

Vroeger voorspelden we de toekomst beter.

Een discussiepaper.

Arnout Schoemakers
Oranjewoud
arnout.schoemakers@oranjewoud.nl

Karst Geurs
Planbureau voor de Leefomgeving
karst.geurs@pbl.nl

**Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk
20 en 21 november 2008, Santpoort**

Samenvatting

Vroeger voorspelden we de toekomst beter. Een discussiepaper

In dit discussiepaper betogen we dat de waarde die aan verkeersprognoses wordt gehecht de afgelopen 10 jaar is toegenomen. Gelijktijdig is de kwaliteit van de verkeersvraagmodellering in de praktijk de afgelopen decennia nauwelijks verbeterd. Ofwel: vroeger voorspelden we de toekomst beter.

De uitkomsten van verkeersmodellen zijn de afgelopen jaren een steeds belangrijkere rol gaan spelen in het besluitvormingsproces over rijksinfrastructuurprojecten. Resultaten van kosten-batenanalyses en milieustudies zijn doorslaggevend geworden voor het al dan niet aanleggen van infrastructuur. De uitkomsten van verkeersmodellen zijn van grote invloed op de uitkomsten van kosten-batenanalyses en toetsen op luchtkwaliteitsnormen. De waarde die aan verkeersprognoses wordt toegekend is daarmee sterk toegenomen.

Als we naar de vernieuwingen in de modellering van de verkeersvraag in de afgelopen 50 jaar kijken dan zien we dat vooral de wijze van modelleren van de verkeersvraag in de gemeentelijke modellenpraktijk nauwelijks is geëvolueerd. De vraag is hoe erg dit is. Dit hangt uiteraard af van de vraagstelling. Als de trend van gebruik van kosten-batenanalyses zich doorzet naar gemeentelijk of regionaal verkeersbeleid dan is volgens ons een forse verbetering nodig. Het grootste probleem is dat nog volop wordt gewerkt met geaggregeerde verkeersmodellen (met een beperkte of geen bevolkingssegmentatie in de modellen voor vervoerwijze en bestemmingskeuze). In de evolutie van Nederlandse verkeersmodellen valt ook op dat er tot nu toe geen regionale en of gemeentelijke modellen ontwikkeld zijn die verkeer in samenhang met ruimtegebruik bezien, iets wat we in andere landen in Europa wel zien. Ook op nationale schaal worden dergelijke modellen nog niet ingezet bij verkennende studies of planstudies.

Er bestaat geen breed geaccepteerd richtlijnenkader voor het kwalificeren van een verkeersmodel. Verschillende overheden en marktpartijen gebruiken verschillende kaders. De enige kwantitatieve richtlijn die in de praktijk wordt gehanteerd is gericht op het benaderen van verkeerstellingen door het verkeersmodel voor het basisjaar. Verder is het vooral 'expert judgement'.

Onzekerheden lijken bij veel verkeersprognoses geen rol te spelen. Er wordt veelal gebruik gemaakt van één model, één set modelinvoer en één omgevingsscenario (met inschatting bevolkingsgroei, economische groei etc.). Maar de modellen kennen een onzekerheid, evenals de invoervariabelen (waaronder ook beleidsvariabelen). Verkeersprognoses zijn intrinsiek onzeker. Wij pleiten ervoor om met minimaal 2 omgevingsscenario's te werken en bandbreedtes in prognoses aan te geven.

De vraag hoe goed Nederlandse verkeersmodellen de daadwerkelijke effecten van infrastructuurprojecten modelleren is tot nu toe niet of nauwelijks beantwoord. Modellen leren onvoldoende van het verleden. Uit buitenlands onderzoek blijkt dat er grote systematische en willekeurige afwijkingen tussen prognoses en realiteit kunnen ontstaan. Wij pleiten er voor om afgeronde projecten systematisch te evalueren. Het gaat hierbij niet om het achteraf 'afrekenen' op fouten of om het terugdraaien van beslissingen maar om onderzoekers en modellers in staat te stellen te leren van het verleden. Het doel moet vooral zijn om met informatie uit ex post evaluaties de kwaliteit van verkeersmodellen en -prognoses te verbeteren, en hiermee de besluitvorming over verkeers- en vervoerbeleid.

1. Inleiding

Op het eerste CVS in 1974 stonden verkeersmodellen al centraal. In de 35 edities van het CVS zijn vele papers over de ontwikkeling en toepassing van verkeersmodellen geschreven. Modellen helpen bij het in beeld brengen van effecten van maatregelen die ingrijpen op het verkeers- en vervoersysteem maar we zien dat de rol van modellen in de loop van de tijd wel is gewijzigd. Tegenwoordig wordt steeds meer waarde gehecht aan de precieze uitvoer van de modellen. De resultaten worden gebruikt als invoer voor milieuberekeningen en voor kosten-baten analyses. De resultaten met verkeersmodellen worden hierbij als exacte voorspellingen gezien. De Nederlandse wet- en regelgeving eist dat de effecten van infrastructuraanleg op hoog detailniveau en exactheid worden berekend met een tijdshorizon van 2015 of 2020. Zijn de verkeersmodellen in de afgelopen 35 jaar echter beter geworden? Kunnen we de toekomst nu beter of exacter voorspellen? Of weten we dat eigenlijk helemaal niet? De vraag is of daar waar de eisen aan de modelresultaten zijn aangescherpt ook de modellen zelf of het gebruik daarvan eenzelfde kwaliteitsslag hebben gemaakt.

Dit is een discussiepaper. Het doel van de paper is het voeren van een discussie over de kwaliteit en het gebruik van verkeersmodellen. De paper is vooral gebaseerd op ervaringen met ontwikkeling en gebruik van landelijke, regionale en lokale verkeersmodellen. We richten ons hierbij op strategische modellen die de vervoervraag modelleren en dan met name op personenmobiliteit. In de paper betogen we in de eerste plaats dat er nauwelijks aanwijzingen zijn dat het voorspellend vermogen van verkeersmodellen ten opzichte van vroeger is toegenomen. De eisen die tegenwoordig aan modellen worden gesteld lijken sterk te zijn toegenomen, maar de kwaliteit van de modellen is niet of nauwelijks geëvolueerd. In de tweede plaats betogen we dat modellen nauwelijks leren van het verleden. Er wordt vrijwel nooit gekeken of de prognoses van modellen daadwerkelijk uitkomen en wat we daarvan kunnen leren.

In de paper gaan we kort in op het gebruik van modelcijfers, de eisen die aan verkeersmodellen in theorie en in de praktijk worden gesteld en de kwaliteit van de modellen zelf. We sluiten af met een aantal vragen en stellingen die hierbij aansluiten.

2. Resultaten verkeersmodellen steeds belangrijker bij besluitvorming

Het verkeers- en vervoersysteem is complex. Zo complex dat het vaak moeilijk is om op rationele wijze de effecten van een maatregel die hierop ingrijpt goed in beeld te krijgen. Denk bijvoorbeeld aan de effecten van het aanleggen van een spoorlijn of een weg, aan de effecten bij een gewijzigde bevolkingsomvang- en samenstelling in de toekomst of aan een gewijzigd economisch klimaat. Om een beeld te krijgen van de effecten van deze complexe vraagstukken – en varianten daarop – worden er mathematische modellen gebouwd die het verkeerssysteem – of delen daarvan – beschrijven en de effecten van wijzigingen hierin in beeld kunnen brengen. Als hulpmiddel om te structureren bij het afwegingskader. Voor sommige vragen volstaat een globaal beeld van de effecten en voor andere vragen is meer precieze informatie nodig.

Verkeersmodellen spelen een belangrijke rol in het besluitvormingsproces. En deze rol is in de loop der tijd alleen maar toegenomen:

- Bij planstudies wordt berekend of het aanleggen of verbreden van infrastructuur leidt tot een overschrijding van luchtkwaliteitsnormen en beleidsuitgangspunten voor geluid. Prognoses met verkeersmodellen vormen een belangrijke input voor lucht- en geluidberekeningen en hiermee ook voor het wel of niet doorgang kunnen vinden van een project. Dit vraagt om betrouwbare verkeerscijfers.
- Voor een rationele afweging van infrastructuurprojecten worden de afgelopen 10 jaar steeds vaker kosten-baten analyses uitgevoerd. Een kosten-batenanalyse volgens de OEI leidraad is al bijna 10 jaar verplicht voor grote infrastructuurprojecten, en in

planstudies is sinds kort een vereenvoudigde kosten-batenanalyse ('OEI bij MIT') verplicht. Ook deze analyses worden gevoed met verkeerscijfers uit verkeersmodellen. Veelal zijn dit het landelijk modelsysteem verkeer en vervoer (LMS) en regionale modellen (NRM) van Rijkswaterstaat. Verkeersmodellen hebben hiermee een belangrijke invloed op de resultaten van economische evaluaties van rijksinfrastructuurinvesteringen en daarmee besluitvorming over infrastructuur in Nederland. Zo zijn reistijdwinsten van weginvesteringen, die zijn ingeschat met behulp van verkeersmodellen, veelal dominant in uitkomst van maatschappelijke kosten-baten analyses. Verkeerscijfers krijgen hiermee een gewicht dat zwaarder is dan bij het gebruik waar ze oorspronkelijk voor zijn ontwikkeld.

Daarnaast wordt door het beleid steeds specifiekere informatie gevraagd, die als tussenresultaat wel gebruikt worden in een model, maar waarop een model vaak niet is toegesneden bij het ontwerp. Denk bijvoorbeeld aan het genereren van informatie over reistijden of de betrouwbaarheid van reistijden.

Stelling: De afgelopen 10 jaar is de waarde die aan verkeersprognoses wordt gehecht sterk toegenomen.

3. Ontwikkeling van verkeersmodellen in Nederland

In de jaren vijftig van de vorige eeuw werden de eerste verkeersmodellen ontwikkeld. Dit waren eenvoudige unimodale modellen (auto) gebaseerd op de zwaartekrachtstheorie. In de jaren '60 en '70 zijn deze modellen ontwikkeld tot de bekende 4 steps multimodale verkeersmodellen.

In de jaren tachtig vond een belangrijke vernieuwing in modellering van de vervoervraag plaats; de gedeasaggregeerde discrete keuze (logit) modellen werden ontwikkeld. Een belangrijk element hierbij is dat rekening wordt gehouden met het feit dat verschillende bevolkingssegmenten een eigen mobiliteitsgedrag vertonen. De modelsystemen van Rijkswaterstaat (Landelijk Modelsysteem - LMS en regionale modellen - NRM) maar ook onderdelen van het huidige gemeentelijke model van Amsterdam (Genmod) werken met deze techniek. De belangrijkste verbetering is het detailniveau van vervoervraagmodellering, zo maakt het LMS onderscheid naar 49 persoonstypen, 9 verplaatsingsmotieven en 3 tijdspanes. Het LMS en NRM's bevatten veel verschillende keuzemodellen (zoals tijdstipkeuzemodellen, routekeuzemodellen, vervoerwijzekeuzemodellen) in onderlinge samenhang. Dit maakt het inzicht in de werking van de modellen complex. Ook is de ontwikkeling, de toepassing en het beheer van dergelijke modellen in vergelijking met de klassieke geaggregeerde zwaartekrachtmodellen tijdrovend en complex.

Vanaf de jaren tachtig zien wij drie belangrijke hoofdrichtingen bij modelontwikkeling. De eerste richting is de integratie van transport en grondgebruik. In Nederland heeft Rijkswaterstaat TIGRIS XL ontwikkeld (Zondag et al., 2005), een grondgebruik-transport-interactie (GTI) model (op basis van discrete keuzetheorie) waarin zowel locatiekeuzegedrag van huishoudens en bedrijven als verplaatsingsgedrag in onderlinge samenhang is gemodelleerd. TIGRIS XL is een landelijk model waarin het LMS als verkeersmodel is opgenomen. GTI-modellen zijn in theorie geschikter voor de evaluatie van verkeersbeleid dan 'stand-alone' verkeersmodellen, omdat ze expliciet rekening houden met de wederzijdse relaties tussen het ruimtelijke systeem (vraag een aanbod van activiteiten, gebouwen en grondgebruik) en het transportsysteem (vervoervraag en de kenmerken van het infrastructuraanbod). Bij 'stand-alone' verkeersmodellen is de ruimtelijke component exogeen. Er zijn in Nederland geen GTI-modellen ontwikkeld specifiek voor de regionale of en lokale schaal. In andere landen in Europa en elders zijn juist op het stedelijke/metropolitane schaalniveau veel grondgebruik-transport-interactie modellen in gebruik.

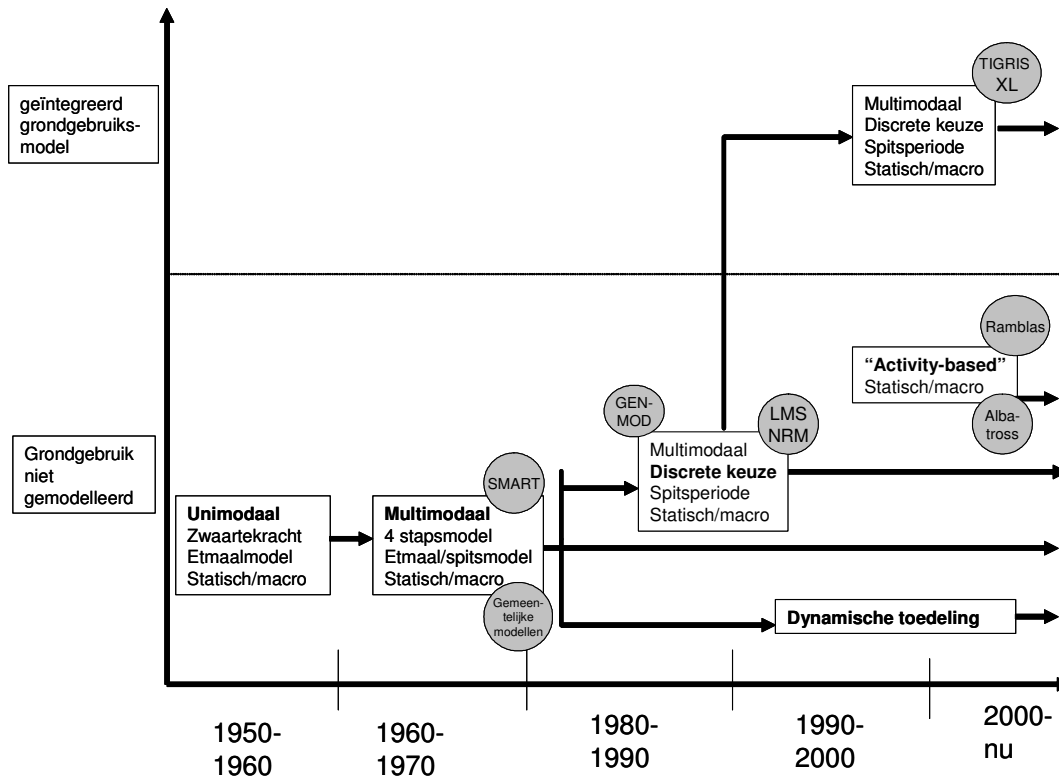
Een tweede ontwikkelrichting is de ontwikkeling van zogenoemde 'activity-based' modellen. Deze modellen modelleren het verplaatsingsgedrag van individuen als afgeleide van activiteitenpatronen van huishoudens (in tegenstelling tot de eerdere 'trip-based' modellen) en zijn daarbij voor de 'klant' erg intuïtief. Voorbeelden zijn het Albatross model van Rijkswaterstaat (Arentze en Timmermans, 2005) en RAMBLAS (Veldhuisen et al., 2005). De modellen zijn echter nog maar beperkt toegepast. Ook in het buitenland zijn nog maar weinig operationele activity-based modellen die het hele verkeers- en vervoersysteem bevatten. De ontwikkeling is tot nog toe voornamelijk gericht op methodologische vooruitgang, en niet op de toepassingen voor de evaluatie van ruimtelijk of transport beleid. Albatross loopt hierin voorop.

De derde ontwikkelrichting richt zich op een specifiek onderdeel van een verkeersmodel: de verkeersafwikkeling. Het gaat dan om de ontwikkeling van dynamische toedelingmethoden waarmee op microniveau (afzonderlijke voertuigen) of mesoniveau (groepen voertuigen) de voertuigafwikkeling op een kruispunt of netwerk wordt berekend en waarbij rekening wordt gehouden met 'bottleneck' effecten. Er zijn inmiddels verschillende softwarepakketten op de markt (zoals Vissim, Paramics) waarmee – gevoed door de verkeersvraag uit de statische/strategische modellen – dynamische simulaties worden gemaakt. Daarnaast zijn er diverse softwarepakketten op de markt waarmee (geaggregeerde) strategische verkeersmodellen relatief eenvoudig gebouwd, toegepast en beheerd kunnen worden (zoals VISUM, Omnitrans, CUBE/Voyager, Questor). Tot slot constateren we dat de verkeerskundige software steeds beter geïntegreerd wordt met GIS. Voor visualisatie-doeleinden, maar een koppeling is bijvoorbeeld vaak ook gewenst voor nabewerkingen cq. berekeningen voor geluid en luchtkwaliteit.

Tot slot wordt er de laatste tijd veel aandacht besteedt aan het meer precies modelleren van vertrektijdstipkeuze.

Een ontwikkeling die de komende jaren vorm gaat krijgen richt zich op een integratie van de dynamische toedeeltechnieken in combinatie met een dynamische vraagberekening (11 tijdperioden in plaats van 3 nu). De volgende generatie NRM's en het LMS (met basisjaar 2008) worden hier – zoals het er nu naar uitziet – mee uitgerust. Dat zou een kwaliteitsimpuls kunnen betekenen voor de vraagberekening.

Figuur 1 geeft de historische ontwikkeling van verkeersmodellen in Nederland weer. Als we naar de vernieuwingen in de modellering van de verkeersvraag kijken dan zien we dat die voor de meeste gemeentelijke modellen niet echt zichtbaar heeft plaatsgevonden. Voor zover wij weten werkt alleen het huidige model (GENMOD) van de gemeente Amsterdam met gedesaggregeerde (multinomiale logit) modellen waarbij rekening wordt gehouden met specifiek verplaatsingsgedrag van homogene groepen personen en huishoudens. De belangrijkste inhoudelijke vernieuwing in de gemeentelijke beleidspraktijk die zichtbaar is, is de toepassing van dynamische toedelingmethoden en verbeterde visualisatietechnieken in het algemeen. Het klassieke geaggregeerde vierstapsmodel (tripgeneratie, distributie, vervoerwijzekeuze en toedeling) wordt in Nederland nog steeds volop toegepast in bijna alle gemeentelijke, een aantal regionale verkeersmodellen (bijvoorbeeld het Verkeersmodel Regio Utrecht, het model van het Samenwerkingsverband Regio Eindhoven en het verkeersmodel Haaglanden).maar ook in het nationale model SMART.



Figuur 1: Ontwikkeling verkeersmodellen in de praktijk in Nederland

Methodische verbeteringen vanaf de jaren negentig (interactie met grondgebruik, op activiteiten gebaseerde modellen) hebben een beperkte weerslag in de (nationale of regionale) modellenpraktijk gekregen. Deze modellen zijn tot nu toe niet of nauwelijks toegepast bij beleidsanalyses. Het model TIGRIS XL is tot nu toe alleen in een verkennende studie van het Planbureau voor de Leefomgeving (voorheen Milieu- en Natuurplanbureau) gebruikt (MNP, 2007; Zondag et al., 2007) en in een verkennende studie naar de ruimtelijk structurerende werking van infrastructuur bij verschillende N18 varianten in de Achterhoek. Het model is tot nu toe niet in planstudies gebruikt. Albatross is sinds anderhalf jaar operationeel. Het KIM heeft met dit modelsysteem een studie uitgevoerd naar de effecten van vergrijzing op het activiteitenpatroon en – daarvan afgeleid – de effecten voor de mobiliteit.

We zien dus dat de methoden en technieken voor het berekenen van de verkeersvraag in de loop der tijd niet sterk is geëvolueerd, met name voor gemeenten. De vraag is wat hiervan de oorzaak kan zijn. Een belangrijke oorzaak ligt vermoedelijk in de beperkte databeschikbaarheid op dit niveau. Er is nauwelijks empirische informatie over het reisgedrag op lokaal niveau. Het gevolg hiervan is dat er veel informatie moet worden 'bijgeschat', veelal op basis van basale rekenregels. Een andere oorzaak is dat er nauwelijks eisen worden gesteld aan het bepalen van de verkeersvraag, maar daarover later meer.

We constateren ook dat de hoeveelheid empirische informatie op basis waarop modellen kunnen worden geschat verder aan het afnemen is. De steekproef van het Onderzoek Verplaatsingsgedrag (OVG) was veel groter dan het Mobiliteitsonderzoek Nederland (MON) van nu. Hetzelfde geldt voor de beschikbaarheid van enquêtes. En ook een volledige set betrouwbare en actuele verkeerstellingen ontbreekt vaak, waardoor er in de praktijk soms op oneigenlijke manieren informatie wordt toegevoegd.

Stelling: De modellering van de verkeersvraag is vanaf de jaren tachtig van de vorige eeuw nauwelijks geëvolueerd in de tijd. Daarnaast is de databeschikbaarheid afgenomen. En er worden wel hogere eisen gesteld aan de resultaten. Dit rijmt slecht.

4. Eisen aan verkeersmodellen – toetsing kwaliteit vooraf

Uit sectie 2 hebben we aangegeven dat de waarde die wordt gehecht aan modeluitkomsten zijn toegenomen. In Sectie 3 stellen we dat de modellering van verkeersvraag maar beperkt is verbeterd. Wat betekent dit voor de geschiktheid van state-of-the-practice verkeersmodellen bij beleidsevaluaties? Welke eisen kun je stellen aan verkeersmodellen? Of: hoe goed is een model. Dat is een vraag die niet eenduidig te beantwoorden is. Hier is geen richtlijnenkader voor. De eisen zijn ook afhankelijk van het doel van het model en de beleidsvraag die je wilt beantwoorden. Je zou het kunnen staven aan de stand der techniek in de wetenschap en aan de hoeveelheid en kwaliteit van de empirische data. Als een model wetenschappelijk state of the art is en wordt gevoed door een zeer complete set empirische informatie dan zou een model in theorie goed moeten zijn.

Voor het LMS is een dergelijke audit al eens uitgevoerd. In 1996 is de kwaliteit van het model getoetst door het instituut Transport Research Laboratory, gelieerd aan het Britse Ministerie van Transport. Het instituut concludeerde dat het LMS volgens de toen laatste wetenschappelijke inzichten is gemaakt.

Maar dergelijke audits worden op dit niveau in Nederland zelden uitgevoerd. Aangezien er geen brede richtlijnen bestaan voor het kwalificeren van een verkeersmodel is het de vraag welke eisen in de praktijk wordt gehanteerd. Daarnaast stellen toepassingen in kosten-batenanalyses en toetsen op luchtkwaliteit verschillende eisen aan modellen.

Kwaliteitseisen in de praktijk

In de praktijk wordt de kwaliteit van een model vaak afgemeten aan een aantal elementen:

- De gebruiker en afnemer moet een goed gevoel hebben bij de resultaten. Dat betekent dat modelresultaten logisch en intuïtief moeten zijn en passen binnen het referentiekader;
- Informatie van het basisjaar van het model wordt gestaafd aan empirische locatie-specifieke kenmerken uit bijvoorbeeld het OVG/MON, denk bijvoorbeeld aan ritlengteverdelingen, ritproducties, modal split, reizen en reizigerskilometers;
- Modelwaarden van het basisjaar moeten in een bepaalde mate overeen komen met getelde waarden;
- De niet gekalibreerde matrix en de daaruit voortvloeiende verkeersstromen moet lijken op de gekalibreerde matrix;
- Soms worden kentekenquêtes gebruikt om de verkeersstromen in een model preciezer in te brengen dan met de generieke rekenregels is bepaald.

Technieken die gebruikt worden om een model te maken, maken geen onderdeel uit van de bepaling van de kwaliteit van een model. Denk hierbij bijvoorbeeld aan de wijze waarop een matrix wordt bepaald of gekalibreerd.

Voor het vergelijk van modelwaarden voor het basisjaar met telwaarden zijn er door Rijkswaterstaat wel een aantal richtlijnen opgesteld en op basis hiervan worden ook vergelijkbare – maar soms iets strakkere – eisen gesteld aan lokale modellen. Zo wordt bij de calibratie van een NRM gesteld dat voor de wegvakken met meer dan 10.000 motorvoertuigen/etmaal niet meer dan 5% mogen afwijken van de telcijfers. Het wordt acceptabel gevonden wanneer maximaal 10% van de telpunten niet binnen deze grenswaarde valt. Hiermee kan de kwaliteitseis worden geïnterpreteerd als een toets met

een 90 % betrouwbaarheidsinterval. Voor wegen met minder verkeer mag de afwijking volgens de gehanteerde kwaliteitsprocedures groter zijn. De kwalificatie van de overige modeluitkomsten is op basis van 'expert judgement'. Verder zien we dat verschillende overheden en marktpartijen verschillende richtlijnenkaders hanteren.

Stelling: Er bestaan nauwelijks kwantitatieve richtlijnen voor het kwalificeren van een verkeersmodel. De enige kwantitatieve richtlijn die in de praktijk wordt gehanteerd is gericht op het benaderen van verkeersstellingen door het verkeersmodel voor het basisjaar. De kwalificatie van de rest is op basis van 'expert judgement'.

Kwaliteitseisen vanuit kosten-batenanalyses

In de literatuur zijn wel enkele aangrijpingspunten voor modeleisen te vinden. Zo wordt in een Engelse review van grondgebruik-transport-interactiemodellen (DSC/ME&P, 1999) een aantal algemene eisen gesteld waaraan een transportmodel idealiter zou moeten voldoen om de transportbaten van infrastructuurprojecten goed in te kunnen schatten. Dit zijn dus eisen waaraan een verkeersmodel moet voldoen om gebruikt te kunnen worden bij kosten-batenanalyses. DSC/ME&P stellen de volgende eisen:

1. Het gesimuleerde verplaatsingsgedrag dient een consistente uitkomst te zijn van mogelijke gedragsreacties op ruimtelijke veranderingen (zoals het wijzigen van locaties van activiteiten) wijzigingen in het transportsysteem, zoals een wijziging in de transportkosten of de kenmerken van het infrastructuraanbod. DSC/ME&P (1999) onderscheiden zes mogelijk gedragsreacties: aanpassing van reisfrequenties, tijdstip van verplaatsing, routekeuze, vervoerwijzekeuze, bestemmingskeuze en het combineren van verplaatsingen en/of activiteiten door individuen. Het model moet ook de verschillende mogelijke gedragsreacties op een realistische manier aan elkaar koppelen.
2. De personenvervoermodellen moeten in voldoende mate een segmentatie van de bevolking kennen om verschillen in verplaatsingsgedrag tussen bevolkingsgroepen te kunnen representeren, zoals een onderscheid sociaaleconomische kenmerken (inkomen, leeftijd, etc.).
3. De sterkte van gedragsreacties moet zijn geschat op basis van werkelijk verplaatsingsgedrag;
4. Het model moet zowel gevoelig zijn voor financiële veranderingen (waarbij een verdubbeling van kosten niet perse hoeft te resulteren in een verdubbeling van het effect), zoals prijsverhogingen, als niet-financiële veranderingen, zoals reistijdverbeteringen en verbeteringen van serviceniveaus.
5. Gedragsreacties moeten in het model op een incrementele wijze worden gemodelleerd, rekening houdende met tijdsvertragingen in gedragsreacties. Veranderingen moeten ten opzichte van een startpunt (basisjaar) worden gemodelleerd dat een goede weergave geeft van waargenomen gedrag (verplaatsingsgedrag, locaties van activiteiten, etc.).

DSC/ME&P geven aan dat de set van eisen een ideaalbeeld geeft van modellen die in de praktijk - door beperkte budgetten voor modelontwikkeling - wellicht moeilijk is te halen. In dit paper geven we geen gedegen analyse van de mate waarin de huidige verkeersmodellen in Nederland voldoen aan bovenstaande eisen. Wij vermoeden dat alleen de huidige gedesaggregeerde discrete-keuzemodellen (LMS, NRM, GENMOD) voldoet aan deze eisen. Zo hebben deze modellen een groot detailniveau in segmentatie van homogene bevolkingsgroepen en modellering van gedragsreacties. Toch hebben de huidige versies van LMS en NRM een groot aantal verbeterpunten (zie Geurs et al., 2007; van Mourik, 2008). Voor zover wij hebben kunnen achterhalen worden in de geaggregeerde zwaartekrachtsmodellen die op regionale en lokale schaal worden toegepast geen bevolkingssegmentaties gebruikt in vervoerwijzekeuze en distributiemodellen. In het algemeen levert het modelleren van de gedragsreacties van één gemiddelde bevolkingsgroep andere resultaten op dan de som van verschillende

(homogene) bevolkingsgroepen. De gedragseffecten worden hierdoor in het beste geval overschat. Maar volgens DSC/ME&P (1999) kunnen modellen die onvoldoende onderscheid maken naar verschillende bevolkingsgroepen zelfs misleidende resultaten genereren. De conclusie is dat kosten-baten analyses hogere eisen stellen aan de modelfunctionaliteit dan de eisen die momenteel in de gemeentelijke praktijk aan de verkeersmodellen worden gesteld.

Stelling: Als de trend naar kosten-batenanalyses zich doorzet naar gemeentelijk of regionaal verkeersbeleid dan is een forse verbetering in de gemeentelijke modellenpraktijk nodig.

Kwaliteitseisen vanuit toets op luchtkwaliteitsnormen

In Nederland is de EU-regelgeving aanvankelijk geïmplementeerd via het Besluit Luchtkwaliteit 2001. Bijzonder is dat hierin een koppeling werd aangebracht met ruimtelijke besluitvorming. De luchtkwaliteit mag niet verslechteren door nieuwe ruimtelijke projecten zolang niet wordt voldaan aan de EU-grenswaarden. De nationale wetgeving is daarna twee keer bijgesteld, vooral om besluitvorming rond ruimtelijke projecten meer flexibiliteit te bieden bij het toetsen aan de grenswaarden. Het resultaat van deze bijstellingen zijn het Besluit Luchtkwaliteit 2005 en sinds november 2007 de wet Milieubeheer (luchtkwaliteitseisen). Op dit moment geldt nog de eis dat ruimtelijke projecten, waaronder infrastructuurprojecten, die minder dan 1% van de norm bijdragen aan de verslechtering van de lokale luchtkwaliteit (in 2015), individueel moeten worden getoetst. Op het moment dat het zogenoemde Nationaal Samenwerkingsprogramma Luchtkwaliteit (NSL) is vastgesteld wordt deze eis versoepeld naar 3%. Daarnaast kunnen plannen die wel in betekenende mate bijdragen worden opgenomen in het NSL en vervolgens collectief worden getoetst op de luchtkwaliteitsnormen. Tot nu toe stond de Raad van State heel beperkte toenames van de overschrijding van de norm toe. Zo besliste de RvS in de zaak van een Factory Outlet Center in Roosendaal dat een toename van circa $0,1 \mu\text{g}/\text{m}^3$ (0,3% van de grenswaarde) niet zodanig klein is dat wordt voldaan aan de regelgeving. In de zaak van het nieuwe ADO-stadion in Den Haag ging het zelfs over $0,01-0,02 \mu\text{g}/\text{m}^3$ (0,03-0,05% van de grenswaarde) op korte afstand van de rijweg (Diederens en Koelemeijer, 2008). Het aantal voertuigen dat een concentratieverandering van $0,1 \mu\text{g}/\text{m}^3$ tot gevolg heeft, is voor snelwegen geschat op 1000 voertuigen per etmaal (500 per richting) voor NO_2 en 3000 voertuigen voor PM_{10} (Brief Minister van VenW aan de Tweede Kamer, 27 november 2007, VenW/DGP-2007/9369).

De vraag is of het mogelijk is om met een dergelijke mate van exactheid luchtkwaliteits- en verkeersprognoses te maken. Het antwoord hierop is nee. Alleen al onzekerheden in de berekening van actuele luchtkwaliteitconcentraties (door spreiding in meetwaarden, modelberekeningen) lopen op tot meer dan 15% (Diederens en Koelemeijer, 2008). Bovenop deze onzekerheden komen de onzekerheden in verkeersmodellen. Alle verkeersmodellen kennen onzekerheden. Deze zijn afkomstig uit de invoergegevens en kenmerken van de modellen. Door het voormalige Rijkswaterstaat Adviesdienst Verkeer en Vervoer is er een studie uitgevoerd naar de robuustheid van modeluitkomsten van het Landelijk Model Systeem (LMS). Hierbij is enerzijds op een slimme manier gevarieerd met de modelparameters en anderzijds met de modelinvoer. De onzekerheid in verkeersintensiteiten op geselecteerde wegvakken op het hoofdwegennet bleek circa 7-13% te bedragen (De Jong et al., 2007). Deze onzekerheid bleek zowel voort te komen uit onzekerheden in de omgevingsfactoren die invoer zijn voor het model (ontwikkeling inkomens, autobezit, etc.) als kenmerken van de modellen zelf. Volgens het KiM is de geschatte onzekerheid in de invoer en modelparameters groter (Annema en De Jong, 2007).

Volgens een inschatting van het Planbureau voor de Leefomgeving bedraagt de onzekerheid, inclusief de onzekerheid in de verkeersprognoses, in de toekomstige concentraties op een willekeurige plaats in Nederland minimaal 20% (Diederens en

Koelemeijer, 2008). Geen enkel model biedt zekerheden dat een milieunorm op een bepaalde locatie in de toekomst wel of niet wordt overschreden. (Annema en de Jong, 2007).

Stelling: Geen enkel verkeersmodel kan voldoen aan de eisen die momenteel (nog) vanuit de juridische toetsing op luchtkwaliteitsnormen worden gesteld.

4. Toetsing kwaliteit verkeersmodellen en -prognoses achteraf (ex post)

Het voorgaande lezend kan de vraag gesteld worden hoe goed een model nu eigenlijk is. En of we een model wel op een goede manier gebruiken. Wat is goed? Als een model goed kan voorspellen. Dat klinkt logisch. Maar die vraag wordt in Nederland in de praktijk eigenlijk nauwelijks gesteld.

Uit buitenlands onderzoek blijkt dat verschillen tussen prognoses en daadwerkelijke verkeersvolumes aanzienlijk kunnen zijn. Uit een onderzoek naar 210 infrastructuurprojecten uit 14 landen die in de afgelopen decennia zijn gerealiseerd (Flyvbjerg et al., 2006) blijkt dat verkeersprognoses over het algemeen significant afwijken van de realisatie. Vooral bij railprojecten is de afwijking groot. Bij het overgrote deel van de (in totaal 27) railprojecten bleken de vervoerprognoses hoger te liggen dan de daadwerkelijke passagiersaantallen, gemiddeld waren de actuele reizigersaantallen circa de helft van de geprognoseerde aantallen. Bij wegenprojecten bleek dat in de helft van de (ruim 180) projecten het verschil tussen prognose en werkelijke verkeersvolumes meer dan 20% was, en bij een kwart van de projecten was het verschil groter dan 40%. Hierbij werden de verkeersvolumes even vaak onderschat als overschat. Flyvbjerg ziet geen verbetering in de ontwikkeling van de nauwkeurigheid van modelprognoses. Zijn conclusie is "if techniques and skills for arriving at accurate demand forecasts have improved over time, as often claimed by forecasters, this does not show in the data". Uiteraard is het zo dat de realiteit er altijd anders uit zal zien dan de verkeersprognose, alleen al omdat scenariogegevens (bevolkingsgroei, economische groei, olieprijsen etc.) die invoer zijn voor verkeersmodellen niet overeen zullen komen met de werkelijkheid.

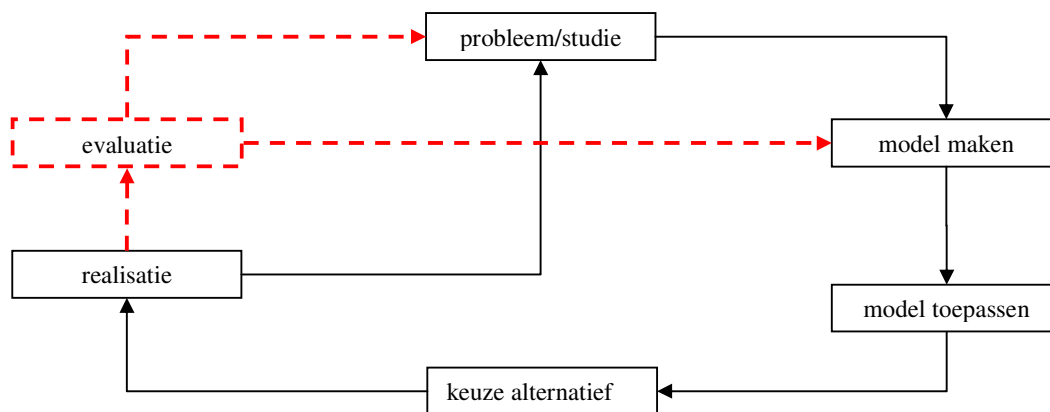
In een aantal landen is ervaring met ex post evaluaties van infrastructuurprojecten. In *Frankrijk* geldt sinds 1982 al een verplichting om infrastructurele projecten ex post te evalueren. Nadat de systematiek in 2001 werd vervangen door een eenvoudiger, en er meer gecontroleerd werd op naleving, werd de verplichting effectief. In de eerste 4 jaar van de nieuwe systematiek werden er zo'n zestig evaluaties uitgevoerd, die overigens nog niet allemaal afgerond zijn. Hierbij wordt melding gemaakt van verkeerde inschattingen bij vervoersvoorspellingen, waarbij onjuiste aannames in de vervoersmodellen mede een rol gespeeld hebben (Jeannesson-Mange et al., 2007). Of deze bevindingen geleid hebben tot aanpassingen van de modellen of aannames is niet bekend, maar de mogelijkheid is er in ieder geval. Het *Verenigd Koninkrijk* kent (sinds 2001) een Post Opening Project Evaluation (POPE) voor doorgaande wegen. Dit houdt in dat iedere aangelegde weg 1 jaar na opening en 3 tot 5 jaar na opening ex post geëvalueerd wordt. De evaluatie is beperkt tot vervoersvolumes, reistijden en ongevallen. Het *Australische* Department of transport and regional services voert vijf jaar na opening economische ex-post analyses van infra-projecten uit.

Er is tot nu toe in Nederland vrij weinig onderzoek verricht naar de mate waarin verkeersmodellen effecten van infrastructuurprojecten nauwkeurig inschatten. NEI en CE kwamen in 2001 na een uitgebreide zoektocht tot 3 ex post evaluaties van wegenbouwprojecten, 2 evaluaties van railprojecten (Ringlijn Amsterdam, Schiphollijn) en 6 ex post evaluaties van benuttingsmaatregelen (NEI/CE, 2001). Eén ex post evaluatie bestond uit een kosten-baten analyse (aanleg van de ringweg Amsterdam A10). Voor zover bekend is dit ook het enige project waarbij is gekeken of de

verkeersprognose overeenkwamen met de werkelijke situatie. Onderzocht is of de LMS prognose overeen kwam met de werkelijke verkeersintensiteiten na 1 jaar en 5 jaar (de Jong et al., 1998; Kroes et al., 1996). Vijf jaar na aanleg bleek de prognose op het niveau van de gehele dwarsdoorsnede Noordzeekanaal/IJsselmeer redelijk goed te kloppen. Het LMS prognostiseerde een toename van autoverplaatsingen 8%, de werkelijke toename was 7%. De afwijkingen in verkeersintensiteiten op wegvakniveau zijn niet gerapporteerd.

Tot slot is er in het verleden met het LMS een backcast uitgevoerd: hierbij is uit het verleden een prognose gemaakt van een recenter jaar. Er is hierbij een voorspelling gemaakt van 1996 uit 1986 en die is vergeleken met de werkelijkheid in 1996. De kwaliteit van de voorspelling bleek redelijk te zijn. Alleen de groei van het sociaal-recreatief verkeer had het model onderschat. Vanaf dat moment is de ontwikkeling van de huishoudinkomens meegenomen in het model. Een voorbeeld dus waar geconstateerde beperkingen hebben geleid tot een aanpassing van het functioneel ontwerp van het model.

In Nederland kijken we dus bijna nooit of de prognose achteraf goed bleek te zijn en - als dat niet zo was - wat hiervan de oorzaak was. Dit is ook de Commissie Elverding opgevallen, de commissie vindt het opvallend dat er achteraf niet structureel aan effectmeting wordt gedaan. Uit de studie van Flyvbjerg blijkt dat er grote systematische en willekeurige afwijkingen kunnen ontstaan. De vraag hoe goed Nederlandse verkeersmodellen het doen is tot nu toe niet goed beantwoord. Wij achten het zinvol om te onderzoeken wat de verschillen zijn en waardoor deze verschillen worden veroorzaakt. Ook het KiM stelt dat afgeronde projecten systematisch moeten worden geëvalueerd (Annema en de Jong, 2008). Het gaat hierbij niet om het achteraf 'afrekenen' maar om onderzoekers en modellers in staat te stellen te leren van het verleden. Het gaat immers om de vervolgvraag: kunnen we de modellen aanpassen zodat we in de toekomst in staat zijn om de modellen beter te maken (zie onderstaand figuur).



Figuur 2: Ex post evaluatie als onderdeel van modelontwikkeling

Stelling: Verkeersmodelleers leren onvoldoende van het verleden. Een model wordt bijna nooit achteraf getoetst. Met deze ontbrekende informatie zouden we de modellen en daarmee besluitvorming over verkeersbeleid kunnen verbeteren.

5. Omgaan met onzekerheden in verkeersprognoses

In de praktijk zie je dat verkeersprognoses nog steeds vaak 'puntschattingen' zijn: de intensiteit op wegvak x is in het jaar 2020 x voertuigen'. Er wordt veelal gebruik gemaakt van één omgevingsscenario (met inschatting bevolkingsgroei, economische groei etc.), één model en één set modelinvoer. En op basis hiervan wordt gemeentelijk, provinciaal en nationaal beleid geformuleerd. Onzekerheden lijken geen rol te spelen. Er worden 'wetenschappelijke' modellen gebruikt en de resultaten zijn harde getallen.

De toekomst is echter intrinsiek onzeker. Welke economische ontwikkeling worden bijvoorbeeld verondersteld en van welk gerealiseerd beleid moet uitgegaan worden? En welke projecten zijn in 2020 wel en niet gerealiseerd?

Vanuit de modelspecialisten, maar ook bijvoorbeeld door de Commissie Elverding en het KiM, wordt geopperd om meer met bandbreedtes te werken. Bandbreedtes vanuit de betrouwbaarheid van een model maar ook vanuit de betrouwbaarheid van de toekomst. Maar dit stelt wel weer eisen aan de wijze hoe hiermee omgegaan moet worden in het beleid.

Stelling: er wordt onvoldoende rekening gehouden met onzekerheden die de toekomst biedt. Nu is een verkeersprognose bijna altijd een puntschatting.

6. Conclusies

De uitkomsten van verkeersmodellen zijn de afgelopen jaren een steeds belangrijkere rol gaan spelen in het besluitvormingsproces over rijksinfrastructuurprojecten. In het besluitvormingsproces zijn resultaten van kosten-batenanalyses en milieustudies doorslaggevend geworden voor het al dan niet aanleggen van infrastructuur. De uitkomsten van verkeersmodellen zijn van grote invloed op de uitkomsten van kosten-batenanalyses en toetsen op luchtkwaliteitsnormen. De waarde die aan verkeersprognoses wordt toegekend is daarmee sterk toegenomen.

Als we naar de vernieuwingen in de modellering van de verkeersvraag in de afgelopen 50 jaar kijken dan zien we dat vooral de wijze van modelleren van de verkeersvraag in de gemeentelijke modellenpraktijk nauwelijks is geëvolueerd. De vraag is hoe erg dit is. Dit hangt uiteraard af van de vraagstelling. Als de trend van gebruik van kosten-batenanalyses zich doorzet naar gemeentelijk of regionaal verkeersbeleid dan is volgens ons een forse verbetering nodig. Het grootste probleem is dat nog volop wordt gewerkt met geaggregeerde verkeersmodellen (met een beperkte of geen bevolkingssegmentatie in de modellen voor vervoerwijze en bestemmingskeuze). In de evolutie van Nederlandse verkeersmodellen valt ook op dat er tot nu toe geen regionale en of gemeentelijke modellen zijn ontwikkeld die verkeer in samenhang met ruimtegebruik bezien, iets wat we in andere landen in Europa wel zien. Ook op nationale schaal worden dergelijke modellen nog niet ingezet bij verkennende studies of planstudies.

Er bestaat geen breed geaccepteerd richtlijnenkader voor het kwalificeren van een verkeersmodel. Verschillende overheden en marktpartijen gebruiken verschillende kaders. De enige richtlijn die in de praktijk wordt gehanteerd is gericht op het benaderen van verkeersstellingen door het verkeersmodel voor het basisjaar. Uiteraard zijn de eisen die je aan een model stelt afhankelijk van het doel van de studie of toepassing. De kwaliteit van modellering moet zijn afgestemd op de diepgang van het project of studie.

Onzekerheden lijken bij veel verkeersprognoses geen rol te spelen. Er wordt veelal gebruik gemaakt van één model, één set modelinvoer en één omgevingsscenario (met inschatting bevolkingsgroei, economische groei etc.). De modellen kennen een onzekerheid, evenals de invoervariabelen (waaronder ook beleidsvariabelen).

Verkeersprognoses zijn intrinsiek onzeker. Wij pleiten ervoor om met minimaal 2 omgevingsscenario's te werken en bandbreedtes in prognoses aan te geven.

De vraag hoe goed Nederlandse verkeersmodellen de daadwerkelijke effecten van infrastructuurprojecten modelleren is tot nu toe niet of nauwelijks beantwoord. Modellen leren niet van het verleden. Uit buitenlands onderzoek blijkt dat er grote systematische en willekeurige afwijkingen tussen prognoses en realiteit kunnen ontstaan. Uiteraard is het zo dat de realiteit er altijd anders uit zal zien dan de prognose maar het prognosticeren van een 'bijpassende' mobiliteitsvraag moet kwalitatief hoogwaardig zijn. Er wordt in Nederland vrijwel nooit gekeken of de prognoses van modellen daadwerkelijk uitkomen en wat we daarvan kunnen leren. Wij pleiten er voor om afgeronde projecten systematisch te evalueren. Het gaat hierbij niet om het achteraf 'afrekenen' op fouten of om het terugdraaien van beslissingen maar om onderzoekers en modelleers in staat te stellen te leren van het verleden. Het doel moet zijn om met informatie uit ex post evaluaties de kwaliteit van verkeersmodellen en -prognoses te verbeteren, en hiermee de besluitvorming over verkeers- en vervoerbeleid.

7. Discussie

Wij pleiten voor structurele ex post evaluaties en een 'lerend' mechanisme richting de modellen. Volgens ons zou voor rijksinfrastructuurprojecten naast een ex ante toets (OEI, OEI bij MIT) ook een ex post evaluatie verplicht moeten worden (zoals in sommige andere landen al het geval is). Het ligt voor de hand de diepgang van een ex post evaluatie af te stemmen op de omvang van de investering. Maar wat is de juiste diepgang? Moet er ook een verplichting komen voor provinciale of gemeentelijke projecten? De vraag is ook wat er in de praktijk zal gebeuren met de resultaten van ex post evaluaties. Kunnen dergelijke evaluaties – in samenhang met het omgaan met onzekerheden in studies – een deel van de huidige beperkingen van verkeersmodellen wegnemen?

Moet een verkeersmodel een kwaliteitsstempel krijgen? Er bestaan geen breed geaccepteerde richtlijnen voor het kwalificeren van een verkeersmodel. Moet er een instantie zijn die modellen certificeert? Of moeten overheidslagen dit zelf weten (gemeenten, provincies). De commissie Elverding spreekt over 'gecertificeerde berekeningen'. Het MNP heeft in het verleden een beoordelingen gegeven van luchtkwaliteitsmodellen (saneringstool). Moet iets dergelijks ook bij verkeersmodellen gebeuren?

Moet gemeentelijk of regionaal verkeersbeleid ook met kosten-batenanalyses worden beoordeeld? En wat betekent dit voor de huidige modellenpraktijk?

In Nederland krijgt de wederzijdse relatie met ruimte weinig aandacht. Er zijn geen regionale en of gemeentelijke grondgebruik-transport-interactiemodel, en op nationale schaal wordt het beschikbare modelinstrument (TIGRIS XL) niet gebruikt bij verkennende studies. Hoe zorgen we ervoor dat op het 40^{ste} CVS verkeer en ruimte in samenhang worden bekeken?

Referenties

- Annema, J.A., M. de Jong (2007) Milieuschattingen in planstudies. Een voorstel tot vereenvoudiging. Kennisinstituut voor Mobiliteitsbeleid, Den Haag.
- Arentze, T.A., H.J.P. Timmermans (2005) Albatross version 2: A learning-Based Transportation Oriented Simulation System. European Institute of Retailing and Services Studies, Eindhoven University, Eindhoven.

- De Jong, G., A. Daly, M. Pieters, S. Miller, R. Plasmeijer, F. Hofman (2007) Uncertainty in traffic forecasts: literature review and new results for The Netherlands. *Transportation*, 34 (4), 375-395.
- de Jong, G., E. Kroes, H. van Mourik, T. van der Hoorn (1998) The impacts of the Amsterdam Ringroad: five years after. AET European Transport Conference, Loughborough University, UK.
- Diederren, H.S.M.A., R.B.A. Koelemeijer (2008) Onzekerheden en complexiteit van de Nederlandse regeling voor luchtkwaliteit. 500144001/2008, Milieu en Natuurplanbureau, Bilthoven.
- DSC/ME&P (1999) Review of land-use/transport interaction models. Reports to The Standing Advisory Committee on Trunk Road Assessment. David Simmonds Consultancy/Marcial Echenique and Partners/Department of the Environment, Transport and the Regions/, London.
- Flyvbjerg, B., M. K. S. Holm, S. L. Buhl (2006) Inaccuracy in traffic forecasts. *Transport Reviews*, 26 (1), 1-24.
- Geurs, K.T., J.A. Annema, H. van Mourik (2007) Analyse van onzekerheden in de verkeerskundige en wagenparkeffecten van de Eerste stap Anders Betalen voor Mobiliteit. Rapport 500076008, Milieu- en Natuurplanbureau/Kennisinstituut voor Mobiliteitsbeleid, Bilthoven/Den Haag.
- Jeannesson-Mange, E., J.N Chapulut, J.P. Taroux (2007) French Post evaluation practice and results. Paper presented at the 86th TRB Annual Meeting, Washington.
- Kroes, E., A. Daly, H. Gunn, T. van der Hoorn (1996) The opening of the Amsterdam ring road: a case study on short-term effects of removing a bottleneck. *Transportation*, 23, 71-82.
- MNP (2007) Nederland Later. Tweede Duurzaamheidsverkenning, deel Fysieke leefomgeving Nederland. Milieu- en Natuurplanbureau, Bilthoven.
- NEI/CE (2001) Lessen uit het verleden. Ex post evaluaties van verkeer- en vervoerbeleid. NEI, Rotterdam.
- van Mourik, H. (2008) Toets op het verkeersmodel Landelijk Model Systeem. Kennisinstituut voor Mobiliteitsbeleid, Den Haag.
- Veldhuisen, K.J., H.J.P. Timmermans, L.L. Kapoen (2005) Simulating the effects of urban development on activity - Travel patterns: An application of Ramblas to the Randstad North Wing. *Environment and Planning B: Planning and Design*, 32 (4), 567-580.
- Zondag, B., M. de Bok, P. Louter, P. van Eikeren, M. Pieters (2007) Toepassen van TIGRIS XL binnen de studie "Nederland Later". Significance, Leiden
- Zondag, B., M. Pieters, J. Baak, P. Louter, P. van Eikeren, M. Wegener (2005) TIGRIS XL. Proeftoepassing. Rand Europe, Leiden.