

**Simulatiestudie naar Methodebreuken in het Onderzoek
Verplaatsingen in Nederland**

Bianca Wouters
Centraal Bureau voor de Statistiek
bias@cbs.nl

Jan van den Brakel
Centraal Bureau voor de Statistiek
jbrl@cbs.nl

**Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk
25 en 26 november 2010, Roermond**

Samenvatting

Simulatiestudie naar Methodebreuken in het Onderzoek Verplaatsingen in Nederland

Vanaf 2010 wordt het Onderzoek Verplaatsingen in Nederland (OViN) uitgevoerd door het CBS. Voorheen werd dit onderzoek uitgevoerd door Rijkswaterstaat onder de naam Mobiliteitsonderzoek Nederland (MON). De wijzigingen in het onderzoeksdesign zullen een methodebreuk veroorzaken. Omdat het MON en het OViN niet gelijktijdig met elkaar zijn uitgevoerd, zal de breuk moeten worden gekwantificeerd met behulp van een structureel tijdreeksmodel. Omdat er onzekerheid is over de betrouwbaarheid van de data in 2008 en 2009 en er een gedeelte van de gegevens in 2009 ontbreekt, is de robuustheid van de breukschattingen met betrekking tot de vertekening van data en de ontbrekende data onderzocht.

Hiertoe is een simulatiestudie uitgevoerd. De voornaamste variabelen van het OViN worden beschouwd, te weten de totale vervoersprestatie in miljard kilometer per jaar en het aantal verplaatsingen per persoon per dag voor personen vanaf 12 jaar. Voor deze variabelen zijn er verschillende scenario's gedefinieerd voor wat betreft een aantal aannames. Allereerst zijn er verschillende waarden gekozen voor de periode van de beschikbare reeks en de grootte van de methodebreuk vanaf 2010. Voor elke combinatie van periode en breuk is er onderscheid gemaakt met betrekking tot het al dan niet toevoegen van vertekening in de jaren 2008 en 2009. Bovendien is er een scenario gekozen waarin de vertekening niet alleen wordt toegevoegd, maar ook wordt meegenomen in het simulatiemodel. Tenslotte worden de data in 2008 en 2009 ofwel meegenomen ofwel worden ze als volledig ontbrekend beschouwd.

De modelschattingen van de methodebreuken blijken inderdaad gevoelig te zijn voor minder betrouwbare data in de jaren die vooraf gaan aan het herontwerp. Dit kan echter worden ondervangen door de verminderde datakwaliteit van 2008 en 2009 expliciet te modelleren. Ook valt het te overwegen om deze jaren volledig weg te laten. Er moet daarom duidelijkheid komen over de betrouwbaarheid van de data in 2008 en 2009 ten opzichte van de voorgaande jaren. Deze informatie is cruciaal om het juiste tijdreeksmodel te kunnen selecteren. Het te behalen detailniveau van een eventuele correctie moet in een later stadium worden bepaald, wanneer er voldoende data beschikbaar zijn.

1. Inleiding

Sinds meer dan 30 jaar wordt in Nederland het zogeheten verplaatsingsonderzoek uitgevoerd. Eerst werd het onderzoek door het CBS uitgevoerd, vervolgens is het overgenomen door Rijkswaterstaat in samenwerking met Socialdata om vanaf 2010 weer door het CBS te worden gedaan. Het onderzoek levert informatie over de tijdstippen waarop de Nederlandse bevolking deelneemt aan het verkeer en op welke wijze. Belangrijke variabelen zijn de totale vervoersprestatie in kilometer per jaar en het aantal verplaatsingen per persoon per dag. Vanaf de aanvang van het onderzoek zijn er tijdreeksen bijgehouden voor dergelijke variabelen.

De verandering van onderzoeksdesign in 2010 zal onvermijdelijk leiden tot zogeheten methodebreuken in dat jaar. Er is sprake van een methodebreuk in een tijdreeks wanneer veranderingen in het surveyproces systematische effecten hebben op de uitkomsten. Dergelijke breuken verstoren de continuïteit van de tijdreeksen die de afgelopen decennia met het verplaatsingonderzoek zijn opgebouwd. Het is daarom van belang om deze methodebreuken zo goed mogelijk te kwantificeren. In een eerder stadium is besloten om het MON en het OViN niet gedurende een bepaalde periode parallel aan elkaar uit te voeren. Daarom zullen de methodebreuken moeten worden geschat aan de hand van tijdreeksmodellen. Het detailniveau en de kwaliteit van de schatting van de breuken is afhankelijk van een aantal factoren waaronder de kwaliteit van de historische cijfers. Doordat er twijfel is over de kwaliteit van de data van 2008 en er onzekerheid is over de aanwezigheid en de kwaliteit van de data van 2009 is het de vraag of het mogelijk is om een accurate schatting te maken van de breuken die ontstaan bij de overgang op het OViN in 2010. Om deze reden zijn er simulaties uitgevoerd om de robuustheid van de breukschattingen met betrekking tot de vertekening van data en de ontbrekende data in 2008 en 2009 te onderzoeken. Tevens bediscussiëren we het te behalen detailniveau.

In paragraaf 2 lichten we het doel en de geschiedenis van het onderzoek nader toe en bespreken we de wijzigingen in het onderzoeksdesign in 2010. In paragraaf 3 geven we uitleg over structurele tijdreeksmodellen. Paragraaf 4 gaat verder in op de gedane simulaties en in paragraaf 5 bespreken we de resultaten. Tot slot bevat paragraaf 6 een conclusie.

2. Het Onderzoek Verplaatsingen in Nederland

2.1 Beschrijving en historie van het onderzoek

Bij het verplaatsingsonderzoek wordt aan mensen gevraagd om voor één dag bij te houden waar ze die dag heen gaan en met welke vervoermiddelen. Met deze informatie wordt de mobiliteit van de Nederlandse bevolking beschreven en met deze kennis wordt vervolgens het verkeers- en vervoersbeleid gevormd. De resultaten worden bijvoorbeeld gebruikt door de Rijksoverheid voor het verbeteren van de verkeersveiligheid en het openbaar vervoer, het milieubeleid en het bestrijden van files. Ook de gemeenten en provincies gebruiken de informatie voor hun beleid en voor de verbetering van de verkeers- en vervoerssituatie.

Sinds 1978 worden jaarlijks enquêtes uitgevoerd om het verplaatsingsgedrag van de Nederlandse bevolking te onderzoeken. Oorspronkelijk werd dit onderzoek door het CBS uitgevoerd onder de naam Onderzoek Verplaatsingsgedrag (OVG). In 1985 is er een eerste methodebreuk ontstaan welke tot op heden moeilijk te herstellen is. In de beginperiode van het onderzoek is het onderzoeksdesign namelijk jaarlijks aangepast. Bovendien is er onvoldoende documentatie aanwezig over de wijzigingen. Daarom zijn de tijdreeksen pas beschikbaar vanaf 1985. Zie Konen en Swinkels (2000) voor een verdere

uitwijding over deze breuk. Vervolgens is het onderzoek in 2004 overgenomen door de Adviesdienst Verkeer en Vervoer van het ministerie van Verkeer en Waterstaat in samenwerking met Socialdata onder de naam Mobiliteitsonderzoek Nederland (MON). Meer informatie over het onderzoek is te vinden in Molnár-in 't Veld (2007) en projectteam MON (2008).

Momenteel zijn de cijfers gepubliceerd door het CBS tot en met het jaar 2007. De cijfers van 2008 zijn weliswaar beschikbaar, maar er is enige twijfel omtrent de plausibiliteit van de cijfers van dat jaar. Neem bijvoorbeeld de tijdreeks van de totale vervoersprestatie. Deze heeft tot en met 2007 een duidelijk stijgende trend vertoond. In het jaar 2008 is er een relatief grote afname van de totale vervoersprestatie te zien ten opzichte van 2007. Dit ligt niet in de lijn der verwachting aangezien hulpbronnen juist een verdere toename vertonen. Het gaat hier om variabelen uit andere statistieken die een sterke correlatie hebben met de totale vervoersprestatie. Bijvoorbeeld in de tijdreeks van de totale verkeersprestatie van Nederlandse voertuigen wordt de stijging in 2008 verder doorgetrokken, ook naar verschillende voertuigtypen en grondgebieden. Dit roept twijfels op over de betrouwbaarheid van de cijfers van het MON in 2008.

De data van 2009 zullen eind 2010 beschikbaar komen en zouden meer duidelijkheid kunnen geven over het verdere verloop van de MON tijdreeksen. Echter de data van 2009 zijn onvolledig. In dit jaar zijn er slechts gegevens beschikbaar over de maanden januari tot en met de derde week van november. De laatste vijf weken van 2009 ontbreken omdat er in deze periode geen respondenten meer zijn benaderd. Bovendien is het risico groot dat er ook bij deze cijfers twijfel over de kwaliteit ontstaat. De overgang van beschikbare data eind 2009 naar het niet beschikbaar hebben van data in de laatste weken van 2009 is mogelijk geleidelijk verlopen. Verwacht kan worden dat er een overgangperiode is in de laatste veldwerkperiode van het MON waarin de data van minder kwaliteit is.

In 2010 is tenslotte het onderzoek weer door het CBS overgenomen onder de naam Onderzoek Verplaatsingen in Nederland (OViN). Rijkswaterstaat heeft tot samenwerking met het CBS besloten op basis van een pilot die door het CBS is uitgevoerd in 2008 om aan te tonen dat het onderzoek gedeeltelijk ook via internetwaarneming kan worden uitgevoerd, zie Janssen e.a. (2008).

2.2 Methodebreuken in 2010

Bij de overgang van het MON op het OViN in 2010 wordt er een nieuw onderzoeksdesign geïntroduceerd welk in meerdere opzichten verschilt met het design van het MON. Ten eerste is de benaderingsstrategie aangepast. De papieren vragenlijsten maken plaats voor een mixed-mode benadering. Hierbij wordt er eerst een internetwaarneming gedaan. Wanneer er geen respons wordt verkregen, wordt deze gevolgd door telefonische benadering indien er een telefoonnummer beschikbaar is. Wanneer er geen telefoonnummer beschikbaar is, wordt er persoonlijk benaderd. Ten tweede verschilt de onderzoekseenheid tussen het MON en het OViN. De eenheid is gewijzigd van huishoudens naar personen. Verder wordt er gebruik gemaakt van een aangepaste vragenlijst en wordt er extra informatie over een aantal variabelen verzameld uit registers. Zie Bouhuijs (2010) voor verdere beschrijving van de wijzigingen.

Als gevolg van deze wijzigingen zullen er in 2010 methodebreuken ontstaan in de tijdreeksen. Het is van belang om zo goed mogelijk inzicht krijgen in de omvang van deze breuken. Idealiter worden het oude MON en het nieuwe OViN gedurende een aantal maanden parallel aan elkaar uitgevoerd om de methodebreuk zo goed mogelijk te schatten en te corrigeren. Omdat er ditmaal niet is parallel gedraaid, kunnen schattingen worden gemaakt van de methodebreuken van de hoofdvariabelen aan de hand van tijdreeksmodellen. Echter vanwege de problematiek rond de jaren 2008 en 2009

ontstaan er vraagtekens rondom de gevoeligheid van de schattingen op basis van een tijdreeksmodel voor verminderde datakwaliteit in de jaren voorafgaand aan de overgang naar het nieuwe design. Er is daarom een uitgebreidere analyse vereist naar de mogelijkheden van het schatten van de methodebreuken gegeven de databestanden die beschikbaar zijn vanaf 1985 tot en met 2008 en diens kwaliteit. Voor deze periode zijn er tijdreeksmodellen geschat op basis van puntschattingen voor een aantal doelvariabelen. Er zijn simulaties uitgevoerd voor een aantal scenario's met betrekking tot de betrouwbaarheid en aanwezigheid van de data in 2008 en 2009. Aan de hand van de resultaten is de robuustheid van de modelschattingen voor de methodebreuken onder de invloed van de vertekening van data en de ontbrekende data onderzocht.

3. Structurele tijdreeksmodellen

We spreken van een methodebreuk in een tijdreeks wanneer wijzigingen in het statistische proces van zodanige aard zijn dat ze systematische invloed hebben op de publicatiecijfers (Van den Brakel, Kraan en Roels, 2010). Om de continuïteit van opgebouwde reeksen te behouden, is het van belang dat de omvang van een methodebreuk gekwantificeerd wordt. De manier waarop dit gebeurt, wordt mede bepaald door de aard en de plaats van de veranderingen in het surveyproces. Indien wijzigingen plaats vinden in classificaties waarnaar gepubliceerd wordt of wijzigingen worden doorgevoerd in de verwerking kunnen de reeksen worden herberekend met behulp van de bestaande data zoals beschreven in Van den Brakel (2010). Wanneer er (zoals bij het OViN het geval is) sprake is van veranderingen in het veldwerk, kan de breuk het beste worden gekwantificeerd wanneer er gedurende een periode parallel wordt gedraaid met het oude en het nieuwe ontwerp. Bij de overgang van het MON op het OViN is er echter niet dubbel gedraaid. In een dergelijk geval kan er worden getracht om de methodebreuk te schatten met behulp van een zogeheten interventieanalyse aan de hand van een structureel tijdreeksmodel. Dit wordt voorgesteld in Van den Brakel en Roels (2010) en Van den Brakel, Smith and Compton (2008) en is een directe toepassing van de interventie aanpak voorgesteld door Harvey en Durbin (1986).

Met een structureel tijdreeksmodel wordt de reeks beschouwd als zijnde opgebouwd uit een trend component, een seizoenscomponent, andere cyclische componenten, een regressiecomponent en een storingsterm. Voor elke component wordt een stochastisch model aangenomen, zodat de componenten tijdsafhankelijk kunnen zijn. Indien nodig kunnen ARMA (Autoregressive – Moving Average model) componenten worden toegevoegd om eventuele autocorrelatie tussen storingstermen te modelleren. In een interventieanalyse wordt in de regressiecomponent een interventievariabele toegevoegd. De interventievariabele is een soort stapfunctie waarmee wordt aangegeven op welk tijdstip er wordt overgegaan op het nieuwe ontwerp. De regressiecoëfficiënt kan dan worden geïnterpreteerd als de omvang van de methodebreuk ten gevolge van de overgang van het oude naar het nieuwe ontwerp. Dit kunnen we stellen onder twee aannames. Ten eerste wordt aangenomen dat de werkelijke tijdreeks goed wordt benaderd door het tijdreeksmodel. Ten tweede wordt verondersteld dat op het tijdstip dat wordt overgegaan op het nieuwe ontwerp (ingeval het OViN in 2010) de werkelijke ontwikkeling niet afwijkt van het veronderstelde tijdreeksmodel.

Stel \hat{y}_t is de geschatte doelvariabele in het OVG, MON of OViN op tijdstip t . De vorm van het structureel tijdreeksmodel is afhankelijk van de betrouwbaarheid van de data in 2008 en 2009. Ofwel er wordt aangenomen dat deze vergelijkbaar is met voorgaande jaren, ofwel de verminderde datakwaliteit wordt expliciet gemodelleerd in de simulaties.

In het eerste geval, waarbij ervan wordt uitgegaan dat de betrouwbaarheid van de data in 2008 en 2009 vergelijkbaar is met die van de voorgaande jaren, wordt het structureel tijdreeksmodel als volgt geschreven

$$\hat{y}_t = L_t + \beta\delta_t + \varepsilon_t \quad (1)$$

met L_t de trend op tijdstip t (het niveau), δ_t een interventievariable die aangeeft of het het oude of het nieuwe ontwerp betreft ofwel

$$\delta_t = \begin{cases} 1 & (t \geq 2010) \\ 0 & (t < 2010) \end{cases} \quad (2)$$

β de regressiecoëfficiënt voor de interventievariabele en ε_t de storingsterm. Dit tijdreeksmodel volstaat omdat er geen overlap is tussen de jaarlijkse steekproeven en er dus geen correlatie bestaat tussen de verschillende steekproeffouten. We nemen aan dat de storingstermen onderling onafhankelijk zijn en normaal verdeeld zijn volgens

$$\varepsilon_t \cong N\left(0, \frac{\sigma_\varepsilon^2}{n_t}\right) \quad (3)$$

met n_t de steekproefomvang (aantal huishoudens). Door de variantie omgekeerd evenredig te kiezen met de steekproefomvang wordt er rekening gehouden met sterk fluctuerende steekproefomvang. Er wordt met dit variantiemodel geen rekening gehouden met lokale oversampling die in verschillende jaren zijn toegepast. Zoals beschreven in paragraaf 2.2 gaat het OViN van een huishoudsteekproef over op een personensteekproef. In (3) is n_t vanaf 2010 gelijk aan het aantal responderende personen in plaats van huishoudens. De variantiestructuur van het model kan worden verbeterd door de variantie in (3) evenredig te kiezen aan de variantie van de geschatte doelvariabele. Zie Van den Brakel en Roels (2010) voor technische details.

De stochastische trend wordt gemodelleerd met een zogeheten smooth trend model

$$\begin{aligned} L_t &= L_{t-1} + R_{t-1} \\ R_t &= R_{t-1} + \eta_t \end{aligned} \quad (4)$$

met R_t de stochastische hellingsparameter en η_t storingstermen waarvan verondersteld wordt dat ze onderling onafhankelijk zijn en normaal verdeeld volgens

$$\eta_t \cong N(0, \sigma_\eta^2) \quad (5)$$

Het tijdreeksmodel kan in toestandsruimteform, ook wel state space form genoemd, worden geschreven. Vervolgens kunnen schattingen worden gemaakt van de onbekende parameters en de doelvariabelen met behulp van het Kalman filter zoals beschreven in Durbin en Koopman (2001) en Harvey (1989).

Wanneer, in het tweede geval, de verminderde kwaliteit van de data in 2008 en 2009 expliciet wordt gemodelleerd, wordt dat gedaan met behulp van een tweede interventievariabele. Het structureel tijdreeksmodel wordt uitgebreid met een term voor de vertekening in 2008 en 2009:

$$\hat{y}_t = L_t + \beta_1\delta_t^1 + \beta_2\delta_t^2 + \varepsilon_t \quad (6)$$

met

$$\delta_t^1 = \begin{cases} 1 & (t \in [2008, 2009]) \\ 0 & (t \notin [2008, 2009]) \end{cases} \quad (7)$$

en

$$\delta_t^2 = \begin{cases} 1 & (t \geq 2010) \\ 0 & (t < 2010) \end{cases} \quad (8)$$

Hier is L_t de trend gemodelleerd volgens het stochastisch trendmodel (4) en (5), δ_t^1 en β_1 zijn de interventievariable en de regressiecoëfficiënt voor de vertekening in 2008 en 2009, δ_t^2 en β_2 zijn de interventievariable en de regressiecoëfficiënt voor de breuk in 2010 en ε_t is de storingsterm. Merk op dat hier de veronderstelling is gemaakt dat de

vertekeningen in de jaren 2008 en 2009 gelijk zijn. Men zou kunnen overwegen om het model verder uit te breiden door aparte interventievariabelen te nemen voor de vertekening in 2008 en 2009. In dit onderzoek wordt dit uitgebreide model niet beschouwd. In één van de simulatiescenario's wordt namelijk een model gebruikt waarin de waarden in 2008 en 2009 als missing worden aangenomen. Aangezien deze twee modellen gelijk zijn, heeft het gebruik van twee aparte interventievariabelen geen toegevoegde waarde.

Ook dit tijdreeksmodel wordt in toestandsruimteform geschreven waarna schattingen worden gemaakt van de onbekende parameters en de doelvariabelen met behulp van het Kalman filter.

4. Simulaties

We beschouwen de voornaamste variabelen van het OViN, te weten de totale vervoersprestatie in miljard kilometer per jaar en het aantal verplaatsingen per persoon per dag voor personen vanaf 12 jaar. Voor iedere variabele worden de volgende stappen uitgevoerd.

Allereerst wordt een model gefit op de tijdreeks zover beschikbaar tot en met 2008 en maken een schatting van de onbekende parameters van dit model.

$$\hat{y}_t = L_t + \beta\delta_t + \varepsilon_t \quad (9)$$

met

$$\delta_t = \begin{cases} 1 & (t = 2008) \\ 0 & (t < 2008) \end{cases}. \quad (10)$$

Het model bestaat uit een trend component en een interventie voor de vertekening in 2008 ten gevolge van de verminderde data kwaliteit in dat jaar. Dit geeft een schatting van de onbekende modelparameters voor de trend en de vertekening in 2008. Aan de hand van deze schattingen wordt het volgende trendmodel gedefinieerd:

$$\hat{y}_t = L_t + \varepsilon_t \quad (11)$$

met L_t gedefinieerd in (4) en (5).

Dit model specificeert een verdeling waaruit een onbeperkt aantal tijdreeksen van uiteenlopende lengtes kunnen worden getrokken. In iedere reeks kan vervolgens een vertekening worden toegevoegd aan de waarde in 2008 en 2009 of een waarneming als missing worden aangenomen. Tevens kan een methodebreuk worden toegevoegd vanaf het jaar 2010 ten gevolge van de overgang naar het OViN. Vervolgens worden deze reeksen geanalyseerd met de modellen beschreven in paragraaf 3. Op deze wijze kan voor verschillende scenario's onderzocht worden hoe goed een tijdreeksmodel een vooraf ingestelde methodebreuk terugvindt. Zie Van den Brakel en Roels (2010) voor meer technische details van een dergelijke simulatie studie. Samengevat hebben we de volgende te variëren parameters:

1. De lengte van de beschikbare reeks: we beschouwen de periode $[1985, t_i]$ met $t_1=2010$, $t_2=2011$ en $t_3=2012$.
2. De grootte van de methodebreuk vanaf het jaar 2010: de breuk is gelijk aan b_i met $i = 1, 2, 3$.
3. Er is wel/geen vertekening in het jaar 2008 en/of 2009.
4. Er zijn wel/geen ontbrekende data in het jaar 2008 en/of 2009.

De eerste twee punten te weten de periode van de reeks en de grootte van de methodebreuk leveren negen mogelijke combinaties op. Voor ieder van deze combinaties worden de vertekening in de data en de ontbrekende data in een vijftal simulatiescenario's verwerkt. In ieder scenario wordt de breuk vanaf 2010 meegenomen

in het simulatiemodel. Wanneer er vertekening wordt toegevoegd in 2008 en/of 2009, wordt de waarde genomen die volgt uit de analyse van de waargenomen reeks tot en met 2008 met model (9) in de eerste stap van de simulatie.

De volgende scenario's zijn onderzocht:

1. Er wordt geen vertekening toegevoegd in 2008 en 2009. Dit is de ideale situatie om een methodebreuk te schatten via een interventiemodel.
2. Er wordt wel vertekening toegevoegd aan de gesimuleerde data in 2008 en 2009.
3. Er wordt alleen vertekening toegevoegd aan de gesimuleerde data in 2008. De data in 2009 worden als missing aangenomen.
4. De data in 2008 en 2009 worden als missing aangenomen.
5. Er wordt vertekening toegevoegd aan de gesimuleerde data in 2008 en 2009. In het simulatiemodel wordt niet alleen de breuk in 2010 meegenomen, maar ook de vertekening in 2008 en 2009.

Voor de scenario's 1 tot en met 4 wordt het interventiemodel (1) gehanteerd met de interventievariabele beschreven in formule (2) uit paragraaf 3.

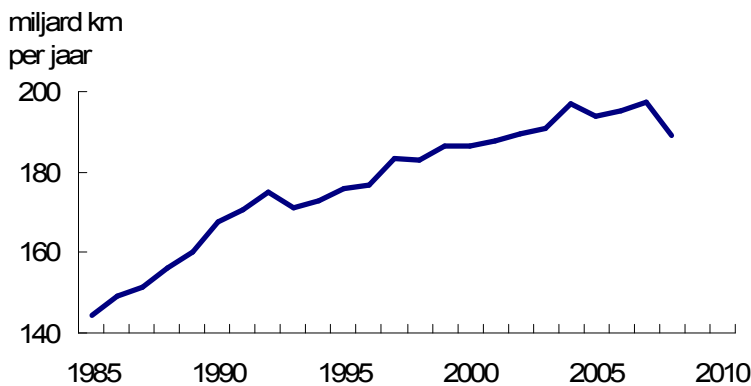
Scenario 5 heeft dezelfde input als scenario 2, maar er wordt een uitgebreider interventiemodel gehanteerd door de vertekening in 2008 en 2009 mee te nemen in het model zoals beschreven in de formules (6), (7) en (8).

Daarmee zijn er per variabele 45 scenario's. Voor ieder scenario zijn er telkens 5000 tijdreeksen gegenereerd volgens de hierboven beschreven procedure. Iedere geresampelde reeks wordt geanalyseerd met een tijdreeksmodel en levert een schatting voor σ_ε^2 , σ_η^2 , een schatting van de methodebreuk vanaf 2010 en diens standaardfout en tenslotte indien van toepassing (simulatiescenario 5) een schatting van de vertekening in 2008 en 2009 en diens standaardfout. De verdeling van deze 5000 resamples geeft een goed beeld van de mate waarin de vooraf ingestelde waarden voor de methodebreuk vanaf 2010 onder de verschillende scenario's worden teruggevonden.

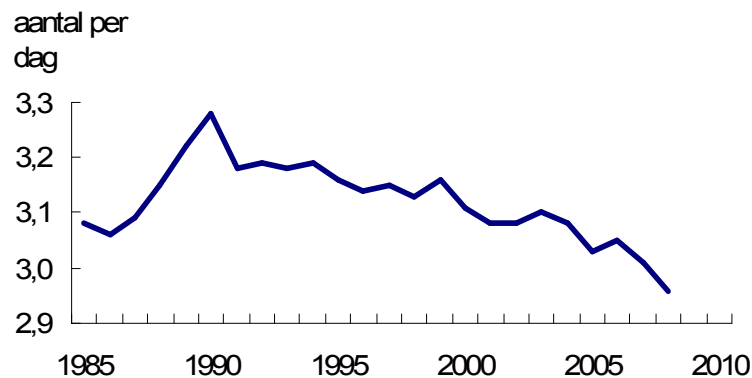
5. Resultaten

In deze paragraaf worden de resultaten besproken voor de variabelen totale vervoersprestatie en het aantal verplaatsingen per persoon per dag voor personen vanaf 12 jaar. In Figuur 1 zijn de beschikbare tijdreeksen voor deze variabelen weergegeven.

Totale Vervoersprestatie



Verplaatsingen per persoon



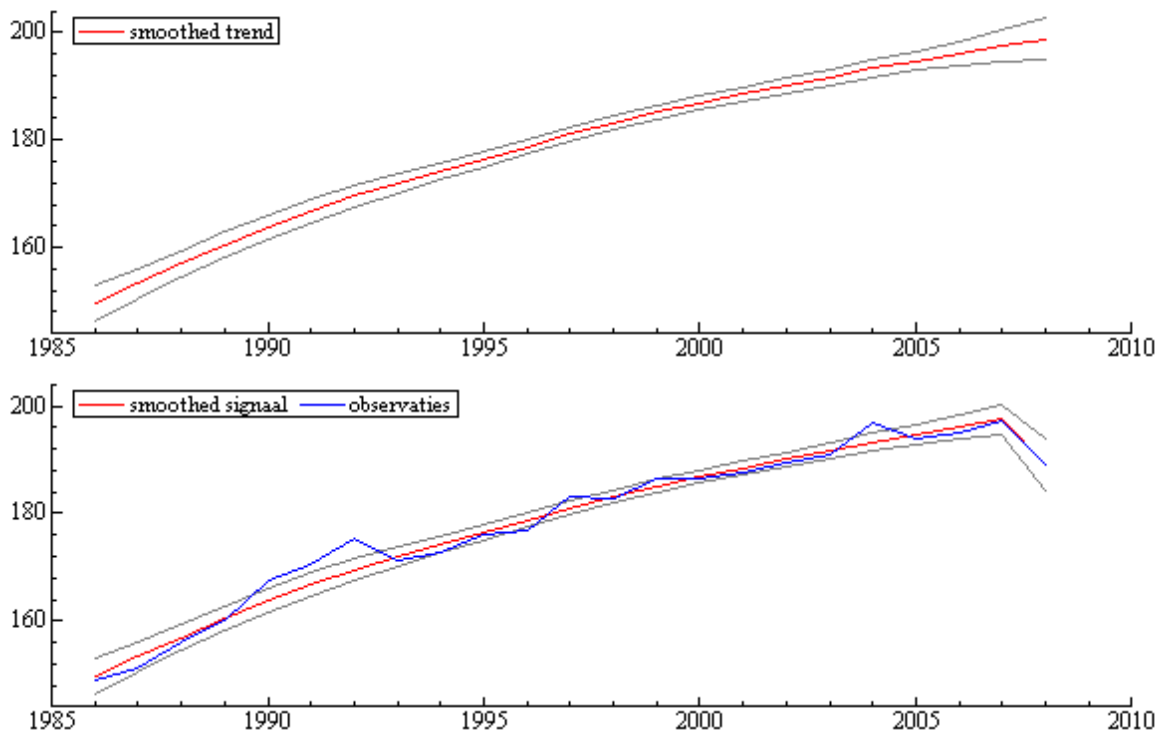
Figuur 1. Beschikbare tijdreeksen van de beschouwde doelvariabelen.

In Tabel 1 zien we de resultaten van de eerste analysestap op basis van model (9). Voor de totale vervoersprestatie per jaar wordt er een vertekening vanaf 2008 geconstateerd van -9,7 met een marge van 3,17. Voor het aantal verplaatsingen per dag is de vertekening -0,04 met een marge van 0,04.

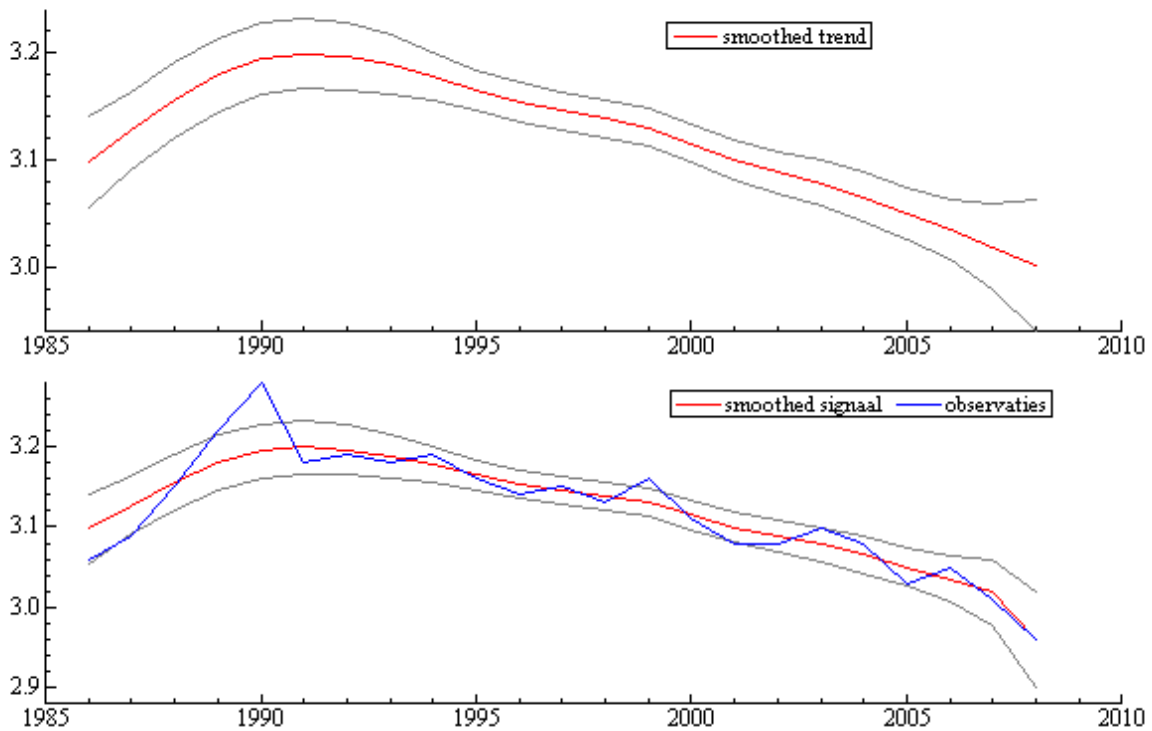
Tabel 1. Resultaten van de schattingen voor beide doelvariabelen.

	vervoer	verplaatsingen
σ_{η}	0,432	0,0096
σ_{ε}	329	4,05
$\hat{\beta}_{2008}$	-9,7	-0,04
$\sigma_{\hat{\beta}_{2008}}$	3,17	0,04

In Figuur 2 en Figuur 3 zijn de resultaten van beide variabelen weergegeven. Naast de (gesmoothde) schattingen van de trend wordt het zogeheten signaal getoond welk gelijk is aan de trend plus de geschatte vertekening. De marges van de trend en het signaal zijn in de figuur toegevoegd.



Figuur 2. Schatting van de trend en het signaal van de totale vervoersprestatie per jaar.



Figuur 3. Schatting van de trend en het signaal van het aantal verplaatsingen per persoon per dag.

De schattingen van de parameters in Tabel 1 worden gebruikt om een tijdreeksmodel te specificeren waarmee in de simulaties herhaald tijdreeksen gegenereerd worden. Iedere reeks wordt geanalyseerd met één van de modellen uit paragraaf 3. De gemiddelden en standaardfouten van de breukschattingen worden getoond in Tabel 2 en Tabel 3. De kolom 'breukschatting' toont het gemiddelde over de 5000 simulaties van de geschatte breuk per simulatie. De kolom 'standaardfout' toont zowel de modelstandaardfout als de gesimuleerde standaardfout. De modelstandaardfout wordt berekend door de standaarddeviatie te nemen van de 5000 schattingen per simulatie voor de standaardfout van de breuk. De gesimuleerde standaardfout wordt berekend door het gemiddelde te nemen van de 5000 standaardfouten die per simulatie door het model zijn bepaald. De standaardfouten geven een indicatie van welke breuken met deze methodiek waarneembaar zijn. Breuken waarvan de absolute waarde groter is dan twee keer de standaardfout zijn significant waarneembaar met een significantieniveau van 5%. Een vergelijk van gesimuleerde standaardfout met modelstandaardfout laat zien dat het model de standaardfout enigszins onderschat. Verder is binnen ieder scenario de gemiddelde standaardfout voor de drie breuken hetzelfde. Dit komt doordat er voor iedere simulatie uit wordt gegaan van dezelfde set van 5000 tijdreeksen. Omdat binnen één scenario alleen gevarieerd wordt met verschillende waarden voor de methodebreuk, zijn de geschatte standaardfouten identiek.

Tabel 2. Gemiddelde van de breukschattingen en van de standaardfout van de breuk in de totale vervoersprestatie.

scenario	breukschatting			standaardfout ¹	
	$b_1=-25$	$b_2=0$	$b_3=25$	sim	model
gesimuleerde reeks tot en met 2010					
1	-25,04	-0,04	24,96	2,73	2,41
2	-14,91	10,09	35,09	3,76	3,06
3	-18,17	6,83	31,83	4,69	3,55
4	-25,09	-0,09	24,91	4,09	3,44
5	-25,11	-0,11	24,89	4,16	3,99
gesimuleerde reeks tot en met 2011					
1	-25,03	-0,02	24,98	2,79	2,32
2	-15,32	9,62	34,62	3,30	2,95
3	-18,81	6,19	31,19	3,96	3,30
4	-25,05	-0,05	24,95	4,11	3,27
5	-25,05	-0,05	24,95	4,14	3,74
gesimuleerde reeks tot en met 2012					
1	-25,00	0,00	25,00	2,71	2,26
2	-16,23	8,77	33,77	3,12	2,81
3	-19,53	5,48	30,48	3,57	3,11
4	-25,05	-0,05	24,95	3,71	3,07
5	-25,06	-0,06	24,94	3,93	3,60

¹ De standaardfout van b_i is gelijk voor verschillende i .

Tabel 3. Gemiddelde van de breukschattingen en van de standaardfout van de breuk in aantal verplaatsingen.

scenario	breukschatting			standaardfout	
	$b_1=-0,5$	$b_2=0$	$b_3=0,5$	sim	model
gesimuleerde reeks tot en met 2010					
1	-0,501	0,001	0,499	0,039	0,035
2	-0,454	0,046	0,546	0,040	0,036
3	-0,462	0,038	0,538	0,054	0,048
4	-0,501	-0,002	0,498	0,068	0,059
5	-0,502	-0,002	0,498	0,067	0,061
gesimuleerde reeks tot en met 2011					
1	-0,500	0,000	0,500	0,040	0,034
2	-0,456	0,044	0,544	0,040	0,035
3	-0,467	0,033	0,533	0,052	0,044
4	-0,501	-0,001	0,499	0,063	0,053
5	-0,501	-0,001	0,499	0,065	0,059
gesimuleerde reeks tot en met 2012					
1	-0,500	0,000	0,500	0,037	0,032
2	-0,460	0,033	0,540	0,037	0,033
3	-0,471	0,029	0,529	0,046	0,041
4	-0,500	-0,000	0,496	0,053	0,049
5	-0,501	-0,001	0,499	0,058	0,055

Om te beginnen is te zien dat met scenario 1 in staat zijn om een goede schatting van de breuk te maken. Voor de totale vervoersprestatie wordt bijvoorbeeld de breuk $b_3=-25$ geschat op 24,96 bij een reeks tot en met 2010. In dit scenario wordt er geen vertekening toegevoegd onder de aanname dat de jaren 2008 en 2009 van gelijke betrouwbaarheid zijn als de voorgaande jaren. Onder dit gegeven is de reeks goed te modelleren en de breuk vanaf 2010 logischerwijs goed te schatten.

Het is echter waarschijnlijker dat de data in 2008 en 2009 van mindere kwaliteit zijn. In scenario 2 wordt het effect hiervan op de schattingen van de methodebreuken onderzocht door vertekening toe te voegen aan de gesimuleerde data in de jaren 2008 en 2009. Het is te zien dat de schattingen ver naast het doel schieten. Bijvoorbeeld voor de totale vervoersprestatie wordt de breuk $b_2=0$ zelfs voor de reeks tot en met 2010 nog geschat op 8.72. Door de verminderde kwaliteit in 2008 en 2009 is er geen goede schatting meer van de methodebreuk te maken. Dit wordt veroorzaakt door het feit dat het model ervan uitgaat dat de data in 2008 en 2009 van vergelijkbare kwaliteit zijn ten opzichte van de voorgaande jaren. Het model veronderstelt dat de daling die vanaf 2008 inzet een reële ontwikkeling is.

Ook in scenario 3 wordt aangenomen dat de data in 2008 van mindere kwaliteit zijn. De data in 2009 worden echter weggelaten. De schatting wordt iets beter, maar zit er nog ver naast. Het lijkt beter te zijn om helemaal geen data in 2009 te gebruiken dan data van mindere kwaliteit. We zien echter ook de gemiddelde standaardfout van de geschatte breuk toenemen.

Wanneer de data in 2008 en 2009 volledig worden weggelaten zoals in scenario 4, is de gemiddelde schatting van de breuk zeer goed te noemen. Maar ook hier heeft de schatting een zeer grote gemiddelde standaardfout.

In het laatste scenario wordt net als in scenario 2 de vertekening toegevoegd aan de gesimuleerde data in 2008 en 2009 en wordt deze tevens meegenomen in het simulatiemodel. Ook dit scenario geeft zeer goede schattingen van de methodebreuk in 2010 met een grotere fout. De standaardfout is echter groter dan in scenario 4, wat doet vermoeden dat het beter is om de data weg te laten dan deze expliciet te modelleren. Het is een opmerkelijk resultaat dat het weglaten van informatie resulteert in nauwkeurigere schattingen voor de methodebreuk. Dit resultaat heeft vermoedelijk twee oorzaken. Ten eerste is aan het tijdreeksmodel een extra parameter toegevoegd om de vertekening voor de jaren 2008 en 2009 te beschrijven. Deze parameter moet worden geschat waardoor de variantie van de overige parameterschattingen zal toenemen. Ten tweede zijn de 5000 reeksen gegenereerd uit een verdeling die wordt gespecificeerd met model (11) en geparametriseerd met de gegevens uit Tabel 1. Het schatten van een breuk in reeksen waar twee jaren ontbreken gaat wellicht beter omdat de reeksen uit dezelfde verdeling worden gegenereerd. Dit resultaat is niet zondermeer te generaliseren naar de praktijk van het OViN omdat de waargenomen reeks niet perfect voldoet aan een vooraf gespecificeerd model.

Verder kan er worden gesteld dat wanneer ook de data in 2011 en 2012 worden meegenomen, de schatting steeds beter wordt en vooral de standaardfout afneemt. Met slechts één OViN jaar voorhanden is het moeilijk om een betrouwbare schatting te maken voor de methodebreuk. Maar deze schatting zal waarschijnlijk steeds beter zijn naarmate er meer onderzoeksjaren volgens het nieuwe OViN design beschikbaar komen. En nog veel belangrijker is het om betrouwbare en volledige data ter beschikking te hebben van 2008 en 2009.

Deze afwegingen gelden vanzelfsprekend voor de gekozen breuken b_i . Het onderzoek zou verder kunnen worden uitgebreid door te kijken naar de invloed van de keuze van een grotere of kleinere breuk. Dit heeft echter weinig toegevoegde waarde, omdat de marges onveranderd blijven bij verschillende breuken.

6. Conclusies en aanbevelingen

Er is een simulatiestudie uitgevoerd om de robuustheid van de breukschattingen met betrekking tot de vertekening van data en de ontbrekende data van het OViN te onderzoeken. We kunnen het volgende concluderen:

- Het blijkt dat de schattingen zeer gevoelig zijn voor de verminderde kwaliteit van de gegevens, indien een model gebruikt wordt dat er (onterecht) vanuit gaat dat de betrouwbaarheid van de data in 2008 en 2009 vergelijkbaar is met voorgaande jaren.
- Het model werkt goed als 2008 en 2009 als missing worden aangenomen.
- Het model werkt ook goed als de verminderde datakwaliteit expliciet gemodelleerd wordt. Dit resultaat geldt echter alleen onder de aanname dat de vertekening in 2008 en 2009 gelijk zijn. Indien deze veronderstelling niet

plausibel is, is het wellicht beter om informatie uit 2008 en 2009 op missing te zetten.

Het is daarom van belang dat op basis van externe informatie een beslissing wordt genomen over de betrouwbaarheid van 2008 en 2009. Op basis hiervan kan een geschikt tijdreeksmodel geselecteerd worden om een zo goed mogelijke schatting voor de methodebreuken te maken.

Bij het schatten van methodebreuken via een tijdreeksmodel wordt verondersteld dat de ontwikkeling van de doelvariabele niet afwijkt van het veronderstelde tijdreeksmodel in het jaar waarop het surveyproces gewijzigd wordt. Systematische afwijkingen van de echte ontwikkeling ten opzichte van het veronderstelde tijdreeksmodel, worden door het tijdreeksmodel onterecht opgevat als een methodebreuk. Wanneer de data in 2008 en 2009 als missing worden aangenomen, wordt tevens de veronderstelling gemaakt dat de trend van de voorgaande jaren zich voortzet in 2008 en 2009.

Een ander aandachtspunt bij het gebruik van tijdreeksmodellen voor het kwantificeren van methodebreuken is dat in de komende jaren steeds meer waarnemingen onder het nieuwe design beschikbaar komen. Hierdoor is het mogelijk om steeds nauwkeurigere schattingen te maken voor de methodebreuken. Een consequentie is echter dat de schattingen voor de breuken en de eventuele gecorrigeerde tijdreeksen ieder jaar gereviseerd kunnen worden. In dit kader moet daarom worden nagedacht over een praktisch hanteerbare publicatiestrategie voor de methodebreuken en de eventuele gecorrigeerde reeksen. Dit punt geldt in extreme mate voor de eerste waarneming die in 2010 onder het nieuwe design beschikbaar komt. Voor dit eerste jaar is de schatting voor de methodebreuk gelijk aan het verschil tussen de voorspelling op basis van het tijdreeksmodel onder het oude design en de waarneming onder het OViN.

Voor wat betreft het te behalen detailniveau bij correctie kan er worden gezegd dat correctie op het niveau van een beperkt aantal kernvariabelen waarschijnlijk haalbaar is. Over lagere niveaus kunnen we nog geen uitspraken doen. De variatie op het totaal kan wellicht klein zijn zodat de hoofdvariabele goed te modelleren is, maar de data op het niveau van subvariabelen zullen veel grilliger zijn. Om een goede inschatting van het haalbare detailniveau te maken moeten er voldoende data beschikbaar zijn, in ieder geval van de jaren 2009 en 2010.

Referenties

- Bouhuijs, I. (2010). *Het nieuwe OViN – wijzigingen ten opzichte van OVG / MON*. CBS-notitie, Centraal Bureau voor de Statistiek, Heerlen.
- Van den Brakel, J.A., P. Smith en S. Compton (2008). *Quality procedures for survey transitions – experiments, time series and discontinuities*. Survey Research Methods, 2, 123-141.
- Van den Brakel, J.A. (2010). *Sampling and estimation techniques for the implementation of new classification systems*. Survey Research Methods, In Press.
- Van den Brakel, J.A., T. Kraan en J. Roels (2010). *Methodenreeks: Thema: Methodebreuken*. CBS Rapport, BPA-nr. DMH-2010-02-21-JBRL-TKRN-JRLS, Centraal Bureau voor de Statistiek, Heerlen.
- Van den Brakel, J.A. en J. Roels (2010). *Intervention analysis with state-space models to estimate discontinuities due to a survey redesign*. Annals of Applied Statistics, In Press.
- Durbin, J, en S.J. Koopman (2001). *Time series analysis by state space methods*. Oxford: Oxford University Press.
- Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.
- Harvey, A.C. en J. Durbin (1986). *The effects of seat belt legislation on British road casualties: a case study in structural time series modelling*. Journal of the Royal Statistical Society, Series A, 149, 187-227.

- Janssen, B. e.a. (2008). *Pilot Mobiliteitsonderzoek Nederland*. CBS Rapport, BPA-nr. SDV-2008-H096, Centraal Bureau voor de Statistiek, Heerlen.
- Konen, H.J. en H. Swinkels (2000). *Trendbreukanalyse Onderzoek Verplaatsingsgedrag vanaf 1985*. CBS nota, BPA-nr. H01190-02-SAH, Centraal Bureau voor de Statistiek, Sector Statistische Analyse personen, Heerlen.
- Molnár -in 't Veld, H. (2007). *Mobiliteitsonderzoek Nederland – Methodologische beschrijving*. Interne CBS nota, Centraal Bureau voor de Statistiek, Voorburg/Heerlen.
- Projectteam MON (2008). *Mobiliteitsonderzoek Nederland 2007*. Ministerie van Verkeer en Waterstaat, Rijkswaterstaat Dienst Verkeer en Scheepvaart.

Kennisgeving

De in dit rapport weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.