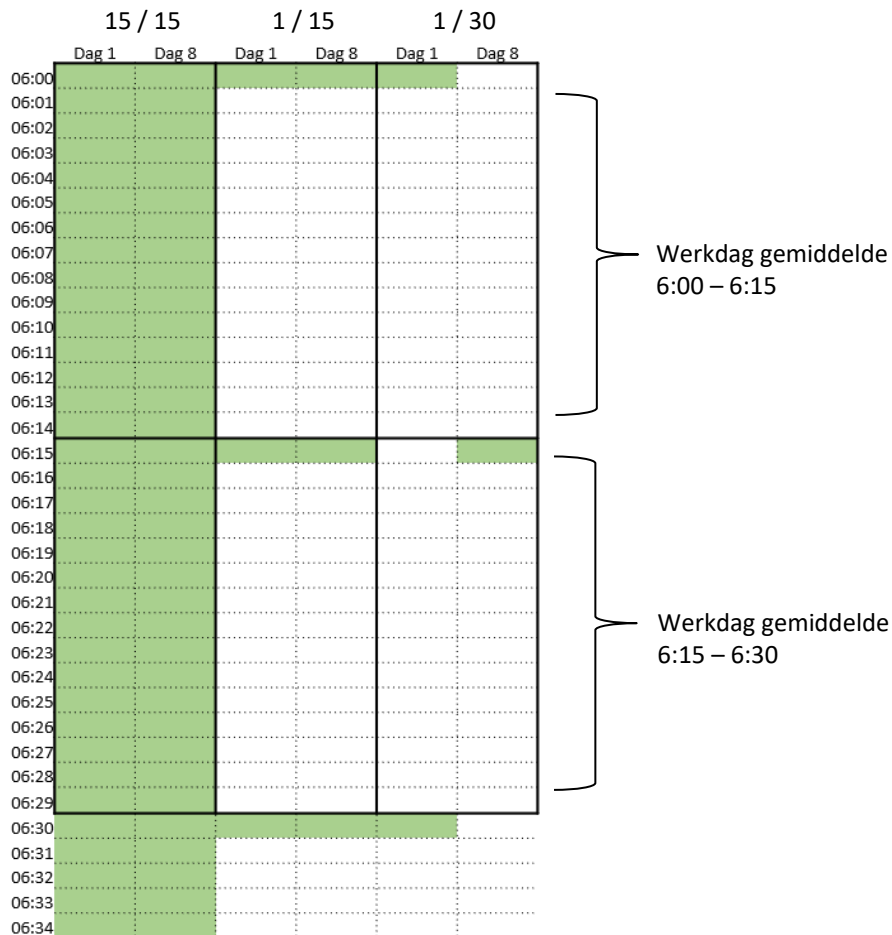
Naast de reistijdinformatie, is voor elke route ook een groot aantal route-specifieke variabelen opgeslagen. Deze informatie is geaggregeerd vanuit de LMS-routes shapefile die is beschreven in de vorige stap.



Figuur 32 Dekking floating car data op het LMS-netwerk na koppeling.

Bijlage 2 Verschillende selectiemethoden voor bepalen gemiddelde reistijden

De bijlage geeft een overzicht van de verschillende selectiemethoden om gemiddelde reistijden te bepalen. De met groen aangegeven vakken zijn vertrektijden die zijn opgenomen verschillende subsets die beschreven zijn in paragraaf 5.3. De waarnemingen in deze subsets zijn gebruikt voor het bepalen van de gemiddelde reistijd voor een bepaalde werkdag en vertrektijd, en voor het bepalen van standaarddeviaties rond die gemiddelden.



Figuur 33 Schematische weergave van de verschillende selectiemethoden om gemiddelde reistijden te bepalen.

Bijlage 3 Complete schattingsresultaten voor de modellen met tijdsdummies

Tabel 21 Volledige resultaten fixed-effects model voor verklaren van temporele variatie in onverwachte reistijd in minuten.

Weersomstandigheden			
Mist	-0,031** (0,012)	Week 37	0,165** (0,008)
Regen	0,963** (0,004)	Week 38	0,473** (0,008)
Sneeuw	2,769** (0,082)	Week 39	0,964** (0,008)
IJsvorming	-0,175** (0,024)	Week 40	1,540** (0,008)
Wegwerkzaamheden			
Aanwezig	0,739** (0,062)	Week 41	1,143** (0,008)
Op werkdagen	1,081** (0,137)	Week 42	0,712** (0,008)
Dag			
Maandag	0 (referentie)	Week 43	0,786** (0,008)
Dinsdag	0,304** (0,005)	Week 44	1,319** (0,008)
Woensdag	0,030** (0,005)	Week 45	0,517** (0,008)
Donderdag	0,276** (0,005)	Week 46	0,298** (0,008)
Vrijdag	0,126** (0,005)	Week 47	1,019** (0,008)
Waarnemingen	5.745.480	Week 48	1,455** (0,009)
Adjusted R ²	0,035	Week 49	1,220** (0,008)
Vertrekkwartier			
06:00	0 (referentie)	12:30	-0,035* (0,015)
06:15	0,005 (0,015)	12:45	-0,028 (0,015)
06:30	0,007 (0,015)	13:00	-0,033* (0,015)
06:45	0,016 (0,015)	13:15	-0,032* (0,015)
07:00	0,064** (0,015)	13:30	-0,029 (0,015)
07:15	0,105** (0,015)	13:45	-0,024 (0,015)
07:30	0,156** (0,015)	14:00	-0,045** (0,015)
07:45	0,179** (0,015)	14:15	-0,053** (0,015)
08:00	0,211** (0,015)	14:30	-0,037* (0,015)
08:15	0,232** (0,015)	14:45	-0,030* (0,015)
08:30	0,218** (0,015)	15:00	0,014 (0,015)
08:45	0,173** (0,015)	15:15	0,018 (0,015)
09:00	0,089** (0,015)	15:30	0,006 (0,015)
09:15	0,034* (0,015)	15:45	0,016 (0,015)
09:30	-0,002 (0,015)	16:00	0,016 (0,015)
09:45	-0,023 (0,015)	16:15	0,033* (0,015)
10:00	-0,031* (0,015)	16:30	0,066** (0,015)
10:15	-0,035* (0,015)	16:45	0,084** (0,015)
10:30	-0,038* (0,015)	17:00	0,095** (0,015)
10:45	-0,036* (0,015)	17:15	0,101** (0,015)
11:00	-0,029* (0,015)	17:30	0,080** (0,015)
11:15	-0,033* (0,015)	17:45	0,063** (0,015)
11:30	-0,032* (0,015)	18:00	0,035* (0,015)
11:45	-0,030* (0,015)	18:15	0,015 (0,015)
12:00	-0,030* (0,015)	18:30	0,013 (0,015)
12:15	-0,036* (0,015)	18:45	0,006 (0,015)

Significantie codes coëfficiënten: * p < 0,1** p < 0,01. Standaardfouten tussen haakjes.

Tabel 22 Volledige resultaten voorgesteld model ter verklaring van reistijdonbetrouwbaarheid op basis van random-effects regressie en onder meer het aandeel van de route met minder dan drie rijstroken.

Route kenmerken		Regressie kenmerken	
Ln lengte (km)	-0,014 (0,028)	Model	Rijstroken
Ln freeflow snelheid (km/u)	0,619** (0,127)	Waarnemingen	6.240
Ln < 3 rijstroken	-0,237** (0,044)	Adjusted R ²	0,791
Vertragingkenmerken			
Ln vertraging	0,925** (0,016)		
Ln vertraging OWN	-0,110** (0,019)		
Constante			
Constante	-4,244** (0,478)		
Vertrekkwartier dummies			
06:00	0 (referentie)	12:30	-0,127** (0,034)
06:15	-0,081* (0,032)	12:45	-0,149** (0,034)
06:30	-0,098** (0,033)	13:00	-0,145** (0,034)
06:45	-0,039 (0,033)	13:15	-0,071* (0,034)
07:00	0,018 (0,034)	13:30	-0,002 (0,034)
07:15	0,085* (0,034)	13:45	-0,007 (0,034)
07:30	0,142** (0,035)	14:00	0,003 (0,034)
07:45	0,150** (0,036)	14:15	-0,021 (0,034)
08:00	0,184** (0,036)	14:30	-0,043 (0,034)
08:15	0,208** (0,036)	14:45	-0,038 (0,034)
08:30	0,224** (0,036)	15:00	-0,006 (0,035)
08:45	0,256** (0,035)	15:15	-0,006 (0,035)
09:00	0,269** (0,034)	15:30	-0,002 (0,035)
09:15	0,207** (0,034)	15:45	-0,019 (0,036)
09:30	0,109** (0,033)	16:00	0,005 (0,036)
09:45	0,020 (0,033)	16:15	0,023 (0,037)
10:00	-0,056 (0,033)	16:30	0,048 (0,037)
10:15	-0,092** (0,033)	16:45	0,049 (0,037)
10:30	-0,161** (0,033)	17:00	0,057 (0,037)
10:45	-0,163** (0,033)	17:15	0,074* (0,037)
11:00	-0,152** (0,033)	17:30	0,152** (0,036)
11:15	-0,174** (0,033)	17:45	0,243** (0,035)
11:30	-0,154** (0,033)	18:00	0,278** (0,034)
11:45	-0,153** (0,033)	18:15	0,277** (0,033)
12:00	-0,127** (0,034)	18:30	0,228** (0,033)
12:15	-0,130** (0,034)	18:45	0,125** (0,033)

Significantie codes coëfficiënten: * p < 0,1** p < 0,01. Standaardfouten tussen haakjes.

Tabel 23 Volledige resultaten voorgesteld model ter verklaring van reistijdonbetrouwbaarheid op basis van random-effects regressie en onder meer de verkeershoeveelheid.

Route kenmerken		Regressie kenmerken	
Ln lengte (km)	0,011 (0,027)	Model	Verkeershoeveelheid
Ln freeflow snelheid (km/u)	0,406** (0,125)	Waarnemingen	6.240
Ln verkeershoeveelheid	0,288** (0,027)	Adjusted R ²	0,794
Vertragingskenmerken			
Ln vertraging	0,916** (0,016)		
Ln vertraging OWN	-0,103** (0,019)		
Constante			
Constante	-5,455** (0,458)		
Vertrekkwartier dummies			
06:00	0 (referentie)	12:30	-0,124** (0,033)
06:15	-0,079* (0,032)	12:45	-0,146** (0,033)
06:30	-0,095** (0,033)	13:00	-0,142** (0,033)
06:45	-0,035 (0,033)	13:15	-0,068* (0,033)
07:00	-0,059 (0,034)	13:30	0,002 (0,033)
07:15	0,010 (0,035)	13:45	-0,003 (0,034)
07:30	0,068 (0,036)	14:00	0,007 (0,034)
07:45	0,077* (0,036)	14:15	-0,017 (0,034)
08:00	0,111** (0,036)	14:30	-0,038 (0,034)
08:15	0,134** (0,036)	14:45	-0,033 (0,034)
08:30	0,151** (0,036)	15:00	0,0003 (0,034)
08:45	0,181** (0,036)	15:15	0,002 (0,035)
09:00	0,275** (0,034)	15:30	0,007 (0,035)
09:15	0,211** (0,033)	15:45	-0,009 (0,036)
09:30	0,113** (0,033)	16:00	-0,081* (0,037)
09:45	0,023 (0,033)	16:15	-0,063 (0,037)
10:00	-0,053 (0,033)	16:30	-0,037 (0,037)
10:15	-0,090** (0,033)	16:45	-0,036 (0,038)
10:30	-0,158** (0,033)	17:00	-0,028 (0,038)
10:45	-0,161** (0,033)	17:15	-0,011 (0,038)
11:00	-0,149** (0,033)	17:30	0,066 (0,037)
11:15	-0,172** (0,033)	17:45	0,155** (0,036)
11:30	-0,152** (0,033)	18:00	0,285** (0,034)
11:45	-0,151** (0,033)	18:15	0,282** (0,033)
12:00	-0,124** (0,033)	18:30	0,231** (0,033)
12:15	-0,127** (0,033)	18:45	0,126** (0,033)

Significantie codes coëfficiënten: * p < 0,1** p < 0,01. Standaardfouten tussen haakjes.

Tabel 24 Volledige resultaten voorgesteld model ter verklaring van reistijdonbetrouwbaarheid op basis van random-effects regressie en onder meer verkeershoeveelheid en verandering in vertraging.

Route kenmerken		Regressie	
Ln lengte (km)	-0,049 (0,026)	Model	Verandering verkeershoeveelheid
Ln freeflow snelheid (km/u)	0,617** (0,122)	Waarnemingen	6.240
Ln verkeershoeveelheid	0,282** (0,027)	Adjusted R ²	0,777
Ln Δ verkeershoeveelheid	-0,251** (0,019)		
Vertragingskenmerken			
Ln vertraging	0,991** (0,013)		
Ln vertraging OWN	-0,048** (0,018)		
Ln gradiënt vertraging	-0,251** (0,019)		
Constantes			
Constante	-6,276** (0,453))		
Ochtendspits	0,118** (0,012)		
Avondspits	0 (referentie)		
Rest dag	-0,050** (0,015)		
Overige dag	0,119** (0,015)		

Significantie codes coëfficiënten: * p < 0,1 ** p < 0,01. Standaardfouten tussen haakjes.