



Adviesdienst Verkeer en Vervoer

**GSM als informatiebron
voor Verkeersbeleid
Simulatiestudie**

- Uitgebreide samenvatting -

Modelit
Rotterdamse Rijkweg 126
3042 AS Rotterdam
Telefoon +31 10 4623621



info@modelit.nl
www.modelit.nl

Adviesdienst Verkeer en Vervoer

GSM als informatiebron voor Verkeersbeleid Simulatiestudie

- Uitgebreide samenvatting -

Datum 30 September 2003

Modelit
KvK Rotterdam 24290229



Documentatiepagina

Oprachtgever	Adviesdienst Verkeer en Vervoer
Titel	GSM als informatiebron voor Verkeersbeleid Simulatiestudie - Uitgebreide samenvatting -
Datum	30 September 2003
Projectteam opdrachtgever	drs.ing. J. de Haan, ir. N.S. Schmorak, ir. R.J.Smit
Projectteam Modelit	dr.ir. N.J. van der Zijpp
Projectomschrijving	Het grootschalige gebruik van GSM telefonie in het wegverkeer biedt mogelijkheden om verkeersinformatie af te leiden. Door middel van een simulatiestudie wordt nagegaan wat de mogelijkheden en beperkingen zijn van deze techniek. Om deze simulatiestudie te kunnen uitvoeren is een model voor het simuleren van wegverkeer en telefoniedata gebouwd, en zijn een mapmatch algoritme, reistijdschatter en intensiteitsschatter ontwikkeld.
Trefwoorden	Floating Car Data, GSM, Map matching, Probe-vehicles, Traffic monitoring, Travel time estimation

Inhoud

Samenvatting	5
1 Inleiding	6
2 Opbouw van het simulatiemodel.....	6
2.1 Statisch verkeersmodel	6
2.2 Trajectoriën generator.....	7
2.3 Measurement Report generator.....	7
2.4 Map match algoritme	8
2.5 Reistijd schatter (schakelniveau)	10
2.6 Reistijd schatter (HB-niveau).....	11
2.7 Intensiteitsschatter.....	11
2.8 HB-schatter	12
3 Evaluatiemethode.....	12
3.1 Evaluatie systematiek.....	12
3.2 Methode van beoordeling	13
3.3 Pre-assesment.....	15
4 Resultaten voor het studiegebied Rotterdam.....	17
4.1 Het mapmatchproces.....	17
4.2 De geschatte reistijden op schakel niveau.....	19
4.3 De geschatte reistijden op route niveau	22
4.4 De geschatte intensiteiten	23
4.5 Resultaten voor het studiegebied Breda-Tilburg	24
5 Conclusies.....	25
5.1 Kwaliteit van de ingewonnen verkeersgegevens	25
5.2 Voorbehoud bij de simulatieresultaten	26
5.3 Opgeleverd simulatietool en schattingsalgoritme	27
5.4 Kosten-baten analyse	27
5.5 Rol van de overheid	27
5.6 Verder onderzoek.....	28

Samenvatting

Het inwinnen van verkeersinformatie door gebruik te maken van plaatsbepalingsgegevens die ontleent zijn aan mobiele telefonie lijkt een welkome aanvulling op de huidige manier van dataverzameling die vooral gebruik maakt van wegkantsystemen. Door middel van een simulatiestudie wordt de haalbaarheid van een dergelijk systeem onderzocht.

Een analyse/simulatie model is ontwikkeld en wordt tegelijk met dit rapport opgeleverd. Dit model bestaat uit de volgende onderdelen: een statisch verkeersmodel, een trajectoriën generator, een Measurement Report generator, een mapmatch algoritme, een reistijd schatter; een intensiteitsschatter, een evaluatie module, en een grafische interface.

Als evaluatie systematiek is de Converge methodologie toegepast die in het kader van het Telematics Applications Programme van het 4e kader programma van DGXIII is ontwikkeld. Als onderdeel van deze methodiek zijn een aantal indicatoren bepaald. Deze indicatoren zijn onderverdeeld in: reistijdindicatoren op schakel niveau, reistijdindicatoren op route niveau, intensiteit indicatoren, en technische indicatoren.

Ook zijn circa 60 varianten opgesteld die van elkaar verschillen voor wat betreft parameters die vallen onder de noemers: studiegebied, waarneemperiode, belgedrag, methode van plaatsbepaling en schattingsalgoritme. Vervolgens is geconcludeerd dat de invloed van iedere parameter is terug te voeren op de manier waarop deze parameter de volgende aspecten beïnvloed: de kwaliteit van het mapmatchproces, het aantal ritten dat gematcht wordt, en de representativiteit van de gematchte ritten.

Er zijn twee verschillende studiegebieden doorgerekend: het studiegebied Rotterdam en het studiegebied Breda-Tilburg. Het studiegebied Rotterdam staat model voor netwerken met een hoge netwerkdichtheid en stedelijke omgeving. Het studiegebied Breda-Tilburg kan worden gezien als een voorbeeld van een netwerk met meer landelijke wegen. In dit gebied wordt ook de veldtest Brabantse Wegen Beter Zichtbaar (BWBZ) georganiseerd.

De resultaten geven aan dat het voor het Hoofd Wegen Net (HWN) en de drukbereden gedeeltes van het Onderliggende Wegen Net (OWN) mogelijk moet zijn om betrouwbare reistijdschattingen uit GSM data af te leiden. De resultaten worden echter sterk beïnvloed door de gekozen duur van de studieperiode en de GSM penetratiegraad. Bij een voldoende aantal bellers resulteren ook voor wegen met lagere intensiteiten betrouwbare reistijdschattingen. Het algemene beeld is dat geaggregeerde verkeersprestatieindicatoren, zoals uurgemiddelden over lange trajecten met een grotere relatieve nauwkeurigheid kunnen worden geschat dan de meer gedetailleerde indicatoren, zoals reistijd per kwartier op een specifiek wegvak. Het is met de geïmplementeerde algoritmes niet mogelijk gebleken om betrouwbare schattingen te maken van schakel- of HB intensiteiten. Het grootste probleem bestaat in dit geval uit het bepalen van de geschikte ophoogfactoren. Deze zijn sterk plaatsafhankelijk. Bovendien hangen deze in praktijk ook af van een aantal moeilijk waar te nemen gedragsmatige factoren. Bij het bovenstaande moet worden aangetekend dat er tal van factoren zijn waarmee het model (nog) geen rekening houdt. Als gevolg van deze factoren zullen vooral in de stedelijke omgeving de resultaten in praktijk aanmerkelijk minder goed uitvallen dan het model nu suggereert.

De kosten van een landelijk GSM inwinsystemen geschat op circa 10 m€ bij deelname van één operator. Deze kosten zullen naar verwachting op het onderliggende wegennet, maar buiten de bebouwde kom, moeten worden terugverdiend, omdat hier zowel de kwaliteit van een bruikbaar niveau is en overige verkeersinformatie ontbreekt. Omdat de overheid de belangrijkste afnemer is van verkeersgegevens, is een stimulerende of actieve rol van de overheid een voorwaarde voor realisatie van het beoogde systeem.

1 Inleiding

Traditioneel is de overheid bij het verzamelen van gegevens over het gebruik van het wegennet afhankelijk van data die met lusdetectoren op vaste plaatsen worden verzameld. Het in stand houden en uitbreiden van een dergelijk meetsysteem is kostbaar, en dekt bij de beschikbare middelen niet het hele wegennetwerk af. Systemen voor de inwinning van gegevens op basis van Floating Car Data (FCD) lijken daarom een welkome aanvulling op de bestaande systemen. Het doel van de simulatiestudie die in dit rapport wordt beschreven is om vast te stellen of het inwinnen van verkeersgegevens op basis van GSM-plaatsbepalingsgegevens haalbaar is. Onder 'verkeersgegevens' worden intensiteiten, reistijden/snelheden en HB informatie verstaan. In het kader van deze studie is een simulatiemodel ontwikkeld dat tegelijk met dit rapport wordt opgeleverd. Als onderdeel van dit simulatiemodel is een inwinsysteem voor verkeersgegevens op basis van GSM data ontwikkeld en geïmplementeerd binnen de simulatieomgeving.

Deze samenvatting bestaat uit de volgende gedeeltes:

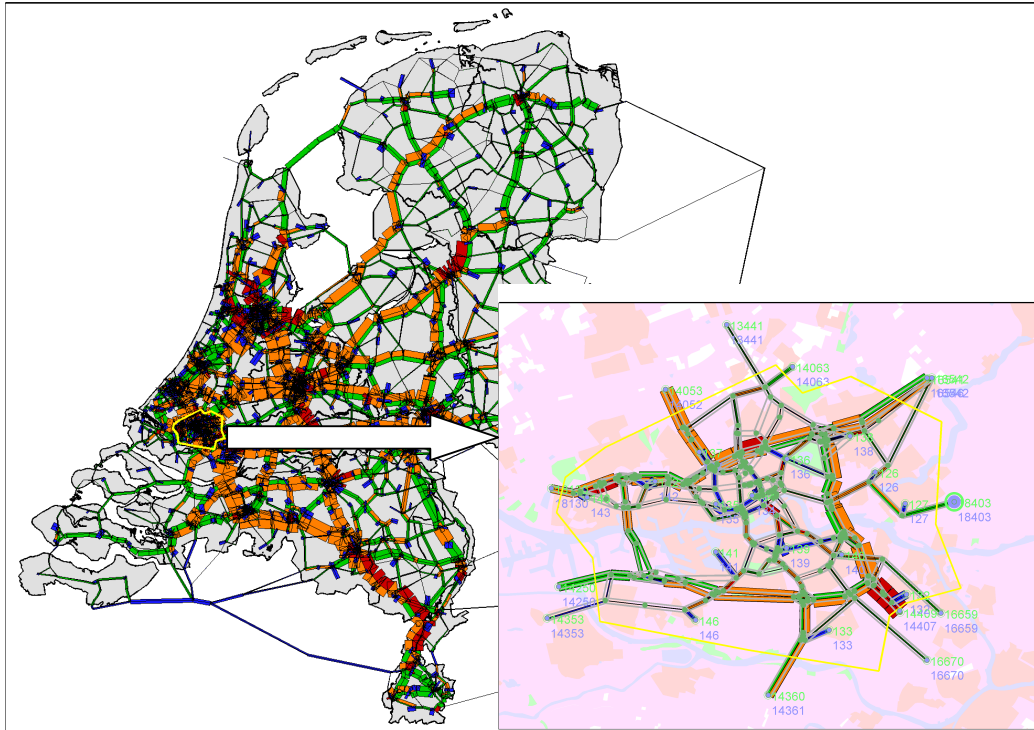
- Een beschrijving van de opzet van het simulatie model en de daarin toegepaste algoritmes (sectie 2) ;
- Een beschrijving van de toegepaste evaluatie en analysemethodiek (sectie 3);
- De resultaten (secties 4 en 4.5);
- De conclusies (sectie 5).

2 Opbouw van het simulatiemodel

2.1 Statisch verkeersmodel

Het uitgangspunt voor de simulatie is een statisch evenwichts-toedelingsmodel met Nederland als toepassingsgebied en een gemiddelde ochtendspits als toepassingsperiode. Dit model berekent uurgemiddelde route- en schakel intensiteiten en uurgemiddelde snelheden. Daarbij wordt consistentie bereikt tussen route-aandelen, schakelintensiteiten, schakelreistijden en route-reistijden. De Herkomst-Bestemmingsmatrix en het netwerk die gebruikt worden zijn met toestemming van TNO Inro ontleend aan het SMART verkeersmodel.

Dit statisch verkeersmodel is in de huidige applicatie gekoppeld aan een zogenaamd uitsnede model. Met behulp van dit uitsnede model kan de berekende landelijke toedeling worden geprojecteerd op iedere willekeurige uitsnede, te definiëren door het tekenen van een contour. Dit maakt het mogelijk om specifieke regio's en situaties als studiegebied te kiezen. In de huidige studie zijn het gebied binnen de ring Rotterdam en het gebied Breda-Tilburg gekozen. Dit laatste gebied komt overeen met het gebied waar de pilotstudie Brabantse Wegen Beter Zichtbaar (BWBZ) wordt uitgevoerd. In deze pilot wordt technologie toegepast die in het kader van Intermezzo wordt ontwikkeld, in het vervolg zal daarom steeds over de Intermezzo veldtest worden gesproken.



Figuur 1: *Met behulp van een statisch evenwichtsmodel en een daaraan gekoppeld uitsnedemodel worden voor het studiegebied de uurgemiddelde belastingen op schakel en routeniveau berekend*

2.2 Trajectoriën generator

Het toedelingsmodel genereert routes en uurgemiddelde routebelastingen. Op basis van deze gegevens worden individuele trajectoriën berekend voor de duur van de waarneemperiode. Uit efficiency overwegingen worden alleen trajectoriën berekend voor probe-ritten, dit zijn ritten waaruit gegevens voortkomen. Het aantal probe-ritten per Herkomst-Bestemmingspaar wordt door loting bepaald op basis van het aantal verplaatsingen en reisduur per HB-paar.

Voor deze probe-ritten wordt eveneens door loting een route bepaald op basis van de vooraf berekende statische route-kansen uit het statische model. Tenslotte worden de passage tijdstippen van iedere knoop op de route berekend door per schakel een reistijd te loten uit een kansverdeling rondom de evenwichtsreistijd.

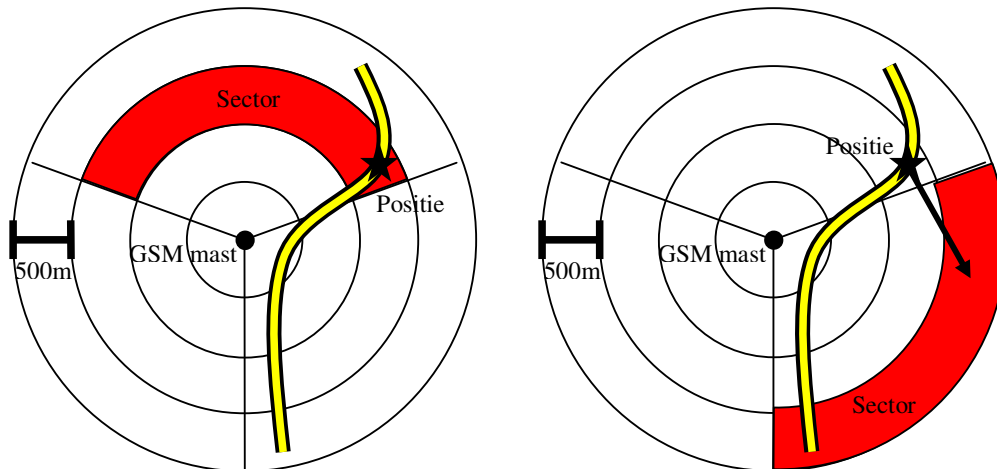
2.3 Measurement Report generator

Voertuigen waaruit gebeld wordt kunnen worden waargenomen doordat de telecom operator continu verschillende parameters moet bijhouden om de verbinding in stand te houden. Zo is bijvoorbeeld in het netwerk bekend via welke antenne en sector een toestel belt en welke Timing Advance van toepassing is. Door deze parameters en eventueel aanvullende parameters met een tussenperiode van bijvoorbeeld 15 seconden op te slaan in zogenaamde Measurement Reports ontstaat een serie waarnemingen op basis waarvan men een schatting kan maken van de route en snelheid van het beschouwde voertuig.

Een Measurement Report bevat naast een herkenningscode voor het bellende toestel en het tijdstip ook een geografische aanduiding. De herkenningscode moet bij voorkeur per toestel en anders per gesprek uniek zijn. De inhoud van de geografische aanduiding hangt af van de waarnemingsmethode die wordt toegepast. De volgende waarnemingsmethodes worden beschouwd:

- De Area-based methode. In dit geval wordt de antenne, sector en Timing Advance parameter opgeslagen. Daarbij kan nog onderscheid gemaakt worden tussen de aanname dat altijd de correcte sector wordt gerapporteerd (AB-exact) en de aanname dat soms de verkeerde sector wordt gerapporteerd (AB-approximate), zie Figuur 2. De Area-based methode is kenmerkend voor wat op dit moment op basis van het GSM systeem met beperkte extra investeringen mogelijk is;
- De Point-based methode. In dit geval wordt de geschatte positie opgeslagen. Dit is kenmerkend voor systemen zoals GPS. Na extra investeringen is een dergelijke plaatsbepalingsmethode ook in combinatie met GSM mogelijk.

Het simulatiemodel genereert de Measurement Reports in een aantal stappen. Eerst worden de gesprekperiodes door loting bepaald. Eventueel kunnen tijdens een rit met een tussenperiode meerdere gesprekken plaatsvinden. Vervolgens worden de tijdstippen bepaald waarop Measurement Reports worden gegenereerd. Vervolgens wordt de inhoud van de Measurement Reports bepaald op basis van de positie waarop het voertuig zich bevindt op het moment dat het Measurement Report wordt gegenereerd.



Area-Based Exact plaatsbepaling: de sector waarin het voertuig zich bevindt wordt gerapporteerd in het Measurement Report

Area-Based Approximate plaatsbepaling: zie Area-based Exact methode, er wordt echter rekening gehouden met een meetfout

Figuur 2: Illustratie van de Area-based Exact en Area-based approximate plaatsbepalingsmethode

2.4 Map match algoritme

Het Mapmatch algoritme koppelt ieder Measurement Report aan een positie op het netwerk. Hiertoe wordt iedere schakel van het netwerk van het netwerk onderverdeeld in korte schakelsegmentjes (zie Figuur 4). Ieder Measurement Report wordt naar het centrum van een dergelijk segment gematcht. Het aantal segmenten per schakel wordt zodanig bepaald dat de afstand tussen twee segment-centroïdes, de zogenaamde 'gap', niet groter is dan een vooraf opgegeven afstand (typisch zo'n 50 meter).

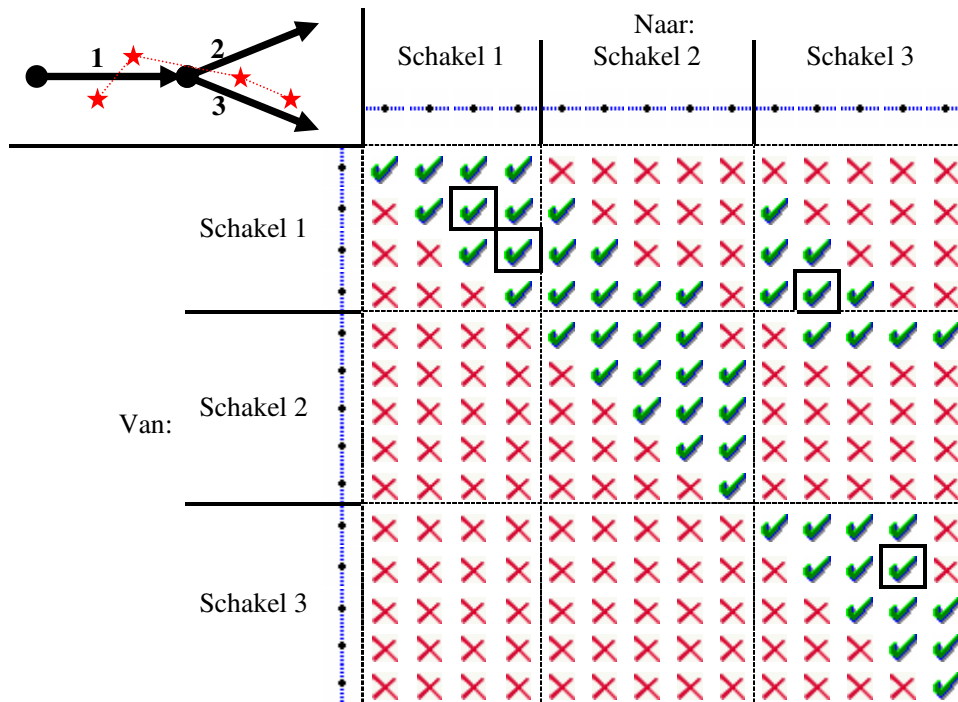
Een karakteristiek van GSM plaatsbepalinggegevens is dat het steeds gaat om betrekkelijk grove plaatsbepaling. In geval van Area-based plaatsbepaling wordt bijvoorbeeld de afstand tot de zendmast in stappen van 500 meter vastgelegd, terwijl de hoek op sectorniveau wordt vastgelegd, waarbij de meeste zendmasten drie sectoren hanteren. In dit geval is dus sprake van een lage *resolutie*. Daarnaast is eventueel ook nog sprake van een *meetfout*, die in dit geval inhoudt dat de verkeerde antenne, sector, of afstand tot de antenne wordt gerapporteerd. Bij de Point-based plaatsbepalingmethode vormt de resolutie geen beperking en is alleen sprake van een meetfout. Een typische waarde voor de standaardafwijking van deze fout is in dit geval 400 meter.

Om op basis van deze gegevens toch zinnige verkeersgegevens af te leiden, kan niet worden volstaan met een algoritme dat ieder Measurement Report simpelweg naar de dichtstbijzijnde positie op het netwerk projecteert, omdat in dit geval waarschijnlijk onzinnige trajecten en onrealistische reistijden resulteren. In plaats daarvan dienen alle Measurement Reports die binnen een rit worden gegenereerd gezamenlijk te worden beschouwd en moet bij het matchen van deze Measurement Reports kennis over mogelijke routes en haalbare snelheden worden toegepast.

Ten behoeve van het mapmatchen is daarom binnen de huidige studie een Maximum Likelihood schatter geïmplementeerd. Dit algoritme kiest de gematchte posities zo dat de gezamenlijke kans van de geïmpliceerde meetfouten minimaal is. Hierbij worden de volgende constraints toegepast:

- Het geïmpliceerde traject tussen twee opeenvolgende gematchte punten moet deel uitmaken van de verzameling vooraf geënumereerde routes. Hierdoor wordt een groot aantal niet realistische routes vooraf uitgesloten;
- De geïmpliceerde snelheden dienen te liggen in een vooraf te definiëren range die per schakel kan worden gedefinieerd. Het minimum van deze range is doorgaans 0 (complete stilstand) en het maximum is typisch 30% hoger dan de free-flow snelheid.

Deze constraints kunnen worden samengevat in een zogenaamde transitie matrix, die voor iedere combinatie van twee schakelsegmenten aangeeft of de transitie tussen deze twee segmenten al dan niet mogelijk is. Voor een voorbeeldnetwerk met drie schakels is de situatie geschetst in Figuur 3. Deze matrix geeft aan welke transities wel, en welke transities niet mogelijk zijn binnen één tijdstap. In de figuur zijn vier cellen omlijnd. Deze vier omlijnde cellen representeren één van de vele mogelijkheden om vier Measurement Reports (aangegeven met rode sterretjes) te matchen zonder de randvoorwaarden te doorkruisen. Het mapmatch algoritme zoekt uit alle toegestane mogelijkheden de meest optimale.



Figuur 3: Toepassing van de transitie matrix op een netwerk met 3 schakels. ✗: transitie is niet mogelijk binnen één tijdstep. ✓: transitie is mogelijk binnen één tijdstep. De vier omliggende hokjes representeren één van de vele mogelijkheden om vier Measurement Reports (★) te matchen.

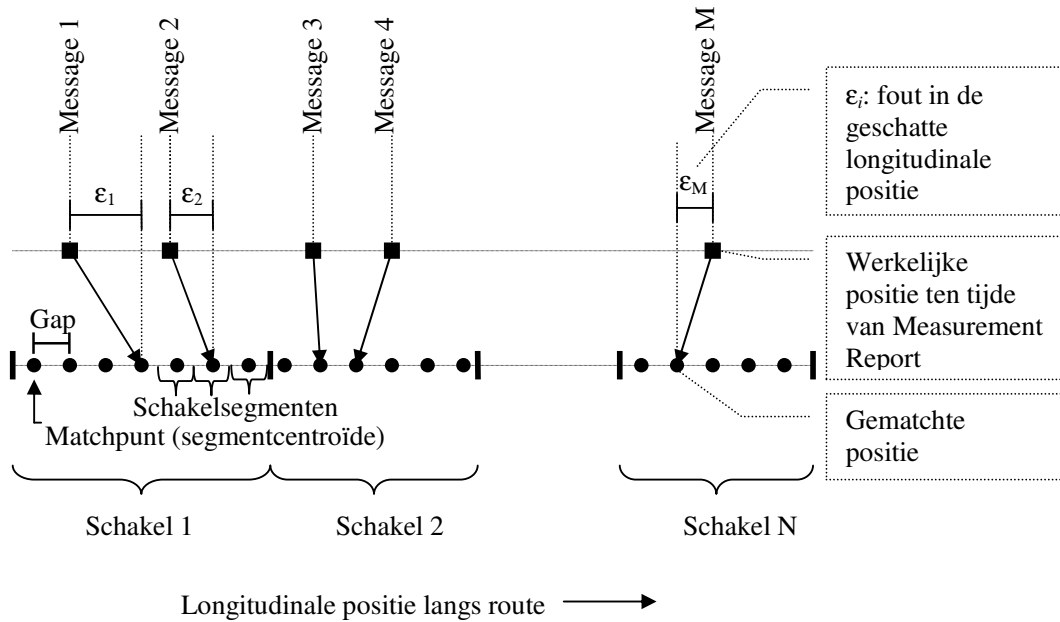
2.5 Reistijd schatter (schakelniveau)

Iedere paar van opeenvolgende gematchte Measurement Reports impliceert dat het tussenliggende traject, bestaande uit een combinatie van schakelsegmenten, in een bepaalde periode door het probe voertuig is afgelegd. Dit maakt het mogelijk om een vergelijking op te stellen met aan de linker kant een lineaire combinatie van schakelreistijden en aan de rechterkant het tijdsverloop tussen de twee Measurement Reports.

Door dit te herhalen voor alle opeenvolgende Measurement Reports ontstaat een uitgebreid stelsel van lineaire vergelijkingen met de schakelreistijden als onbekenden.

Bij elke vergelijking hoort een stoorterm die de mapmatch fouten representeert. Er zijn twee type mapmatch fouten. De ernstigste fout die kan optreden is dat de gematchte punten niet langs de werkelijk gereden route liggen. Deze fout treedt echter met een kleine kans op. Het type fout dat het meest frequent optreedt is dat de Measurement Reports naar de goede route verwijzen, maar naar de verkeerde longitudinale positie. In dit geval wordt er in de vergelijking met teveel of te weinig afgelegde weg rekening gehouden. Dit type fout is serieel gecorreleerd omdat de afgelegde weg waarmee in de ene vergelijking te weinig wordt gerekend in de volgende vergelijking als teveel wordt meegeteld (zie Figuur 4).

Door de seriële correlatie tussen de bovengenoemde stoortermen correct te modelleren, kunnen de waarnemingen die uit de gematchte Measurement Reports zijn afgeleid optimaal worden verwerkt. Dit gebeurt met behulp van een Kalman filter.



Figuur 4: Elke combinatie van twee opeenvolgende Measurement Reports levert één meetvergelijking op. De meetfout in deze meetvergelijking bestaat uit twee plaatsbepalinggerelateerde componenten. Opeenvolgende meetfouten zijn negatief gecorreleerd, doordat ze een van de twee componenten (op het teken na) gemeenschappelijk hebben.

2.6 Reistijd schatter (HB-niveau)

De reistijd tussen herkomsten en bestemmingen wordt berekend als de gewogen som van de route-reistijden tussen deze herkomst en bestemming, waarbij de weegfactoren overeen komen met de route-aandelen zoals in de evenwichtstoedeling berekend. De route-reistijden worden berekend door voor iedere route de schakel-reistijden over de gebruikte schakels te sommeren.

2.7 Intensiteitsschatter

Met het uitvoeren van de mapmatching worden ritten toegedeeld aan de trajecten tussen de gematchte punten. Hieruit kunnen 'pseudo teldata' worden afgeleid. In deze pseudo teldata wordt echter slechts een klein gedeelte van de voertuigen meegenomen in de zin dat deze voertuigen Measurement Reports genereren die vervolgens aan een route worden gekoppeld waarvan de beschouwde wegsectie deel uitmaakt. De 'pseudo teldata' moeten daarom worden opgehoogd met een nog te bepalen ophoogfactor. Bij de bepaling van de ophoogfactor dient de kans berekend te worden dat aan de volgende voorwaarden voldaan is:

- het voertuig moet een GSM aan boord hebben;
- er moet gebeld worden in de periode dat de wegsectie wordt gepasseerd;
- de gegenereerde Measurement Reports moeten worden geclassificeerd als 'niet-stationaire' beller;
- bij het matchen van de Measurement Reports moeten deze aan de goede wegsectie worden gekoppeld.

Los hiervan kan het ook voorkomen dat ten onrechte voertuigen worden waargenomen, bijvoorbeeld doordat Measurement Reports die op een andere wegsectie zijn gegenereerd ten onrechte naar de beschouwde wegsectie worden gematcht.

De ophoogfactor wordt berekend als het product van de kans dat een GSM aan boord is (P[GSM aan boord]), de kans dat gebeld wordt in de beschouwde wegsectie (P[bellen in wegsectie]), en de kans dat het gesprek als niet-stationair wordt herkend (P[classificeren als niet stationair]). Deze factoren kunnen gedeeltelijk op basis van theoretische overwegingen worden bepaald en voor een ander gedeelte direct worden waargenomen in de praktijk, zie Tabel 1. De kans dat het gesprek als niet-stationair wordt herkend is echter op schakelniveau niet direct waarneembaar en evenmin gemakkelijk theoretisch af te leiden. In de huidige studie wordt daarom uitgegaan van de aanname dat deze kans voor alle wegsecties gelijk is, en wordt deze kans gelijk gesteld aan het overall percentage. In de praktijk zou dit overall percentage moeten worden afgeleid door het model te kalibreren op aantal schakels waarvoor telgegevens beschikbaar zijn.

Tabel 1: *Overzicht van de invloed, waarneembaarheid en afleidbaarheid van ophoogfactoren*

Factor	Variatie over wegsecties	Direct waarneembaar	Theoretisch afleidbaar
P[GSM aan boord]	Gering	Ja	Nee
P[bellen in wegsectie]	Groot	Nee	Ja
P[classificeren als niet-stationair]	Groot	Nee	Nee

2.8 HB-schatter

Het schatten van HB-relaties is aanmerkelijk complexer dan het schatten van schakel intensiteiten. Voordat tot het schatten van HB-relaties relaties wordt overgegaan zou daarom eerst de haalbaarheid van het schatten van schakelintensiteiten moeten zijn aangetoond. Na inspectie van de eerste resultaten van de intensiteitsschatter is besloten om geen verder onderzoek naar de HB-schatter uit te voeren.

3 Evaluatiemethode

3.1 Evaluatie systematiek

Door middel van het uitvoeren van een simulatiestudie wordt in de huidige studie nagegaan wat de toepasbaarheid is van een inwinsysteem voor verkeersgegevens dat is gebaseerd op GSM data. Als evaluatiemethode wordt hierbij de systematiek toegepast zoals beschreven in het rapport “Guidebook for Assessment of Transport Telematics Applications” (Zhang, et al., 1998). Deze systematiek komt voort uit het Converge project dat is uitgevoerd als onderdeel van het Telematics Applications Programme van het 4e kader programma van DGXIII (1994-1998). Converge had als doel om een algemeen toepasbare evaluatie methodologie te definiëren voor de telematica projecten die destijds in het kader van het 4e kader programma (inmiddels opgevolgd door het 5e en 6e kaderprogramma) werden uitgevoerd, en bied daarom ook voor de huidige studie een goed houvast.

De Converge methodologie bestaat uit de volgende stappen:

- Definitie van ‘user needs’. Wie zijn de gebruikers en belanghebbenden en wat zijn hun voornaamste behoeftes of belangen?
- Beschrijving van het te evalueren systeem. Welke toepassingen worden geëvalueerd en wat is hun technische, functionele en ruimtelijke bereik?
- Assessment objectives. Definitie van de aspecten op basis waarvan de beoordeling plaatsvindt.
- Pre-assessment of expected impacts. Wat zijn de verwachtingen vooraf van het systeem? Welke hypothesen kunnen vooraf worden geformuleerd?
- Methode van beoordeling. Hoe ziet de experimentele opzet eruit?
- Data analyse. Statistische richtlijnen voor het uitvoeren van de data analyse;
- Rapportage. Richtlijnen voor het rapporteren van de resultaten.

3.2 Methode van beoordeling

In de huidige studie wordt van deze stappen de methode van beoordeling het meest uitgebreid behandeld. Deze stap wordt binnen de Converge methodologie verder opgedeeld in de volgende deelstappen:

- Het definiëren van indicatoren. De indicatoren zijn onderverdeeld in:
 - reistijdindicatoren op schakel niveau;
 - reistijdindicatoren op route niveau;
 - intensiteit indicatoren;
 - technische indicatoren;
- Het definiëren van een referentie schattingsmethode. Om na te gaan of een nieuwe methode iets toevoegt aan reeds bestaande of goedkoop te realiseren alternatieven worden de reistijd indicatoren ook berekend voor een referentiemethode die eruit bestaat dat de free-flow snelheid als schatting wordt gehanteerd. Voor de intensiteitschatter bestaat geen voor de hand liggende referentiemethode;
- Een plan voor dataverzameling (welke parameters worden gemeten?). Omdat het hier een simulatiestudie betreft, gelden hier niet de gebruikelijke beperkingen die gelden bij een veldtest. Elke parameter die deel uitmaakt van het simulatiemodel en van belang is kan worden vastgelegd;
- Het definiëren van varianten. Het doorrekenen van verschillende varianten geeft inzicht in de invloed van de verschillende parameters op de performance van het systeem en fungeert ook als een controle op de juistheid van de toegepaste algoritmes; De te variëren parameters kunnen worden onderverdeeld in verschillende categorieën.

De onderstaande tabel geeft een overzicht van alle parameters die invloed hebben. Niet in alle gevallen is het overigens mogelijk om deze parameters apart te variëren.

Tabel 2: Voornaamste invloedsfactoren op de performance van een GSM gebaseerd inwinsysteem voor verkeersgegevens

Categorie	Parameter
Studiegebied	<ul style="list-style-type: none"> • Netwerk dichtheid (weglengte /km²)
Waarneemperiode	<ul style="list-style-type: none"> • Duur van de waarneemperiode, bijvoorbeeld 60 minuten
Verkeerssysteem	<ul style="list-style-type: none"> • Gemiddelde schakelintensiteit • Gemiddelde ritduur • Spreiding in de snelheden
Belgedrag	<ul style="list-style-type: none"> • Penetratie van GSM toestellen/carkits • Gemiddelde gespreksduur • Gemiddelde gespreksfrequentie
Methode van waarnemen	<ul style="list-style-type: none"> • Nauwkeurigheid van plaatsbepaling (standaardafwijking) • Methode van plaatsbepaling (Area-based of Point-based) • Measurement Report update frequentie
Schattingsalgoritme	<ul style="list-style-type: none"> • Maximale plaatsbepalingsfout waarmee rekening wordt gehouden • Toegepaste discretisatiestap bij mapmatching • Minimale snelheid waarmee rekening wordt gehouden (relatief aan free-flow speed) • Maximale snelheid waarmee rekening wordt gehouden (relatief aan free-flow speed) • Maximale snelheid waarmee rekening wordt gehouden (in absolute zin) • Criterium voor het selecteren van de niet-stationaire ritten (Area-based methode) • Criterium voor het selecteren van de niet-stationaire ritten (Point-based methode) • Vereist aantal waargenomen schakel passages voordat een reistijd schatting wordt afgegeven

Uiteindelijk resulteert dit in een totaal van 52 varianten, die als volgt zijn onderverdeeld:

Gebied	Methode	
	Area-based	Point-based
Breda-Tilburg	breda_000 (basisvariant) breda_001 tm Breda_012 (varianten)	breda_013 (basisvariant) breda_014 tm Breda_025 (varianten)
Rotterdam	rot_000 (basisvariant) rot_001 tm rot_012 (varianten)	rot_013 (basisvariant) rot_014 tm rot_025 (varianten)

- Het kiezen van een analyse methode. In een aparte sectie wordt ingegaan op de interpretatie van de indicatoren. Het probleem met vele (reistijd) indicatoren is dat deze ofwel geen eenvoudige intuïtieve betekenis hebben, ofwel dat de waarde van deze indicatoren net zoveel afhangt van netwerkeigenschappen zoals gemiddelde schakellengte, als van de performance van het systeem. Als compromis wordt gekozen voor de indicator: “kwadratische fout per minuut reistijd”. Deze indicator is, althans in theorie, onafhankelijk van de gemiddelde schakel lengte en snelheid. Aan de hand van enkele voorbeelden wordt 0.1 (10%) als drempelwaarde aangewezen.
- Het opstellen van een meetplan (waar en wanneer wordt gemeten?). Ook in dit geval zorgt het feit dat simulatie als evaluatiemethodiek wordt toegepast ervoor dat de gebruikelijke beperkingen die voor een veldtest gelden niet van toepassing zijn. Metingen worden direct uit het simulatiemodel afgeleid en zijn beschikbaar voor alle schakels en de gehele waarneemperiode (die uit een deel van de ochtendspits bestaat).

Binnen het Intermezzo project dat parallel wordt uitgevoerd gelden de gebruikelijke beperkingen uiteraard wel, en zijn daarom een aantal controle trajecten gedefinieerd waarop metingen zullen worden verzameld. Om de vergelijkbaarheid van beide studies te bevorderen zijn deze controle trajecten ook in de huidige studie gebruikt.

3.3 Pre-assesment

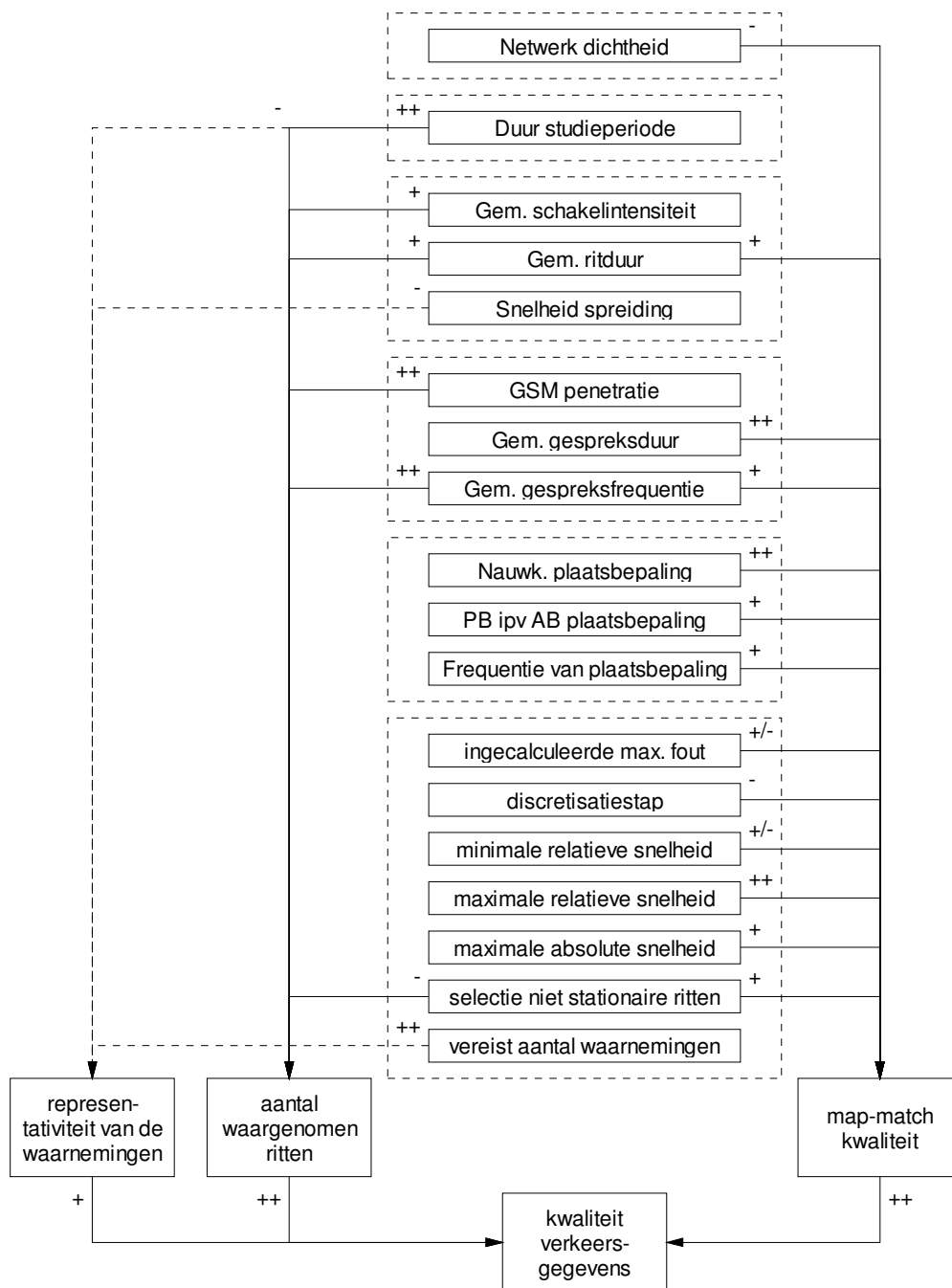
Het is een goede gewoonte om voorafgaande aan het uitvoeren van experimenten, of in ieder geval zonder gebruik te maken van de uitkomsten van deze experimenten, hypothesen op te stellen over de werking van het te implementeren systeem. Zo mogelijk dienen ook uitspraken gedaan te worden over de te verwachten resultaten. Deze analyse vooraf leidt tot een betere opzet van het experiment en helpt bij het kritisch analyseren van de uitkomsten ervan.

Een belangrijke constatering bij de analyse van het GSM gebaseerde inwinsysteem voor verkeersgegevens zoals geschetst in sectie 1 is het feit dat alle ritten onafhankelijk van elkaar naar het netwerk worden gematcht. Op basis van deze gematchte ritten wordt een stelsel van vergelijkingen opgesteld waaruit de schakelsnelheden worden opgelost. De intensiteiten worden geschat door een ophoogfactor toe te passen op het aantal waargenomen schakel passages. De kwaliteit van het eindresultaat hangt daarom in theorie alleen af van:

- De kwaliteit van het mapmatchproces, uit te drukken in de fout in de geschatte locatie na mapmatching;
- Het aantal ritten dat gematcht wordt;
- De representativiteit van de gematchte ritten voor de gemiddelde verkeerssituatie, uit te drukken als de spreiding in reistijd gedeeld door de gemiddelde reistijd.

Door vervolgens na te gaan wat de invloed van de in Tabel 2 genoemde parameters is op deze drie afgeleide grootheden is het mogelijk om uitspraken te doen over de invloed van iedere parameter op het eindresultaat (tenzij één parameter meerdere afgeleide grootheden met tegengesteld teken beïnvloed). Deze aanpak is uitgewerkt in Figuur 5. In absolute zin kunnen in dit stadium echter nog geen uitspraken worden gedaan.

Deze aanpak kan wellicht ook houvast bieden bij het dimensioneren van een in de praktijk toe te passen inwinsysteem op basis van probes. Het aantal invloedsfactoren is immers teruggebracht van een groot aantal, moeilijk kwantificeerbare parameters, tot een drietal goed kwantificeerbare parameters. Idealiter zou zelfs een functioneel verband kunnen worden opgesteld tussen deze drie parameters en de kwaliteit van de resulterende verkeersgegevens, bijvoorbeeld uit te drukken in de hierboven beschreven variantie per minuut reistijd. Het zoeken naar een dergelijke uitdrukking valt echter buiten de scope van de huidige studie.



Figuur 5: Subjectieve analyse van de gevoeligheid van de kwaliteit van de afgeleide verkeersgegevens voor de verschillende systeemparementers. ++: sterke veronderstelde positieve correlatie; +: veronderstelde positieve correlatie; -: veronderstelde negatieve correlatie

4 Resultaten voor het studiegebied Rotterdam

Het studiegebied Rotterdam staat model voor netwerken met een hoge netwerkdichtheid en stedelijke omgeving. Dit brengt een hoog risico met zich mee dat Measurement Reports naar de verkeerde route worden gematcht. Bovendien zijn de snelheden op een groot aantal schakels laag, waardoor het moeilijker is om stationaire en niet stationaire bellers van elkaar te onderscheiden. Daar staat tegenover dat de intensiteiten in een stedelijke omgeving hoger zijn dan in een niet verstedelijkte omgeving.

Voor het netwerk Rotterdam zijn 26 varianten doorgerekend en daarnaast nog 6 in verband met een aangepaste penetratiegraad en waarneemperiode.

4.1 Het mapmatchproces

Na het uitvoeren van een aantal testen, gevolgd door noodzakelijke aanpassingen, rekt het mapmatch algoritme alle 32 varianten (in totaal ongeveer 500.000 gesimuleerde ritten) door zonder foutmeldingen te genereren. Hiermee voldoet de methode aan de eisen met betrekking tot verificatie, in de zin dat de methode de invoer accepteert en uitkomsten in het goede formaat produceert, zonder bijvoorbeeld halverwege te crashen.

Alle gematchte ritten kunnen visueel worden geïnspecteerd. De meerderheid van de ritten wordt op correcte manier gematcht, in de zin dat de rit naar de goede route wordt gematcht en dat de gematchte longitudinale positie van de Measurement Reports niet ernstig afwijkt van de positie waarop deze Measurement Reports zijn gegenereerd door het bellende voertuig. Tabel 3 vat een aantal kentallen van het mapmatchproces samen. Deze kentallen zijn ontleend aan de basisvarianten voor de Area-based methode en Point-based methode voor het studiegebied Rotterdam.

De relatieve residuele locatie fout geeft aan in hoeverre de fout in de geschatte locatie af is genomen na het mapmatchen ten opzichte van de gerapporteerde locatie en is een goede maat voor de kwaliteit van dit proces. Omdat bij de Area-based methode geen locatie wordt gerapporteerd is deze indicator alleen voor de Point-based methode te berekenen. De resultaten geven aan dat na het uitvoeren van het map matchen de locatiefout met een factor vier is afgenomen.

Tussen de 72 en 80 % van de Measurement Reports wordt naar de schakel gematcht waarop ze ook gegenereerd zijn. Tussen de 70 en 75 % van de ritten wordt naar een traject gematcht dat volledig deel uitmaakt van de gereden route. Voor de Point-based methode liggen de percentages iets gunstiger, maar daar staat tegenover dat slechts 33% van de ritten als niet stationair worden aangemerkt.

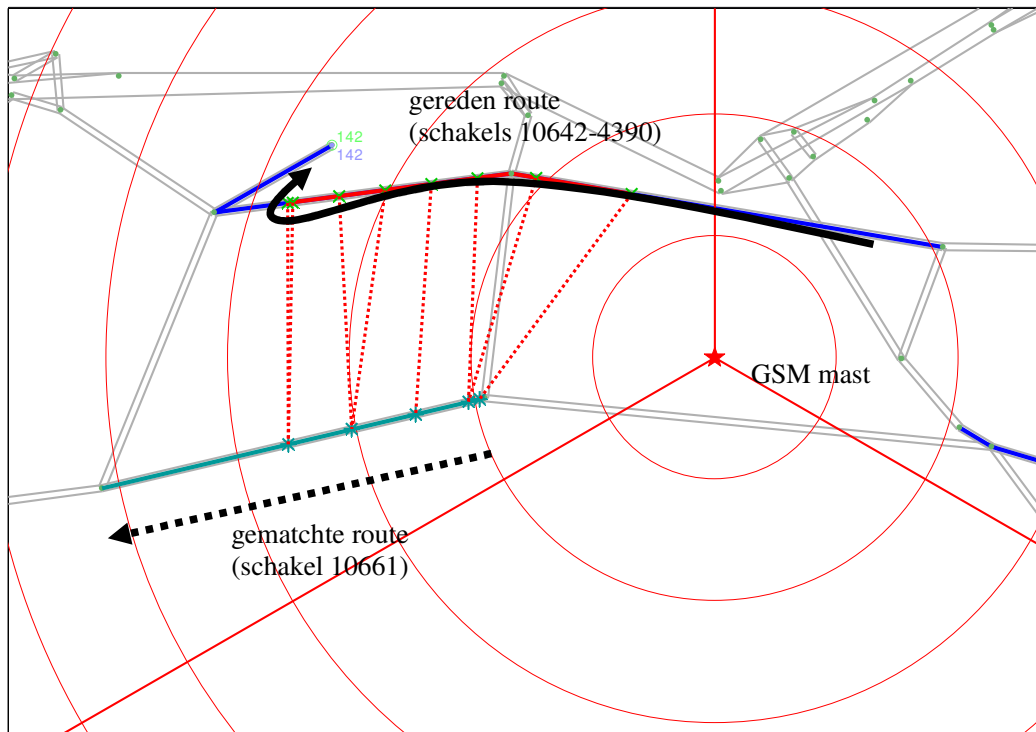
Tabel 3: Kentallen van het mapmatch proces (case Rotterdam)

	Area-based Methode (variant R_000)	Point-based Methode (variant R_013)
relatieve residuele locatie fout	-	0.24
% niet stationair	48.03	33.52
% MR naar correcte schakel	72.28	79.18
% ritten naar correcte route	70.18	74.49

De getallen in Tabel 3 hangen af van de karakteristieken van het netwerk, het belverkeer, de waarnemingsmethode en het mapmatch algoritme, maar zijn niet afhankelijk van het *aantal* te matchen ritten. Hoe beter deze indicatoren uitvallen, hoe beter de reistijd en intensiteitsschatting zullen zijn. Het aantal gematchte ritten is echter de andere belangrijke factor in de kwaliteit van de reistijd en intensiteitsschatting. Aannemende dat alle parameters vast liggen, met uitzondering van de waarneemperiode, zou men deze laatste moeten aanpassen aan de vereiste nauwkeurigheid van de schatting. De getallen in Tabel 3 bepalen in belangrijke mate hoe deze trade-off uitvalt.

Alhoewel het mapmatchen in de meeste gevallen goed verloopt zijn er ook een aantal typische situaties aan te geven waarin fouten optreden.

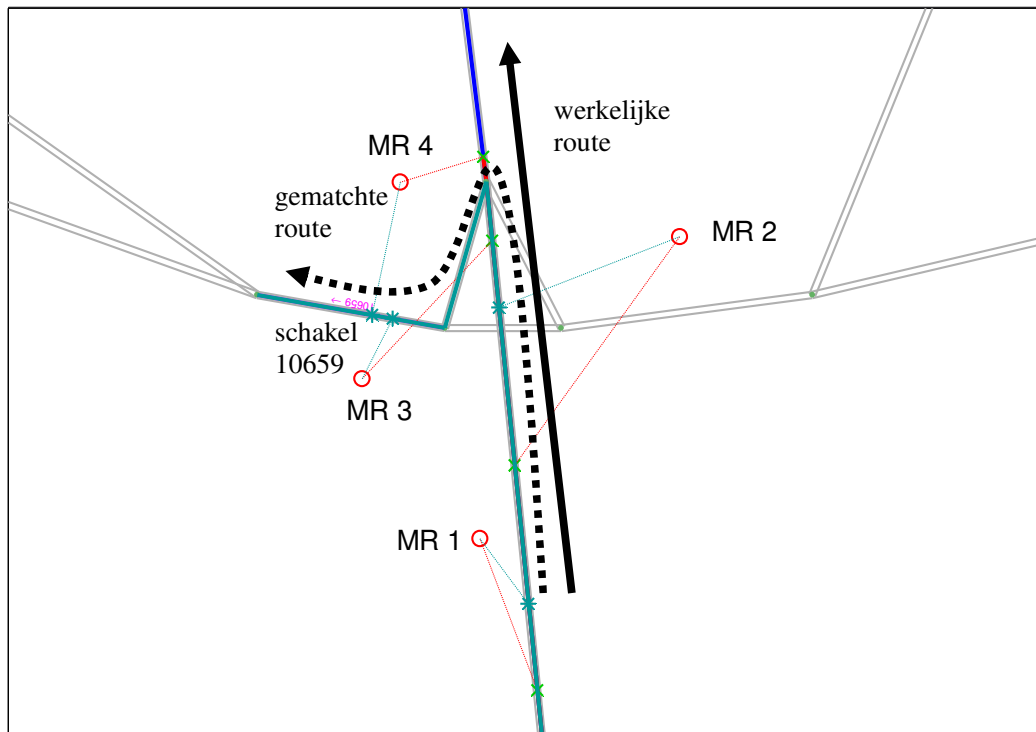
Een voorbeeld hiervan is in Figuur 6 te zien. In dit figuur is te zien dat er twee routes zijn, die op basis van de gerapporteerde sectoren niet van elkaar kunnen worden onderscheiden. Dit is een typische kwaal die kan optreden bij Area-based matching. Omdat als eis wordt toegepast dat een rit in tenminste in drie sectoren is waargenomen voordat deze als niet stationair wordt gekwalificeerd, is het aantal situaties waarin een fout zoals in Figuur 6 optreedt redelijk beperkt. Merk bijvoorbeeld op dat de fout niet zou zijn opgetreden indien het gesprek iets eerder zou zijn begonnen.



Figuur 6: Twee trajecten die op basis van Area-based matching niet van elkaar te onderscheiden zijn. De groene x symbolen representeren de werkelijke positie. De blauwe * symbolen representeren de gematchte positie.

Een ander voorbeeld is te zien in Figuur 7. Dit is een illustratie van Point-based map matching. Door een ongelukkige samenloop van omstandigheden suggereren de laatste twee Measurement Reports dat het voertuig de afslag heeft genomen terwijl dit niet zo is. Ook nu geldt dat de fout waarschijnlijk niet zou zijn opgetreden indien het gesprek iets langer of iets korter zou hebben geduurd.

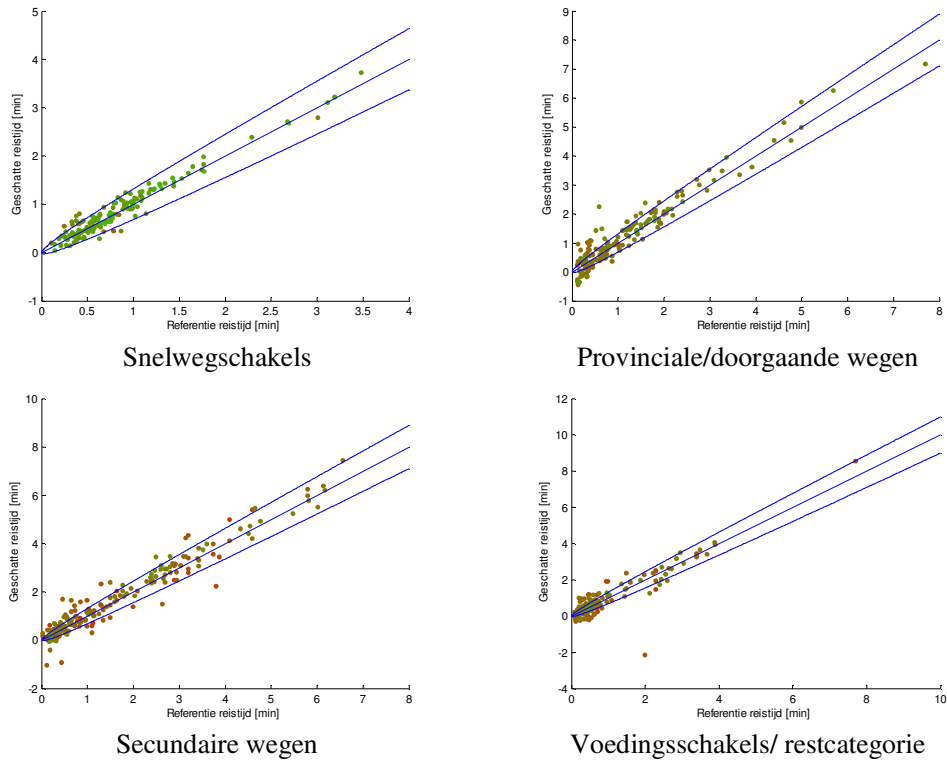
Mapmatch fouten zoals te zien in Figuur 6 en Figuur 7 hebben over het algemeen niet zo veel invloed omdat tegenover iedere foutief gematchte rit vele correct gematchte ritten staan. Dit gaat echter niet altijd op. Wanneer een schakel van zichzelf een lage belasting heeft worden op deze schakels weinig Measurement Reports gegenereerd. In dit geval kan het voorkomen dat het grootste deel van de Measurement Reports die naar deze schakel gematcht worden op een andere schakel zijn gegenereerd, waardoor grote fouten ontstaan. Een typisch voorbeeld van een dergelijke situatie is als een snelweg een parallel weg heeft.



Figuur 7: Voorbeeld van een typische fout bij Point-based map matching. De laatste twee Measurement Reports suggereren dat het voertuig de afslag heeft genomen terwijl dit niet zo is.

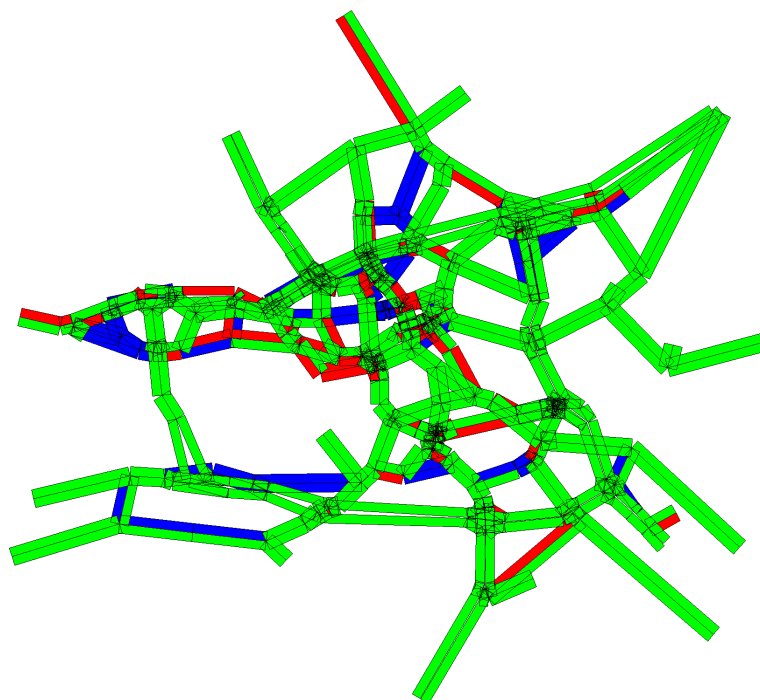
4.2 De geschatte reistijden op schakel niveau

De onderstaande grafieken geven de geschatte versus de werkelijke reistijd weer voor het studiegebied Rotterdam en de Area-based methode. In deze grafieken representeert ieder punt een schakel. Met behulp van een kleurcode die varieert van groen naar rood is het aantal ritten weergegeven waarop de schatting is gebaseerd (groen → veel waarnemingen, rood → enkele waarnemingen)



Figuur 8: *Geschatte versus waargenomen reistijden op schakelniveau, uitgesplitst naar wegtype*

Uit de grafieken blijkt dat voor de categorie ‘snelwegen’ de schattingen uitstekend zijn en dat de kwaliteit iets afneemt naarmate dat de orde van de weg afneemt. De blauwe parabolen komen overeen met de punten waarin de kwadratische fout per minuut reistijd 10% bedraagt. Dit wordt als een acceptabele waarde gezien. Wanneer we alle schakels die buiten de parabool vallen afkeuren, en alle schakels waarvoor onvoldoende waarnemingen beschikbaar zijn beschouwen als niet waargenomen, kan de situatie voor het studiegebied Rotterdam grafisch worden weergegeven. Dit is te zien in Figuur 9.



Figuur 9: *Overzicht van de geaccepteerde(groen), verworpen (rood) en niet waargenomen schakels(blauw) voor de basisvariant van het studiegebied Rotterdam. GSM penetratie: 80%.*

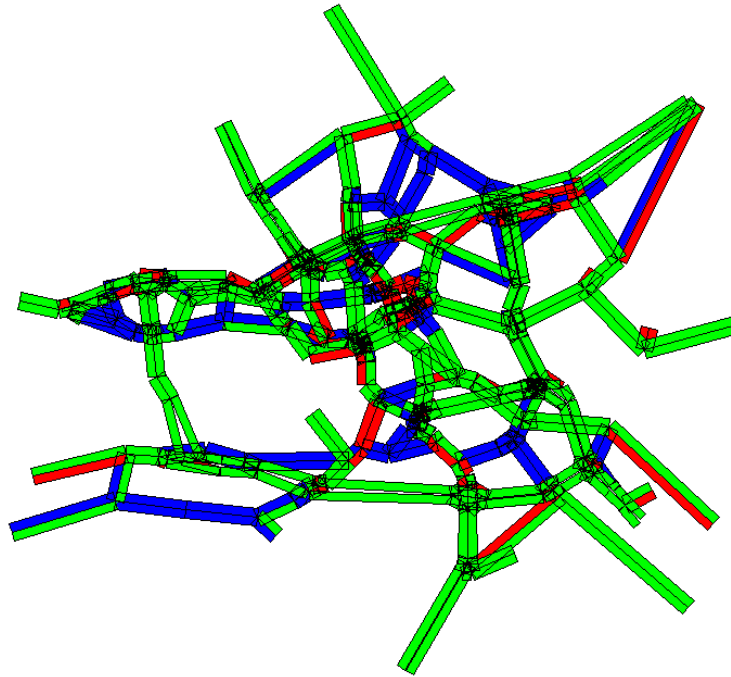
Figuur 9 representeert echter voor wat betreft penetratiegraad wel het meest optimistische scenario omdat is uitgegaan van een aandeel van 80% automobilisten met een GSM in de stand-by stand aan boord. Wanneer we uit gaan van de situatie waarin slechts één telecom aanbieder GSM data beschikbaar stelt, en we rekenen deze aanbieder een marktaandeel van 20% toe, dan daalt de penetratie tot 16%. In dit geval verdubbelt het aantal niet waargenomen schakels van 25 tot 50% en neemt de gemiddelde kwadratische fout per minuut reistijd toe van 10 tot 14%. Het resultaat is te zien in Figuur 10.

In beide gevallen kan de reistijd op de snelwegen met een voldoende nauwkeurigheid worden bepaald. Voor het onderliggende netwerk geldt dat de resultaten in belangrijke mate van de GSM penetratie afhangen.

Figuur 9 en Figuur 10 representeren de basisvariant. De kwaliteit van de reistijdschatting neemt toe of af wanneer met de parameters wordt gevarieerd. De Point-based methode levert, niet onverwacht, over de hele linie betere schattingsresultaten op dan de Area-based methode. De Area-based methode blijkt met name gevoelig voor de nauwkeurigheid waarmee de correcte area wordt gerapporteerd. Het huidige mapmatch algoritme gaat er van uit dat het Measurement Report altijd de correcte ‘area’ rapporteert. Wanneer dit niet gebeurt, lopen de fouten sterk op. Verreweg de duidelijkste verbetering van de reistijd schatting is te zien wanneer extra voorwaarden worden opgelegd met betrekking tot de minimum snelheid. Hierdoor worden alle transities uitgesloten die impliceren dat de snelheid onder 50% van de free-flow snelheid daalt. In het gesimuleerde netwerk komen dergelijke situaties nauwelijks voor, maar in de praktijk kan het waarnemen van dergelijke situaties echter juist een expliciet doel zijn.

Omdat lage intensiteit en sterk gereduceerde snelheden meestal niet samen optreden ligt een verbetering van het mapmatch algoritme voor de hand, waarbij men de te hanteren minimum

snelheid laat afhangen van het aantal waargenomen voertuigen. In dit geval wordt voor snelwegen geen minimum snelheid toegepast en op schaars bereden wegen wel. Deze variant van de methode is echter niet in de huidige studie beproefd

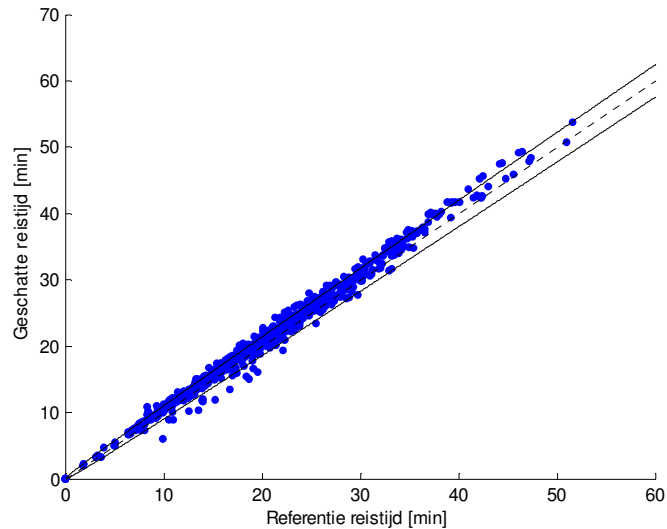


Figuur 10: *Idem als vorig figuur. GSM penetratie: 16%.*

4.3 De geschatte reistijden op route niveau

De reistijden op routeniveau worden getoond in Figuur 11. Iedere stip in dit figuur representeert één relatie. Ook relaties die nauwelijks verkeer genereren zijn in het figuur opgenomen (en zijn verantwoordelijk voor de grootste afwijkingen). Omdat schattingsfouten die op schakelniveau optreden elkaar op routeniveau gedeeltelijk compenseren zijn de relatieve fouten op routeniveau betrekkelijk gering.

Er lijkt sprake te zijn van een lichte bias naar boven. Dit zou kunnen samenhangen met het feit dat ritten waarop langzaam gereden wordt, een iets grotere kans met zich meebrengen dat tijdens de rit twee keer gebeld wordt. Dit zou kunnen leiden tot een oververtegenwoordiging van langzame ritten in de steekproef. Ook de kans dat ten minste één maal gebeld wordt is voor langzame ritten groter, maar de opzet van het simulatie model houdt hier geen rekening mee.



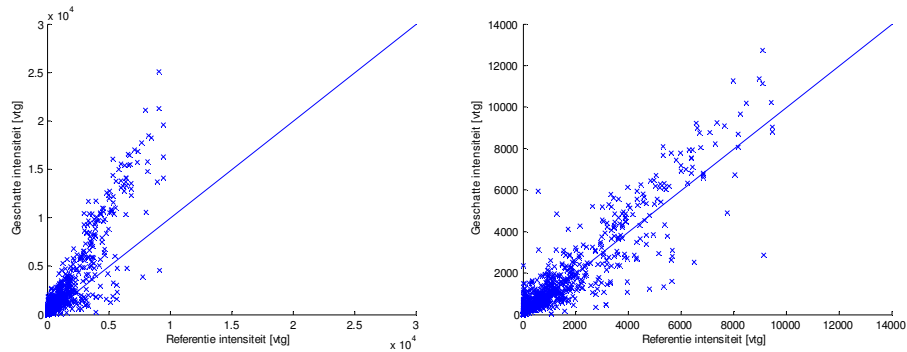
Figuur 11: *Geschatte versus waargenomen reistijd voor de basis variant van het studiegebied Rotterdam (Area-based methode)*

4.4 De geschatte intensiteiten

Het matchen van Measurement Reports naar het netwerk impliceert een aantal ‘virtuele’ tellingen. Deze tellingen worden opgehoogd met ophoogfactoren zoals besproken in sectie 2.7. De opgehoogde tellingen kunnen vervolgens weer worden uitgezet tegen de werkelijke intensiteiten. Het resultaat is te zien in Figuur 12, links. Uit dit figuur blijkt dat de intensiteit behoorlijk wordt overschat. Na enig experimenteren blijkt dat de vertekening alleen optreedt indien een selectie ten behoeve van het identificeren van niet-stationaire bellers wordt toegepast. Wanneer deze selectie achterwege blijft treedt de systematische vertekening niet op, zie Figuur 12, rechts.

Het negeren van bellers die in minder dan het vereiste aantal sectoren zijn waargenomen wordt in de huidige schatter door een enkele ophoogfactor gecompenseerd. Deze ophoogfactor wordt berekend als het totaal aantal bellers gedeeld door het geaccepteerde aantal bellers (met andere woorden: het aantal bellers dat als niet-stationair is aangemerkt). Omdat alleen bellers worden geaccepteerd die in meerdere sectoren zijn waargenomen wordt echter een steekproef gecreëerd waarin de lange ritten oververtegenwoordigd zijn. Omdat lange ritten op meer schakels bijdragen aan de intensiteit ontstaat een overschatting.

Voor schakels waar stationaire bellers geen rol spelen zou men de intensiteiten kunnen schatten op basis van alle ritten. Ook in dit geval blijven de resultaten echter matig tot slecht. Hier komt nog bij dat in de huidige studie verstoringseffecten, zoals fluctuaties in belduur en belfrequentie niet zijn meegenomen. Ook is de invloed van de verkeerssituatie op het belgedrag niet gemodelleerd, terwijl in de praktijk vermoedelijk een correlatie bestaat tussen congestie en belgedrag. Wellicht zijn verdere modelverbeteringen nog mogelijk, maar voorlopig is de conclusie dat GSM data zich niet lenen voor het schatten van intensiteiten, ongeacht het wegtype.



Figuur 12: *Geschatte versus waargenomen intensiteiten. Basis variant, Area-based matching, Rotterdam. Links: Er zijn alleen ritten gematcht die in tenminste drie sectoren zijn waargenomen. Rechts: alle ritten zijn gematcht.*

4.5 Resultaten voor het studiegebied Breda-Tilburg

Het studiegebied Breda-Tilburg beslaat een grotere oppervlakte (1094 km²) dan het studiegebied Rotterdam (364 km²). Ook is de totale lengte van de schakels (819 km) groter dan die voor het studiegebied Rotterdam (669 km). Toch is het totaal aantal ritten kleiner (66739 tegen 91107). Breda-Tilburg kan daarom gezien als een voorbeeld van een landelijk netwerk en Rotterdam als een stedelijk netwerk.

Voor dit studiegebied zijn dezelfde varianten doorgerekend als voor het studiegebied Rotterdam. Het algemene beeld is dat de resultaten ondanks de verschillende schaalgrootte van de gebieden voor een groot deel overeen komen. In het studie gebied Breda-Tilburg komen enkele schakels voor waaraan het toedelingsmodel extreem lage belastingen toekent. Voor deze schakels valt de reistijdschatting zeer slecht uit, omdat de gematchte Measurement Reports voor deze schakels alle ten onrechte zijn toegekend aan deze schakels. Voor de overige schakels valt de reistijdschatting licht beter uit dan de reistijd schatting voor het netwerk Rotterdam. Dit komt vooral tot uiting in het percentage niet waargenomen schakels (17% in de Area-based basisvariant tegenover 25% voor Rotterdam), het percentage Measurement Reports gematcht naar de goede schakel (88% versus 72%) en het percentage ritten gematcht naar de goede route (86% versus 70%).

In het kader van de huidige studie zijn ook een aantal van de trajecten doorgerekend die ook in de veldtest van het parallelle project Intermezzo zullen worden beschouwd. Het betreft de druk bereiden corridor A58 van Breda naar Tilburg en de A27. Op basis van de simulatieresultaten mag verwacht worden dat de reistijdschattingen voor deze trajecten zeer nauwkeurig zullen zijn. In alle gevallen werd een kwadratische fout per minuut van ver onder de 0.1 berekend.

Hierbij moet wel worden aangetekend, dat in de simulatie is uitgegaan van een spreiding van de reistijd van 10% die per combinatie van schakel en rit onafhankelijk geloot wordt. Vooral bij lange schakels en langere waarneemperiodes zou dit kunnen leiden tot een onderschatting van het steekproeffeffect. Wanneer de variatie in reistijd optreedt als gevolg van congestie die tijdens de waarneemperiode opbouwt en zich over meerdere schakels uitspreid is vooral de aanname over de onafhankelijkheid veel te optimistisch. Om tot een correcte modellering te komen zou het huidige model moeten worden uitgebreid met een microscopisch simulatiemodel.

Ook als een aantal pessimistische aannames over het steekproeffeffect worden toegevoegd blijft echter de conclusie staan dat goede reistijdschattingen mogen worden verwacht op de

genoemde trajecten.

5 Conclusies

5.1 Kwaliteit van de ingewonnen verkeersgegevens

Op basis van de huidige simulatiestudie mag geconcludeerd worden dat het voor het Hoofd Wegen Net en de drukbereden gedeeltes van het Onderliggende Wegen Net mogelijk moet zijn om betrouwbare reistijdschattingen uit GSM data af te leiden. Daarbij gaat het om het samenstellen van uurgemiddelde snelheden. Voor wegen waarop de gemiddelde intensiteit laag is en wegen in gebieden met een hoge netwerkdichtheid laat de nauwkeurigheid meer te wensen over.

Wanneer de analyse periode wordt verkort tot een kwartier of de penetratie graad van de bij detectie betrokken GSM's afneemt tot beneden de 20% wordt de toepasbaarheid van de reistijd schattingen beperkt tot alleen de drukke snelwegen met een geïsoleerde ligging.

De relatieve fout neemt toe naarmate de afstand waarover een uitspraak wordt gedaan korter wordt. Voor lange trajecten gelden de kleinste relatieve fouten.

De kwaliteit van de geschatte verkeersgegevens wordt beïnvloed door vele parameters maar is uiteindelijk terug te voeren op de volgende drie factoren:

- de kwaliteit waarmee de mapmatching wordt uitgevoerd;
- het aantal ritten dat gematcht wordt;
- de representativiteit van de gematchte ritten;

Het algemene beeld is dat geaggregeerde verkeersprestatieindicatoren, zoals uurgemiddelden over lange trajecten, met een grotere relatieve nauwkeurigheid kunnen worden berekend dan de meer gedetailleerde indicatoren, zoals reistijd per kwartier op een specifiek wegvak.

De kwaliteit van het mapmatchen wordt door een aantal factoren beïnvloed. Deze kwaliteit neemt toe indien de gemiddelde duur van een gesprek toeneemt, de meetfout afneemt, een Point-based in plaats van een Area-based plaatsbepalingsmethode wordt gebruikt, en wanneer een goede indicatie van de minimum snelheid aanwezig is.

Door de duur van de waarneemperiode op te voeren kan het aantal gematchte ritten naar believen worden opgevoerd, bij een constante kwaliteit van mapmatching. De representativiteit neemt daarbij echter af, omdat naarmate een periode langer duurt er een grotere spreiding in de aangetroffen reistijden optreedt. Voor schakels met een lage intensiteit die liggen in een gebied van hoge netwerkdichtheid treden soms grote fouten bij het mapmatchen op omdat een hoog percentage van de Measurement Reports die naar deze schakels worden gematcht feitelijk gegenereerd zijn op andere schakels. Ook in dit geval helpt het niet op de waarneemperiode te verlengen.

Het is met de geïmplementeerde algoritmes niet mogelijk gebleken om betrouwbare schattingen te maken van schakel- of HB intensiteiten. Het grootste probleem bestaat uit het bepalen van geschikte ophoogfactoren. Deze zijn sterk plaatsafhankelijk. Bovendien hangen deze in praktijk ook af van een aantal moeilijk waar te nemen gedragsmatige factoren.

In de huidige studie is niet gekeken naar de mogelijkheden om andere verkeersprestatieindicatoren te bepalen dan intensiteiten en reistijden, maar op grond van de

ervaringen mag verwacht worden dat het berekenen van indicatoren als verkeersprestatie en voertuigverliesuren op een geaggregeerd niveau mogelijk is.

5.2 Voorbehoud bij de simulatieresultaten

Het gebruikte simulatiemodel gebruikt noodzakelijkerwijs een aantal vereenvoudigingen en een aantal aannames die ten tijde van het uitvoeren van de studie niet geverifieerd konden worden. Een gedeelte van de vereenvoudigingen is bewust gekozen omdat de invloed ervan klein wordt geacht. Een ander deel is nodig omdat een meer gedetailleerde beschrijving niet beschikbaar is of onder verwijzing naar commerciële belangen niet beschikbaar wordt gesteld. Het gaat om de volgende vereenvoudigingen en aannames:

- Aannames over de technische eigenschappen van GSM plaatsbepalingsgegevens:
 - de exacte statistische eigenschappen van de GSM plaatsbepalingsfouten zijn onbekend;
 - er is geen informatie beschikbaar over eventuele correlatie tussen deze fouten;
 - Het aantal en de locaties van de GSM-antennes zijn onbekend;
- Aannames over belgedrag in het algemeen en belgedrag in het wegverkeer in het bijzonder:
 - Er zijn geschatte waarden toegepast voor de gemiddelde duur van een gesprek, de gemiddelde tijd tussen twee oproepen, en het aandeel van de voertuigen dat bereikbaar is via GSM ;
 - Er is geen samenhang tussen belgedrag en rij-omstandigheden verondersteld;
- Aannames over toe te passen algoritmes:
 - Ten behoeve van de huidige studie zijn aparte algoritmes voor map-matching, reistijdschatting en intensiteitschatting ontwikkeld en geoptimaliseerd. Deze kunnen afwijken van de algoritmes die door marktpartijen worden toegepast;
- Verkeerskundige aannames:
 - Er worden geen fietsers en OV reizigers gemodelleerd;
 - Er is geen rekening gehouden met bellende automobilisten die hun rit onderbreken, bijvoorbeeld om te tanken, of hun gesprek voortzetten nadat het reisdoel bereikt is;
 - Er wordt in het model geen rekening gehouden met within-link variatie van voertuigsnelheden;
 - Het gebruikte rekennetwerk is een sterk versimpelde weergave van de werkelijkheid;
 - Er is geen rekening gehouden met eventuele coderingsfouten in het wegennetwerk;
- Overige:
 - Er wordt in het simulatiemodel geen rekening gehouden met (gedeeltelijke) systeemuitval of storing op.

Een aantal van deze aspecten kunnen worden gereduceerd door verfijningen aan het model toe te passen of door extra varianten door te rekenen. Weer een aantal andere factoren, zoals GSM gebruik in het OV, kan moeilijk worden ondervangen, maar heeft alleen een lokale werking. Vooral de kwaliteit van de resultaten voor stedelijke netwerken zullen in de praktijk

slechter uitvallen dan de simulatieresultaten suggereren, doordat een aantal ongunstige factoren door het model worden genegeerd.

5.3 Opgeleverd simulatietool en schattingsalgoritme

Samen met dit rapport wordt ook een simulatietool opgeleverd. Hiermee kan een willekeurig studiegebied binnen Nederland worden geselecteerd en doorgerekend op basis van zelf te kiezen parameters. Het tool beschikt over uitgebreide mogelijkheden om de resultaten te inspecteren.

Als onderdeel van dit tool is een mapmatch algoritme, reistijdschatter en intensiteitsschatter ontwikkeld. Deze algoritmes kunnen desgewenst ook operationeel worden ingezet omdat zij algemeen van opzet zijn. Deze algoritmes maken gebruik van alle beschikbare informatie en kunnen overweg met netwerken van realistische grootte. Informatie over minimum en maximum snelheid per schakel kan worden toegevoegd en er wordt alleen gematched naar routes die een gedragsmatige interpretatie hebben.

5.4 Kosten-baten analyse

De huidige studie geeft enige indicatie van de benodigde investeringskosten voor het tot stand brengen van een inwinsysteem voor verkeersgegevens op basis van GSM data. De belangrijkste kostenposten zijn het verzamelen van de Measurement Reports, het opstellen van een waarheidsgetrouwe digitale kaart en het bewerken van de ruwe gegevens tot verkeersinformatie.

De eerste kostenpost is tijdens de factfinding fase van dit project (zie Van der Zijpp, Haastregt, 2003) geschat op 10 m€ per operator (voor een landelijke dekking). De hoogte van deze kostenpost wordt vooral beïnvloed door het feit dat op ieder antenne locatie hardware aanpassingen nodig zijn. Mogelijk zijn hier nog kosten te besparen door een oplossing te verzinnen die dergelijke hardware investeringen overbodig maakt. De tweede kostenpost, het samenstellen van de digitale kaart, kan grotendeels worden vermeden door van het huidige NWB gebruik te maken. De derde kostenpost, het bewerken van de ruwe gegevens tot verkeersinformatie, kan op basis van de ervaringen in de huidige studie worden gemaximeerd op 300 k€. De jaarlijkse kosten voor onderhoud geschat op 15% van de investering. De afschrijvingstermijn wordt gesteld op 6 jaar.

Deze kosten moeten met name op het onderliggend wegennet worden terug verdiend, omdat het hoofdwegen net uitgerust is en blijft met detectielussen. Omdat mede op basis van de resultaten van de huidige studie nog enige twijfel bestaat over de kwaliteit van de verkeersinformatie die op het onderliggende wegen net geboden kan worden is het aan te raden investeringen uit te stellen tot dat ofwel de kosten drastisch omlaag zijn gegaan, ofwel de kwaliteit van de gegevens is verbeterd. Een dergelijke verbetering kan voortkomen uit het verbeteren van de plaatsbepalingsmethode voor GSM's of door het bewust inzetten van extra bandbreedte, die met de beschikbaarheid van UMTS in ruimere mate voorhanden zal zijn.

5.5 Rol van de overheid

Als grote producent en gebruiker van verkeersgegevens speelt de overheid een belangrijke rol in de ontwikkelingen op het gebied van data-inwinning met behulp van GSM. Wanneer alle

ontwikkelingen aan de markt worden overgelaten komt het inwinsysteem waarschijnlijk niet van de grond. Om realisatie te bevorderen kan gekozen worden voor een stimulerende opstelling. In dit geval worden partijen die op basis van zelf ingewonnen GSM, GPRS of GPS data verkeersinformatie willen genereren gestimuleerd. Dit kan door het geven van subsidies of door het beschikbaar stellen van mapmatch technologie. De overheid is afnemer van deze verkeersinformatie. De overheid kan ook nog een stap verder te gaan door zelf de Measurement Reports en eventueel GPS data bij diverse partijen in te kopen en deze in eigen beheer te bewerken.

Indien exclusief voor de stimulerende aanpak wordt gekozen, dan wordt ook de verantwoordelijkheid voor de kwaliteit van de data uitbesteed, en kan slechts uit een beperkt aantal aanbieders gekozen worden. Ook is men in dit geval van deze zelfde groep aanbieders afhankelijk voor kennis over de eigenschappen van de te verzamelen data. Dit maakt het moeilijk om investeringsbeslissingen op een verantwoorde wijze te nemen. Daarom is het advies om zo nodig als overheid ook de actieve opstelling te kiezen.

5.6 Verder onderzoek

Met het opleveren van dit rapport wordt een periode van intensief onderzoek afgesloten. Er blijven echter vele interessante en vragen over die verder onderzoek waard zijn. In het hoofdrapport worden de volgende onderzoeksopties beschreven:

- Verbetering aan het mapmatch algoritme, met name:
 - Het selectief toepassen van de minimumsnelheid constraint;
 - Het verbeteren van het mapmatch algoritme door gebruik te maken van apriori route kansen;
- Verbeteringen aan de intensiteitsschatter:
 - Het verbeteren van de intensiteitschatting door toepassing van de conservation of flow;
 - Het combineren van GSM gebaseerde intensiteitschattingen en lusdetectie data;
 - De bepaling van betere ophoogfactoren door middel van simulatie;
- Aanvullen/verfijnen scenario's:
 - Het doorrekenen van extra scenario's;
 - het aanbrengen van verfijningen in de modellering van het GSM data-inwinsysteem.
- Het opstellen van ontwerpregels voor het dimensioneren van een GSM gebaseerd inwinsysteem