

LET OP

Deze SWOV-factsheet is gearchiveerd en wordt niet meer bijgewerkt.
Actuele SWOV-factsheets vindt u op swov.nl/factsheets.



SWOV-Factsheet

De analyse van tijdreeksen

Samenvatting

Om de ontwikkelingen in de verkeersveiligheid op kwantitatieve wijze te volgen, beschrijven, verklaren en te voorspellen, kunnen tijdreeksanalysemodellen worden gebruikt. Deze modellen bieden de mogelijkheid om ontwikkelingen in de verkeersveiligheid kwantitatief te modelleren op zodanige wijze dat rekening wordt gehouden met de afhankelijkheden in de observaties. Deze factsheet geeft uitleg over tijdreeksanalyse en gaat daarbij in op de functionaliteit van ARMA-, ARIMA-, DRAG-, en state-spacemodellen. Daarbij wordt gebruikgemaakt van relatief simpele beschrijvende analyses van tijdreeksen.

Achtergrond en inhoud

Een van de belangrijke taken van verkeersveiligheidsonderzoek is het volgen, beschrijven, verklaren en voorspellen van ontwikkelingen in de verkeersveiligheid. Als dit op een kwantitatieve manier gebeurt, dan zijn er analysetechnieken nodig om deze activiteiten op verantwoorde wijze te kunnen uitvoeren. Zoals de term 'ontwikkelingen' impliceert, betreft het hier de analyse van een heel speciaal type gegevens. Deze gegevens bestaan altijd uit herhaalde metingen in de tijd van een bepaald aspect van het verkeersproces. Een voorbeeld zijn de jaarlijkse aantallen verkeersdoden geobserveerd over een bepaalde periode. Zo'n variabele wordt een *tijdreeks* genoemd.

Een belangrijke eigenschap van de observaties in een tijdreeks is dat ze meestal *niet* onafhankelijk zijn van elkaar: het geobserveerd aantal verkeersdoden van vorig jaar is immers meestal een vrij goede indicator voor het aantal verkeersdoden van dit jaar. Aangezien standaardtechnieken uitgaan van onafhankelijke observaties, resulteert de analyse van tijdreeksen met standaardtechnieken heel vaak in residuen die eveneens onderling gecorreleerd zijn. Dit terwijl statistische toetsen en betrouwbaarheidslimieten gebaseerd zijn op de cruciale aanname dat de gevonden residuen random zijn, en dus onafhankelijk van elkaar.

Deze factsheet geeft een overzicht van de technieken voor het analyseren van tijdreeksen. Bovendien worden deze technieken met elkaar vergeleken op hun geschiktheid voor het modelleren van ontwikkelingen in de verkeersveiligheid. De factsheet is bedoeld voor onderzoekers die geïnteresseerd zijn in de analyse van tijdreeksen op diverse vakgebieden. Voor een goed begrip van de factsheet is elementaire kennis van statistiek en van met name klassieke lineaire regressie (zie bijvoorbeeld McCall, 1998) noodzakelijk.

Waarvoor zijn tijdreeksanalysemodellen een oplossing?

Tijdreeksanalysemodellen bieden de mogelijkheid om ontwikkelingen in de verkeersveiligheid kwantitatief te modelleren op zodanige wijze dat rekening wordt gehouden met de afhankelijkheden in de observaties.

Voor wie zijn tijdreeksanalysemodellen bedoeld?

Tijdreeksanalysemodellen zijn bedoeld voor onderzoekers in alle vakgebieden waar herhaalde metingen in de tijd worden verricht. Dat is dus zeker niet alleen de verkeersveiligheid, maar ook de economie, geschiedenis, de sociale wetenschappen, geneeskunde, biologie, en dergelijke.

Hoe werken tijdreeksanalysemodellen?

Van oudsher zijn er verschillende typen modellen in gebruik voor het analyseren van tijdreeksen. Deze modellen hebben met elkaar gemeen dat ze – in principe – niet alleen de ontwikkeling in een geobserveerde tijdreeks kunnen beschrijven, maar dat we ze ook kunnen gebruiken om verklaringen voor de ontwikkeling in een geobserveerde tijdreeks te vinden, en om de toekomstige waarden van de ontwikkeling in een geobserveerde tijdreeks te voorspellen. Hun geschiktheid voor het verwezenlijken van deze doelstellingen loopt echter sterk uiteen.

Deze factsheet geeft een overzicht van de verschillende mogelijkheden en gaat met name in op de voor- en nadelen van iedere aanpak. De verschillende modellen worden geïllustreerd aan de hand van hun resultaten bij de analyse van de volgende tijdreeks: het totaal aantal jaarlijkse verkeersdoden in Nederland in de periode 1950 t/m 2003, zoals onder andere weergegeven in *Afbeelding 1*.

Klassieke lineaire en niet-lineaire regressiemodellen

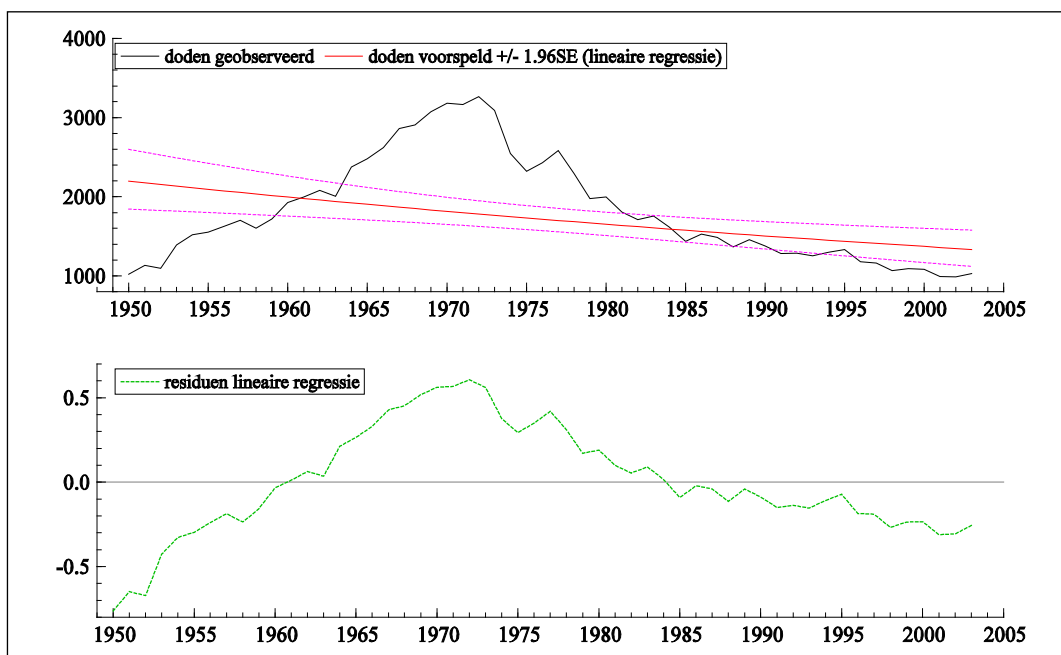
Van oudsher zijn klassieke lineaire en niet-lineaire regressiemodellen gebruikt om tijdreeksen te analyseren. Bij beschrijvende klassieke (niet-)lineaire regressieanalyses wordt de tijd als onafhankelijke variabele gebruikt voor het modelleren van de trend, en worden onafhankelijke dummyvariabelen gebruikt om eventuele seizoenseffecten (bij kwartaal- of maandgegevens bijvoorbeeld) in kaart te brengen. Zonder seizoenseffect is het beschrijvende klassieke lineaire regressiemodel simpelweg

$$y_t = a + bt + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (1)$$

voor tijdstippen $t = 1, \dots, n$. Hierbij is y_t de afhankelijke variabele (de geobserveerde tijdreeks), is t de onafhankelijke variabele (de tijd), zijn a en b de onbekende te schatten parameters, en is ε_t het residu. Parameter a is het intercept, en geeft het punt aan waar de regressielijn lijn de y -as snijdt; parameter b is het regressiegewicht en geeft de hellingshoek aan van de regressielijn met de x -as. Achter de regressievergelijking in (1) staan de *aannames* van het model: de residuen worden geacht normaal en onafhankelijk verdeeld (Normally and Independently Distributed) te zijn met gemiddelde nul, en

variantie σ_ε^2 . Daarnaast wordt aangenomen dat de residuen van het model homoscedastisch zijn, dat wil zeggen in de tijd een constante variantie hebben. Van deze drie aannames is hun volgorde van belangrijkheid voor correcte statistische conclusies als volgt: eerst onafhankelijkheid, dan homoscedasticiteit, en als laatste normaliteit.

Het resultaat van een lineaire regressie van de logaritme van het totaal aantal jaarlijkse verkeersdoden in Nederland op de variabele 'tijd' is weergegeven in *Afbeelding 1*. Uit deze figuur blijkt dat de regressielijn niet alleen grote afwijkingen van de observaties in de tijdreeks vertoont, maar ook dat de residuen (zoals weergegeven onderin *Afbeelding 1*) volstrekt niet onafhankelijk zijn: in de jaren 1950 t/m 1960 worden de observaties consequent overschat, in de jaren 1961 t/m 1984 worden ze consequent onderschat, om ten slotte in de jaren 1985 t/m 2003 weer voortdurend te worden overschat. Aangezien de berekening van het 95%-betrouwbaarheidsinterval rond de regressielijn wordt uitgevoerd onder de aanname van onafhankelijke residuen, betekent dit dat het betrouwbaarheidsinterval bovenin *Afbeelding 1* eveneens volstrekt inaccuraat is. Dit blijkt ook uit het feit dat 40 van de 54 observaties buiten het 95%-betrouwbaarheidsinterval liggen, terwijl dit er op grond van kans ongeveer 3 zouden moeten zijn.

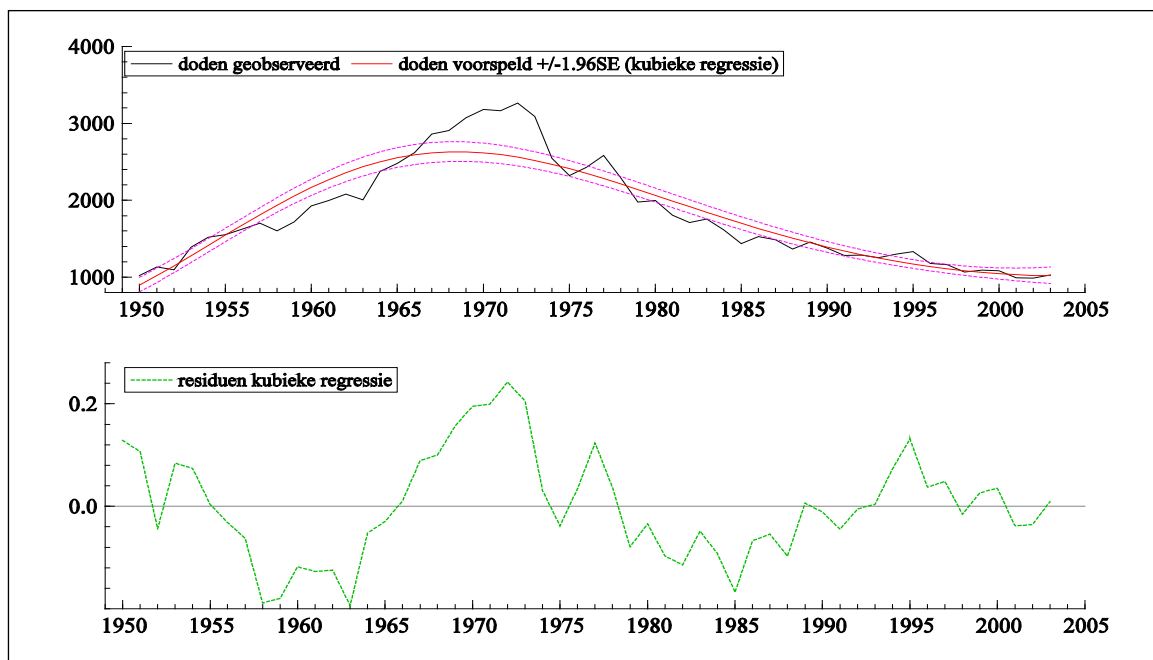


Afbeelding 1. Resultaten van lineaire regressieanalyse.

Ook de analyse van tijdreeksen met niet-lineaire regressiemodellen resulteert vaak in residuen die niet voldoen aan de aanname van onafhankelijkheid. In deze factsheet betekent de term 'niet-lineair model' zowel een model waarbij de relatie tussen de onafhankelijke en afhankelijke variabele niet-lineair is (maar het model zelf mogelijk nog wel 'lineair in de parameters' is), als een model dat niet-lineair in de parameters is. De analyse van de logaritme van de tijdreeks in *Afbeelding 1* met een zogenaamde kubieke trend

$$y_t = a + b_1t + b_2t^2 + b_3t^3 + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (2)$$

voor tijdstippen $t = 1, \dots, n$, waarbij a , b_1 , b_2 , en b_3 de onbekende en te schatten parameters zijn, levert bijvoorbeeld de resultaten getoond in *Afbeelding 2*. Hoewel de residuen van dit model kleiner zijn dan die in *Afbeelding 1* (zoals blijkt uit vergelijking van de schaal op de y-as van de plots met residuen in *Afbeeldingen 1* en *2*), en de modelpredicties dus beter passen bij de data, zijn de residuen nog steeds sterk serieel gecorreleerd, zoals te zien is in de plot onderin *Afbeelding 2*. Weer blijken de residuen in achtereenvolgende periodes consequent te worden over- dan wel onderschat. Dit betekent andermaal dat het 95%-betrouwbaarheidsinterval rond de kubieke trend bovenin *Afbeelding 2* inaccuraat is. Dit is ook te zien aan het feit dat nog steeds veel te veel observaties zich buiten dit interval bevinden.



Afbeelding 2. Resultaten van kubieke regressieanalyse.

Een ander voorbeeld van de problemen die optreden bij het toepassen van niet-lineaire regressiemodellen op tijdreeksen is te vinden in Commandeur & Koornstra (2001). Daar werd de ontwikkeling in de Nederlandse mobiliteit gemodelleerd met een S-vormige groeicurve (het zogeheten Gompertz-model), het fatale risico (gedefinieerd als het aantal verkeersdoden gedeeld door de mobiliteit) werd gemodelleerd met een exponentieel dalende curve, en de aantallen verkeersdoden werden voorspeld uit het product van de predicties van deze twee modellen. Bij deze analyses bleken de residuen eveneens sterk serieel gecorreleerd te zijn, en dus niet te voldoen aan de aanname van onafhankelijkheid. Ook Broughton et al. (2000) onderkennen het probleem van serieel gecorreleerde residuen bij de toepassing van lineaire regressiemodellen op de Britse ontwikkelingen in de verkeersveiligheid.

ARMA-, ARIMA- en DRAG-modellen

ARMA- en ARIMA-modellen (Box & Jenkins, 1976) en DRAG-modellen (Gaudry, 1984; Gaudry & Lassarre, 2000) zijn speciaal ontwikkeld voor de analyse van tijdreeksen, en houden over het algemeen dan ook veel beter rekening met de afhankelijkheden in de observaties dan klassieke (niet-) lineaire regressiemodellen.

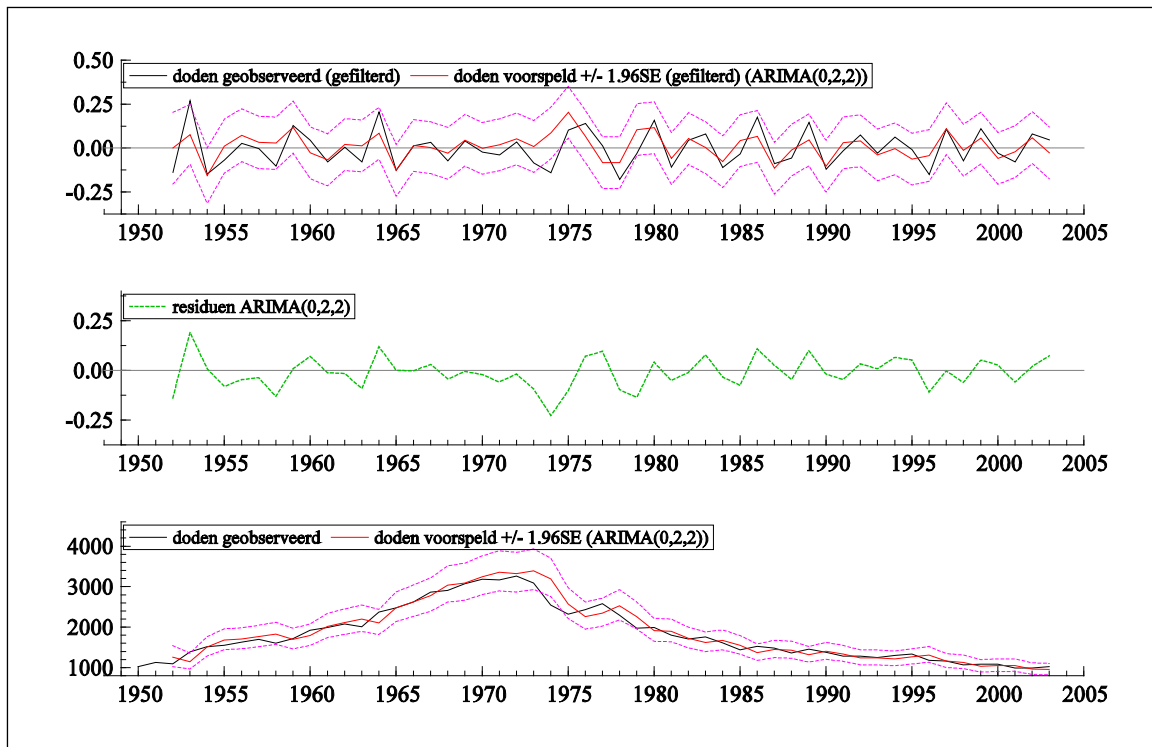
Algebraïsch kan het zogenaamde ARMA-model als volgt worden geschreven:

$$y_t = b_1 y_{t-1} + \dots + b_p y_{t-p} + \varepsilon_t + a_1 \varepsilon_{t-1} + \dots + a_q \varepsilon_{t-q}, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \quad (3)$$

voor tijdstippen $t = 1, \dots, n$, waarbij b_1, \dots, b_p en a_1, \dots, a_q de $p+q$ onbekende parameters zijn. In het eerste deel rechts van het is-gelijkteken van (3) wordt de afhankelijke variabele y_t voorspeld op grond van in de tijd verschoven waarden van zichzelf. Dit wordt dan ook het (regressie op zichzelf) *autoregressieve* (AR) deel van de vergelijking genoemd. In het tweede deel rechts van het is-gelijkteken wordt de afhankelijke variabele voorspeld op grond van in de tijd verschoven waarden van de residuen, en dit is het voortschrijdend gemiddelde (MA, *moving average*) deel van de vergelijking. Voor een gegeven tijdreeks is het de taak van de onderzoeker om een ARMA(p, q)-modelspecificatie te bepalen die een balans vormt tussen een zo goed mogelijk passend model en een spaarzaam gebruik van parameters, en die bovendien resulteert in residuen die aan de aannames van normaliteit en onafhankelijkheid voldoen.

Een belangrijke vereiste van ARMA-modellen is echter, en dat is ook hun zwakke punt, dat de geanalyseerde tijdreeks *stationair* moet zijn (dat wil zeggen een over de tijd constant gemiddelde en constante variantie moet hebben) vóór de eigenlijke analyse kan worden uitgevoerd. Aangezien tijdreeksen in de praktijk vaak bestaan uit niet-stationaire componenten als een trend of een seizoenseffect (zoals bijvoorbeeld bij kwartaal- of maandgegevens) moeten de observaties bij ARMA-modellen meestal eerst worden gefilterd. Dit betekent dat de trend en het seizoenseffect eerst uit de reeks moeten worden verwijderd, door het nemen van verschillen tussen observaties, om stationariteit te verkrijgen. Dit resulteert dan in zogenaamde ARIMA(p, d, q)-modellen waarbij de waarde van d aangeeft hoe vaak verschillen tussen opeenvolgende observaties moeten worden genomen voordat stationariteit van de tijdreeks is verkregen. De eigenlijke analyse bestaat vervolgens uit het bepalen welke ARMA(p, q) het beste bij de gefilterde (en hopelijk stationaire) tijdreeks past.

Uit analyse van de logaritme van het totaal aantal jaarlijkse verkeersdoden in Nederland in de periode 1950 t/m 2003 met ARIMA-modellen blijkt dat een ARIMA(0,2,2)-specificatie een goede beschrijving geeft van deze tijdreeks. In *Afbeelding 1* is duidelijk te zien dat de tijdreeks niet stationair is: eerst stijgt het jaarlijks aantal verkeersdoden vrijwel voortdurend, om daarna vrijwel voortdurend te dalen. Deze tijdreeks moet dus eerst stationair worden gemaakt. In dit geval lijkt voldoende stationariteit te worden verkregen door de eigenlijke analyse uit te voeren op de tweedeordeverschillen tussen de opeenvolgende observaties van de oorspronkelijke serie, zodat d in dit geval gelijk is aan 2. Deze gefilterde versie van de tijdreeks is bovenin *Afbeelding 3* weergegeven, samen met de voorspellingen verkregen met het ARIMA(0,2,2)-model. Het resultaat van de ARIMA(0,2,2)-modelspecificatie, terugvertaald naar de oorspronkelijke tijdreeks, wordt onderin *Afbeelding 3* getoond, en de residuen staan in het midden van de figuur. Deze residuen voldoen voor het eerst wel aan de belangrijke aanname van onafhankelijkheid. Dit blijkt eveneens uit het feit dat de meeste observaties nu binnen de 95%-betrouwbaarheidslimieten vallen.



Afbeelding 3. Resultaten van analyse met het ARIMA(0,2,2)-model.

DRAG-modellen (Gaudry, 1984; Gaudry & Lassarre, 2000) zijn uitbreidingen van ARIMA-modellen waarbij:

- expositie, ongevallen en ongevalsernst in drie aparte analysestappen worden gemodelleerd;
- veel verklarende variabelen aan ieder van deze drie stappen worden toegevoegd;
- zowel de afhankelijke variabele als de onafhankelijke variabelen mogen worden getransformeerd door zogenaamde Box-Cox-transformaties.

Bovenop de reeds gemelde stationariteitseisen voor ARIMA-modellen zijn de nadelen van DRAG-modellen daarom dat:

- expositie, ongevallen en ongevalsernst niet simultaan worden gemodelleerd;
- het grote aantal verklarende variabelen strijdig lijkt met het principe van parsimonie;
- de Box-Cox-transformaties van de variabelen (die worden gepresenteerd als een voordeel) kunnen leiden tot identificatie- en interpretatieproblemen.

State-spacemodellen

Een meer recente ontwikkeling op het gebied van de analyse van tijdreeksen vormen de zogenaamde structurele tijdreeksmodellen (Harvey, 1989; Durbin & Koopman, 2012; Commandeur & Koopman, 2007), ook wel ongeobserveerde componentenmodellen of state-spacemodellen genoemd. In tegenstelling tot de hierboven besproken ARIMA- en DRAG-modellen veronderstellen structurele tijdreeksmodellen geen stationariteit van tijdreeksen, en bieden ze de mogelijkheid om tijdreeksen expliciet te ontbinden in componenten als een trend en een seizoenseffect.

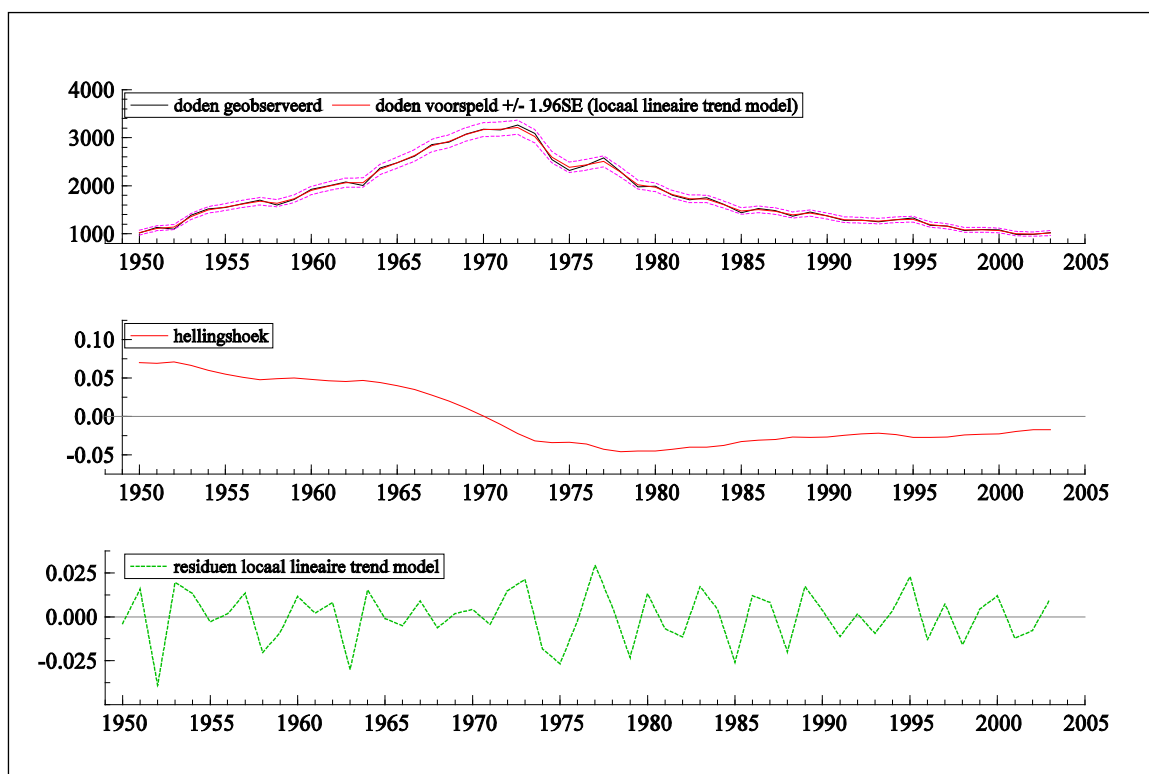
Alle state-spacemodellen kunnen in een algemene notatie worden weergegeven. Aangezien deze notatie echter kennis van matrixalgebra vereist, zullen we ons hier beperken tot het bespreken van één state-spacemodel, en wel dát model dat een goede beschrijving geeft van de ontwikkeling in de tijdreeks van Nederlandse verkeersdoden: het lokaal lineaire trendmodel.

Het lokaal lineaire trendmodel wordt verkregen door het intercept a en het regressiegewicht b in het klassieke lineaire regressiemodel (1) te laten variëren in de tijd, en wel op de volgende wijze:

$$\begin{aligned}
 y_t &= a_t + \varepsilon_t, & \varepsilon_t &\sim NID(0, \sigma_\varepsilon^2) \\
 a_{t+1} &= a_t + b_t + \xi_t, & \xi_t &\sim NID(0, \sigma_\xi^2) \\
 b_{t+1} &= b_t + \zeta_t, & \zeta_t &\sim NID(0, \sigma_\zeta^2)
 \end{aligned}
 \tag{4}$$

voor tijdstippen $t = 1, \dots, n$. Hierin is a_t de trend, die bestaat uit de som van een tijdvariërend intercept en van een tijdvariërende hellingshoek b_t . Het lokaal lineaire trendmodel is dus te beschouwen als een generalisatie van het klassieke lineaire regressiemodel, met dit belangrijke verschil dat de parameters die in model (1) globaal zijn (dat wil zeggen niet mogen veranderen in de tijd), in model (4) lokaal worden gemaakt (en dus per tijdstip mogen variëren). De onbekende en te schatten parameters in het

lokaal lineaire trendmodel zijn de errorvarianties σ_ε^2 , σ_ξ^2 en σ_ζ^2 , en de initiële waarden a_1 en b_1 op tijdstip $t = 1$ van het intercept en van de hellingshoek.



Afbeelding 4. Resultaten van analyse met het lokaal lineaire trendmodel.

De resultaten van de analyse van de logaritme van de jaarreeks met Nederlandse verkeersdoden met model (4) zijn weergegeven in *Afbeelding 4*. Bovenin de figuur staan de observaties en de tijdvariërende trend, samen met het 95%-betrouwbaarheidsinterval, in het midden van de figuur wordt het verloop van de tijdvariërende hellingshoek getoond, en onderin zijn de residuen van het model weergegeven. Dat de hellingshoek van 1950 tot en met 1970 positief is, toont dat het aantal verkeersdoden in die periode voortdurend aan het stijgen was, terwijl deze trend voortdurend dalend was in de periode na 1970 omdat de hellingshoek daarna steeds negatief is. Het feit dat de hellingshoek aan het einde van de serie langzaam aan weer richting nul kruipt, betekent dat de dalende ontwikkeling daar steeds meer afvlakt. De residuen onderin *Afbeelding 4* voldoen aan alle modelaannames van onafhankelijkheid, homoscedasticiteit, en normaliteit, zodat de 95%-betrouwbaarheidslimieten rond de modelvoorspellingen bovenin de figuur ook correct zijn.

State space-modellen zijn heel flexibel: zij kunnen makkelijk ontbrekende observaties aan, en zijn eenvoudig uit te breiden naar multivariate tijdreeksanalysemodellen. Een ander voordeel van structurele tijdreeksmodellen is dat ze, zoals reeds eerder opgemerkt, klassieke regressie als speciaal

geval hebben. Door de errorvarianties σ_ε^2 en σ_ζ^2 in (4) op nul te fixeren, ontstaat namelijk automatisch het klassieke lineaire model (1). Dit maakt het mogelijk om de resultaten van analyses met deze twee typen modellen direct met elkaar te vergelijken. Ook op ARIMA-modellen kunnen state-spacemethoden worden toegepast. Alle in deze factsheet gepresenteerde analyses zijn dan ook met state-spacemethoden uitgevoerd.

Het is om deze redenen dat structurele tijdreeksmodellen te verkiezen zijn boven andere modellen als het gaat om het beschrijven, verklaren en voorspellen van ontwikkelingen in de verkeersveiligheid. Toepassingen van deze modellen op het gebied van de verkeersveiligheid zijn onder meer te vinden in Harvey & Durbin (1986), waar het effect van de invoering van de wet op gordelgebruik wordt geëvalueerd in Groot-Brittannië, en in Ernst & Brüning (1990) waar hetzelfde wordt gedaan voor Duitsland. Op basis van de analyse van maandgegevens concluderen Harvey & Durbin onder andere dat de invoering van de wet op gordelgebruik op 31 januari 1984 in Groot-Brittannië heeft geleid tot een daling van 23% in het aantal doden en zwaargewonden bij autobestuurders en van 18% in het aantal doden bij autobestuurders. Ernst & Brüning vinden voor West-Duitsland onder meer een daling van 28% in het aantal doden bij auto-inzittenden, en een daling van 21% in het aantal zwaargewonden onder auto-inzittenden na invoering van de wet op gordelgebruik op 1 augustus 1984. Lassarre (2001) heeft structurele tijdreeksmodellen gebruikt voor het onderling vergelijken van de ontwikkelingen in de verkeersveiligheid van tien Europese landen. Op grond van de analyse van jaarcijfers over verkeersdoden en motorvoertuigkilometers, vindt Lassarre voor de tien Europese landen in 1994 een gemiddelde daling van 6% in fataal risico per jaar, met als kleinste percentuele daling die van 4,7% in Frankrijk en Finland, en als grootste percentuele daling die van 13,4% in Spanje.

Ord & Young (2003) hebben structurele tijdreeksmodellen toegepast bij de analyse van de effecten van de aanslagen in de Verenigde Staten op 11 september 2001 op ontwikkelingen in de luchtvaart en het treinverkeer. Zij laten zien dat een combinatie van drie effecten (een kortetermijn-, een tijdelijke en een permanente verschuiving) het snelst in staat is om op maandniveau de veranderingen na de aanslagen in het lucht- en treinverkeer van de Verenigde Staten op te pikken. Bij de SWOV zijn structurele tijdreeksmodellen onder andere toegepast bij het voorspellen van het aantal verkeersdoden in de EU-lidstaten in 2010 (Ecorys Transport & SWOV, 2005) en bij het voorspellen van het aantal verkeersdoden in Nederland in 2020 (Wesemann, Van Norden & Stipdonk, 2010). Hierbij is gebruikgemaakt van het multivariate structurele tijdreeksmodel gepubliceerd in Bijleveld et al. (2008), en besproken in Bijleveld (2008). In het kader van het Europese project DaCoTA zijn deze modellen eveneens toegepast om tot prognoses te komen van het aantal verkeersdoden voor alle lidstaten van de Europese Unie, zie Martensen & Dupont (2010) en Lassarre, Dupont & Antoniou (2012).

Wat levert de analyse van tijdreeksen op?

Om didactische redenen hebben we ons in deze factsheet beperkt tot relatief simpele beschrijvende analyses van tijdreeksen. Daarnaast kunnen tijdreeksanalyses echter ook gebruikt worden om:

1. de effecten van maatregelen en van andere variabelen op de ontwikkeling in de verkeersveiligheid te evalueren;
2. te onderzoeken of nieuw gepubliceerde cijfers al dan niet afwijken van de verwachting op grond van het verleden;
3. prognoses te doen over toekomstige ontwikkelingen in de verkeersveiligheid.

Conclusie

Vergelijken we klassieke (niet-)lineaire regressiemodellen, ARIMA- en DRAG-modellen, en state-spacemodellen op hun merites bij het analyseren van tijdreeksen, dan verdienen state-spacemodellen de voorkeur. Niet alleen kunnen ze aangewend worden om een tijdreeks expliciet te ontbinden in voor onderzoek interessante componenten zoals een trend of een seizoenseffect, maar door hun flexibiliteit zijn ze doorgaans ook goed in staat om rekening te houden met de afhankelijkheden in de observaties van een tijdreeks. Bovendien gaan ze transparant om met ontbrekende gegevens, en zijn ze gemakkelijk uit te breiden naar de multivariate analyse van tijdreeksen.

Publicaties en bronnen

Bijleveld, F. (2008). [*Time series analysis in road safety research using state space methods*](#). Proefschrift Vrije Universiteit VU, Amsterdam. SWOV-Dissertatiereeks, Leidschendam.

Bijleveld, F.D., Commandeur, J.J.F., Gould, P. & Koopman, S.J. (2008). [*Model-based measurement of latent risk in time series with applications*](#). In: Journal of the Royal Statistical Society Series A, vol. 171, nr. 1, p. 265-277.

Box, G.E.P. & Jenkins, G.M. (1976). [*Time series analysis, forecasting and control*](#). Holden-Day, San Francisco.

- Broughton, J., Allsop, R.E., Lynam, D.A. & McMahon, C.M. (2000). [*The numerical context for setting national casualty reduction targets*](#). TRL Report 382, Transport Research Laboratory, Crowthorne.
- Commandeur, J.J.F. & Koopman, S.J. (2007). [*An introduction to state space time series analysis*](#). Oxford Practical Econometrics Series, nr. 1. Oxford University Press, Oxford.
- Commandeur, J.J.F. & Koornstra, M.J. (2001). [*Prognoses voor de verkeersveiligheid in 2010: mobiliteit en slachtoffers op grond van de ontwikkelingen in de jaren 1948-1998*](#). R-2001-9. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV, Leidschendam.
- Durbin, J. & Koopman, S.J. (2012). [*Time series analysis by state space methods*](#). Second Edition. Oxford University Press, Oxford.
- Ecorys Transport & SWOV (2005). [*Impact Assessment Road Safety Action Programme; Assessment for mid term review*](#). Ecorys Transport/SWOV, Rotterdam/Leidschendam.
- Ernst, G. & Brüning, E. (1990). [*Fünf Jahre danach: Wirksamkeit der Gurtanlegepflicht für Pkw Insassen ab 1.8.1984*](#). In: Zeitschrift für Verkehrssicherheit, vol. 36, nr. 1, p. 2-13.
- Gaudry, M. (1984). [*DRAG, un modèle de la Demande Routière, des Accidents et de leur Gravit , appliqu  au Qu bec de 1956   1982*](#). Publication CRT-359. Centre de Recherche sur les Transports CRT, Universit  de Montr al, Montr al.
- Gaudry, M. & Lassarre, S. (eds.) (2000). [*Structural road accident models; The international DRAG family*](#). Pergamon, Amsterdam.
- Harvey, A.C. (1989). [*Forecasting, structural time series models and the Kalman filter*](#). Cambridge University Press, Cambridge.
- Harvey, A.C. & Durbin, J. (1986). [*The effects of seat belt legislation on British road casualties: A case study in structural time series modelling*](#). In: Journal of the Royal Statistical Society A, vol. 149, nr. 3, p. 187-227.
- Lassarre, S. (2001). [*Analysis of progress in road safety in ten European countries*](#). In: Accident Analysis and Prevention, vol. 33, p. 743-751.
- Lassarre, S., Dupont, E. & Antoniou, C. (eds.) (2012). [*Forecasting road traffic fatalities in European countries: Towards an integrated European model*](#). Deliverable 4.7 of the EC FP7 project DaCoTA. European Commission, Brussels.
- Martensen, H. & Dupont, E. (eds.) (2010). [*Forecasting road traffic fatalities in European countries: Model and first results*](#). Deliverable 4.2 of the EC FP7 project DaCoTA. European Commission, Brussels.
- McCall, R.B. (1998). [*Fundamental statistics for behavioral sciences*](#). Seventh Edition. Brooks/Cole Publishing Company, Pacific Grove.
- Ord, K. & Young, P. (2004). [*Estimating the impact of recent interventions on transportation indicators*](#). In: Journal of Transportation and Statistics. vol. 7, nr. 1.
- Wesemann, P., Norden, Y. van & Stipdonk, H. (2010). [*An outlook on Dutch road safety in 2020; future developments of exposure, crashes and policy*](#). In: Safety Science, vol. 48, nr. 9, p. 1098-1105.