

Monitoring van verkeersveiligheid

Drs. F.D. Bijleveld

Monitoring van verkeersveiligheid

Beschrijving van een rekeninstrument voor het volgen van ontwikkelingen in de verkeersveiligheid

R-99-20

Drs. F.D. Bijleveld

Leidschendam, 1999

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV

Documentbeschrijving

Rapportnummer:	R-99-20
Titel:	Monitoring van verkeersveiligheid
Ondertitel:	Beschrijving van een rekeninstrument voor het volgen van ontwikkelingen in de verkeersveiligheid
Auteur(s):	Drs. F.D. Bijleveld
Onderzoeksmanager:	Ing. J.A.G. Mulder
Projectnummer SWOV:	51.051
Projectcode opdrachtgever:	PRDVL 98.013
Opdrachtgever:	Ministerie van Verkeer en Waterstaat, Directoraat-Generaal Rijkswaterstaat, Adviesdienst Verkeer en Vervoer
Trefwoord(en):	Mathematical model, prototype, program (computer), safety, trend (stat), statistics, age, transport mode, mobility (pers), injury, forecast, Netherlands.
Projectinhoud:	Voor de monitoring van de verkeersveiligheid is een rekenmodel ontwikkeld, op basis van analysetechnieken die reeds bij de SWOV aanwezig waren. Het model is ontwikkeld voor de analyse van verschillende soorten gegevens die aan verkeersveiligheid zijn gerelateerd, en van de samenhang tussen deze gegevens. Op grond van deze analyse kunnen er voorspellingen gedaan worden over ontwikkelingen in de verkeersveiligheid. Dit rapport beschrijft een prototype van dit model, zijn werking en mogelijkheden, alsmede enkele voorbeelden van toepassingen.
Aantal pagina's:	48 + 45 blz.
Prijs:	f 30,-
Uitgave:	SWOV, Leidschendam, 1999

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
Postbus 1090
2260 BB Leidschendam
Telefoon 070-3209323
Telefax 070-3201261

Samenvatting

In opdracht van de Adviesdienst Verkeer en Vervoer (AVV) is door de SWOV een rekenmodel ontwikkeld voor de monitoring van de verkeersveiligheid. De basis hiervan vormen analysetechnieken die reeds bij de SWOV aanwezig waren. Het model is ontwikkeld voor het analyseren van verschillende soorten gegevens die aan verkeersveiligheid zijn gerelateerd, en van de samenhang tussen deze gegevens. Dit rapport beschrijft een prototype van dit model, zijn werking en mogelijkheden, alsmede enkele voorbeelden van toepassingen.

Globale doelstelling

Het rekenmodel is bedoeld als instrument dat behulpzaam kan zijn bij de beantwoording van beleidsvragen. Dit instrument is vooralsnog een (werkend) prototype, waarvan de definitieve vorm nog moet worden vastgesteld op basis van de specifieke beleidsdoeleinden waarvoor het zal worden gebruikt.

Het gaat om de volgende typen beleidsvragen:

- Hoe heeft de verkeersveiligheid zich in het verleden ontwikkeld?
- Welke trends zijn te ontdekken in de ontwikkeling van de verkeersveiligheid?
- Zijn de trends te extrapoleren naar de toekomst en met welke mate van onzekerheid gaan deze predicties gepaard?
- Hoe verhouden deze ontwikkelingen en predicties zich ten opzichte van 'de taakstelling'?
- Welke deelaspecten in deze ontwikkelingen vereisen aanvullende verkeersveiligheidsmaatregelen?
- Hoe groot is het effect van reeds genomen maatregelen?

Globale werking

De werking van het instrument is primair gericht op de analyse van verschillende reeksen in de tijd geordende gegevens, en van de samenhang tussen deze gegevens en relevante indicatoren van verkeersveiligheid. Op grond van de ontwikkeling in deze tijdreeksen analyses kunnen er voorspellingen gedaan worden over ontwikkelingen in de verkeersveiligheid. Dit is van belang voor de beantwoording van bovenstaande beleidsvragen. Ook is het noodzakelijk dat het instrument verschillende ontwikkelingen gelijktijdig kan analyseren.

Het 'front end' van het instrument, de wijze waarop de gebruiker ermee wordt geconfronteerd, bestaat uit een invoermodule, een (beperkte) modelspecificatie-module en een uitvoermodule. In de invoermodule moet de invoer worden gespecificeerd en een invoerbestand worden gekozen; in de uitvoermodule wordt de uitvoer opgeslagen. De uitvoer bestaat uit cijfermatige uitkomsten (modelparameters, modelschattingen, prognosecijfers, enzovoort) en grafieken, waarin diverse aspecten van de ontwikkelingen en prognoses kunnen worden vastgelegd (de algemene risicotrend, de seizoentrend enzovoort). Dit 'front end' is ontworpen voor de meest voorkomende (soorten) analyses. Voor meer specifieke analyses moet een daarop afgesteld programma worden uitgevoerd. Het prototype is geïmplementeerd in *Mathematica*TM

versie 3.0. en kan onder verschillende computerplatforms worden gebruikt, waaronder *Microsoft Windows '95*TM.

Mogelijkheden en vereisten invoergegevens

Het prototype heeft op dit moment de volgende mogelijkheden en vereisten:

Het gebruik van *seizoenseffecten* is mogelijk. Hierdoor zijn de praktische basiseenheden het jaar, het kwartaal of de maand. In feite wordt een hoofd-tijdseenheid gekozen, meestal het jaar. Deze eenheid kan vervolgens in een geheel aantal delen worden opgedeeld, zodat de basiseenheid jaar (één deel), kwartaal (vier delen), maand (twaalf delen) of zelfs week (52 delen) zou kunnen zijn.

De lengte van de *prognoseperiode* kan opgegeven worden. Dit gebeurt in aantallen tijdseenheden van de observaties. Dus bij een seizoenseffect van kwartalen wordt de prognoselengte in kwartalen uitgedrukt.

De analyse betreft *verkeersgegevens en ongevalsgegevens* gelijktijdig. Hierbij wordt gedacht aan expositiecijfers (bijvoorbeeld reizigerskilometers) en twee typen ongevalsgegevens (bijvoorbeeld aantallen doden en gewonden, of aantallen ongevallen met dodelijke afloop en ongevallen met letsel). De algemene implementatie kan iedere combinatie aan.

Van alle gebruikte cijfers zijn (relatieve-) *betrouwbaarheidsgegevens* nodig. Meestal moeten deze bij de verkeersgegevens apart worden ingevoerd, terwijl deze bij de slachtoffergegevens via de Poisson-assumptie geschat zouden kunnen worden.

De invoergegevens die het instrument gebruikt moeten telkens betrekking hebben op *gelijkwaardige, opeenvolgende perioden*. De opeenvolgende periodes hoeven niet noodzakelijkerwijs aaneensluitend te zijn. Wel moeten de omstandigheden waarop de ene observatie betrekking heeft een relatie vertonen met de omstandigheden waarop een volgende observatie betrekking heeft, zodat kennis over een bepaalde waarneming gebruikt kan worden voor het maken van een voorspelling voor een volgende waarneming. Technisch gezien werkt het instrument zo, dat het de voorspellingen uit het verleden vergelijkt met bijbehorende geobserveerde gegevens. Op basis van deze vergelijking(en) worden de voorspellingen gedaan en wordt de kwaliteit ervan vastgesteld.

De invoergegevens kunnen *verschillende bronnen* hebben en hoeven ook niet steeds voor alle waarnemingen beschikbaar te zijn. Zo kunnen bijvoorbeeld gordeldraagpercentages uit verschillende onderzoeken afkomstig zijn. Het is wel noodzakelijk dat de gegevens steeds voor meerdere tijdstippen beschikbaar zijn, zodat enig gebruik gemaakt kan worden van de ontwikkeling van de gegevens in het verleden. Het instrument bezit een grote tolerantie voor ontbrekende observaties, hoewel de voorspelling daar natuurlijk wel gevoelig voor is: hoe meer ontbrekende gegevens, hoe meer onzekerheid.

Mogelijke typen uitvoergegevens

Het is nuttig gebleken de ontwikkeling van de verkeersonveiligheid te kunnen ontbinden in een *trend*-datgene waarin men vaak het meest geïnteresseerd is en een *seizoenseffect*. Ook kunnen bijvoorbeeld invloeden van exogene grootheden als mobiliteit op het risico onderzocht worden.

De werkelijke data van een waargenomen ontwikkeling over een bepaalde (meestal recente) periode kan vergeleken worden met een prognose voor die ontwikkeling over diezelfde periode, waarbij die data niet bij de berekeningen is gebruikt. Dit kan voor alle data in een bepaalde periode gebeuren, maar ook voor een deel daarvan. Op deze wijze kan men de recente ontwikkeling van de verkeersveiligheid vergelijken met wat er 'uit het verleden te verwachten was'.

Een andere mogelijkheid is dat men bepaalde data, bijvoorbeeld de mobiliteitscijfers, wel in de berekeningen meeneemt. Op deze manier kan men de prognose voor het ongevalsrisico corrigeren voor de effecten van een eventueel afwijkende ontwikkeling in deze cijfers, in dit geval de mobiliteit.

Een derde mogelijkheid is uiteraard een *zuivere prognose*. Deze kan worden gegenereerd over een toekomstige periode, waarover nog geen gegevens bekend zijn.

Andere soorten uitkomsten van het model, waaronder toetsen voor de relevantie van verschillende in het model toegepaste componenten zijn op dit moment nog niet beschikbaar.

Voorbeelden van toepassing

Naast een theoretisch voorbeeld, aan de hand waarvan de specifieke eigenschappen van het model worden uitgelegd, is in het rapport een aantal voorbeelden uitgewerkt met invoergegevens, die in principe gedisaggregeerd zijn (leeftijd en wijze van vervoer). Hierbij zijn voor een aantal groepen verkeersdeelnemers de ontwikkelingen in het aantal verkeersslachtoffers (doden en ziekenhuisgewonden apart) gerelateerd aan de mobiliteit in termen van reizigerskilometers.

Bij de voorbeelden zijn de aantallen reizigerskilometers en aantallen doden en ziekenhuisgewonden verzameld uit bestanden van respectievelijk het onderzoek verplaatsingsgedrag (OVG), de Afdeling Basisgegevens van de AVV (AVV/BG) en de Landelijke Medische Registratie (LMR). Deze aantallen zijn voor elk kwartaal van de periode 1985 tot en met 1997 verzameld, evenals schattingen van hun varianties.

De gegevens zijn telkens onderscheiden naar vervoerwijze en leeftijdsklasse. De categorieën slachtoffers per vervoerwijze zijn personenauto-inzittenden (passagier en bestuurder samengevoegd), motorfietsers, bromfietsers, fietsers, voetgangers, overige vervoerwijzen en alle vervoerwijzen samen. De categorieën slachtoffers per leeftijdsklasse zijn 0- tot 14-jarigen, 15- tot 24-jarigen, 25- tot 49-jarigen, 50- tot 64-jarigen, 65 jaar en ouder, en alle leeftijdsklassen samen.

Verzameling van de gegevens leverde voor iedere combinatie van wijze van vervoer en leeftijdsklasse plus marginalen (alle vervoerwijzen en leeftijden)

een tabel op met geobserveerde aantallen slachtoffers en reizigerskilometers per kwartaal, en een gelijkvormige tabel met bijbehorende betrouwbaarheidsmaten. Deze tabellen zijn als invoer voor het instrument gebruikt.

Voor iedere combinatie is geprobeerd de laatste acht observaties (de jaren 1996 en 1997) te voorspellen uit de eerdere observaties (tot en met 1995). Daarna is ongeveer een zelfde analyse uitgevoerd gebruikmakend van gegevens van de in werkelijkheid geobserveerde reizigerskilometers voor de jaren 1996 en 1997. Tenslotte is een prognose voor de komende jaren (1998, 1999) uitgevoerd, met behulp van de gegevens tot en met 1997.

De volgende combinaties zijn in het rapport uitgewerkt:

1. alle leeftijden en alle wijzen van vervoer;
2. 0- tot 14-jarigen in de personenauto;
3. 25- tot 49-jarige personenauto-inzittenden;
4. 15- tot 24-jarige bromfietzers.

Daarnaast is de registratie van AVV/BG-ziekenhuisgewonden ten opzichte van LMR-ziekenhuisgewonden aan de hand van een aantal combinaties vergeleken. In een laatste voorbeeld is een (beperkte) poging gedaan de invloed van de mobiliteit van het snelverkeer op aantallen slachtoffers onder het langzaam verkeer te onderzoeken.

Resultaten gedisaggregeerde modellen

Geval 1: alle leeftijden en alle wijzen van vervoer

Bij de analyses van de combinatie ‘alle leeftijden’ en ‘alle wijzen van vervoer’ is het volgende opgevallen:

- Er doet zich een geleidelijke daling voor van het risico om gedood te worden in het verkeer; hierin doen zich geen markante variaties voor.
- Het risico voor de (LMR-)ziekenhuisgewonden ontwikkelt zich minder gelijkmatig dan het risico voor de verkeersdoden. Dit verschijnsel doet zich ook voor als AVV/BG-ziekenhuisgewonden worden geanalyseerd in plaats van LMR-ziekenhuisgewonden.
- Het seizoenseffect van de slachtoffers varieert sterker dan het seizoenseffect van de reizigerskilometers. Dit duidt op een seizoensinvloed op het risico, die wellicht meer met het weer en de invloed ervan op de ‘modal split’ te maken heeft dan met seizoenswijzigingen in het totale mobiliteitsvolume.
- In deze analyse voor alle leeftijden samen is een correctie niet nodig gebleken voor de categorie kinderen van 0 tot 11 jaar, die tot 1994 in het OVG ontbraken.

Geval 2: 0- tot 14-jarigen in de personenauto

De belangrijkste conclusies uit de analyse van de combinatie ‘0- tot 14-jarigen’ en ‘personenauto-inzittenden’ zijn:

- Er moet gecorrigeerd worden voor de verandering in het OVG in 1994, toen de bijdrage van kinderen van 0 tot 11 jaar is gaan meetellen.

- Deze correctie komt redelijk overeen met de ontwikkeling in de verhouding van aantallen jongeren van 0 tot 11 jaar en 12 tot 14 jaar.

Geval 3: 25- tot 49-jarige personenauto-inzittenden

De analyse van de combinatie '25- tot 49-jarigen' en 'personenauto-inzittenden' had de volgende resultaten:

- Opvallend is de onzekerheid in de prognoses voor het aantal ziekenhuisgewonden.
- Er is een lichte overschatting van het aantal reizigerskilometers in 1996 en 1997.
- De amplitude van het seizoenseffect blijkt nauwelijks kleiner te worden in de geobserveerde periode. Dit is in tegenstelling tot vorige voorbeelden.

Geval 4: 15- tot 24-jarige bromfietzers

De analyse van de combinatie '15- tot 24-jarigen' en 'bromfietzers' had als resultaten:

- Er is geen sprake van een daling van de slachtofferrisico's.
- Er is wel een dalende ontwikkeling in het aantal reizigerskilometers.
- De jaren 1985 tot 1988 hebben een ander seizoenseffect dan de rest van de jaren. Dit zou een gevolg kunnen zijn van het snorfietsgebruik.
- Het model heeft 'geconstateerd' dat afwijkingen in de reizigerskilometers niet in de slachtofferaantallen terugkomen, en het model heeft blijkbaar deze cijfers min of meer gecorrigeerd.

Registratie AVV/BG- versus LMR-ziekenhuisgewonden

Voor verschillende vervoerwijze zijn zowel geregistreerde AVV/BG- als LMR-ziekenhuisgewonden meegenomen. Gekeken is naar de verhouding tussen de verschillende registraties, en de ontwikkeling daarin. Een aantal conclusies, met soms gewaagde interpretaties, zijn:

- Het blijkt dat er sprake kan zijn van een seizoensgebonden verhouding tussen LMR- en AVV/BG-gewonden. Het lijkt hier dat de seizoenseffecten van beide registraties in wezen dezelfde vorm hebben, maar dat het seizoenseffect bij de LMR-gewonden sterker is (een grotere amplitude). Dit zou misschien voor een deel veroorzaakt kunnen worden door een overbelastingsprobleem bij de registratie van de ziekenhuisgewonden, wanneer het om hogere pieken gaat.
- Er bestaat een plotselinge (sterkere) daling in de verhouding tussen AVV/BG-gewonden en LMR-gewonden aan het einde van 1989 of het begin van 1990. Deze daling lijkt te stagneren vanaf eind 1992, en de verhouding neemt weer plotseling toe in 1996. Deze toename is waarschijnlijk niet het gevolg van een plotselinge afname in LMR-cijfers. Mogelijk is er een relatie tussen deze afname in de AVV/BG-registratie (op basis van politiegegevens) ten opzichte van de LMR-registratie, en de reorganisatie van de (verkeers)politie in ongeveer dezelfde periode.
- Het valt op dat het patroon van de ontwikkeling van de relatieve registratiegraad voor bromfietsen ongeveer hetzelfde is als voor alle vervoerwijzen samen.

Invloed mobiliteit snelverkeer op aantallen slachtoffers langzaam verkeer

Het aantal reizigerskilometers in personenauto's (als maat voor de mobiliteit van snelverkeer) is vergeleken met de slachtoffers in het langzaam verkeer. Bij deze analyses moet worden bedacht dat het om alle slachtoffers onder het langzaam verkeer gaat, en niet alleen om slachtoffers van ongevallen tussen langzaam verkeer en motorvoertuigen of personenauto's. Voor deze laatste groep zijn waarschijnlijk grotere effecten te verwachten. Een andere kanttekening is dat de mobiliteit van alle motorvoertuigen, in plaats van de reizigerskilometers van personenauto's zou kunnen worden gebruikt. Idealiter zou dan ook de mobiliteit op hogere-ordewegen (waar weinig contact met langzaam verkeer mogelijk is) hiervan afgetrokken moeten worden.

Resultaten van de analyse zijn:

- Door de toevoeging van de reizigerskilometers van personenauto-inzittenden worden de voorspellingen van de aantallen doden onder het langzaam verkeer sterker verbeterd dan die van de aantallen ziekenhuisgewonden. Dit geldt misschien niet voor de slachtoffers onder fietsers.
- De geschatte invloed van reizigerskilometers van personenauto-inzittenden op het aantal slachtoffers onder het langzaam verkeer, lijkt in principe per wijze van (langzaam) vervoer te verschillen. Hierbij lijkt de invloed groter te zijn op de slachtofferaantallen onder de jongere en oudere verkeersdeelnemers, dan de onder de middengroepen.

Vergelijking met andere methoden

De statistische methode die in het instrument worden gebruikt, kan worden gezien als een gewone regressiemethode met een aantal aanpassingen en uitbreidingen. Gedeelten van de gebruikte methodiek komen in vele alternatieve technieken terug. De verschillen tussen diverse technieken kunnen in twee hoofdcategorieën worden ingedeeld:

- verschil in benadering en probleemstelling;
- verschil in statistische aannames en methodiek.

Een van de specifieke eigenschappen van het instrument is dat het gebaseerd is op een betrekkelijk expliciete specificatie van de samenhang tussen de gegevens. Hierin onderscheidt het zich van bijvoorbeeld 'autobox'-achtige (AR(I)MA)-technieken, waarin getracht wordt een zo goed mogelijk passend model te vinden uit een klasse van modellen. Op het punt van de betrekkelijk expliciete specificatie komt het model juist wel overeen met de kwartaalanalyse en de macroscopische modellen. De macroscopische modellen echter zijn bedoeld voor langere-termijnprognoses, terwijl dit nieuwe instrument in principe voor korte-termijnprognoses is bedoeld.

Wat betreft de statistische aannames en methodiek is een belangrijk verschil van de gebruikte techniek met bijvoorbeeld de gebruikelijke vormen van AR(I)MA-modellen, dat niet uitgegaan wordt van gelijke onzekerheid in de observaties. Waar het instrument (nog) niet optimaal mee om kan gaan is de statistische verdeling van de ongevals cijfers, dit in tegenstelling tot bijvoorbeeld de macroscopische modellen.

In een groot aantal gevallen blijkt het instrument goed effecten te kunnen schatten of voor afwijkingen te kunnen corrigeren. Dit gaf zijn weerslag in betrekkelijk nauwe betrouwbaarheidsmarges. Het instrument leverde in enkele gevallen echter prognoses op, waar het zelf weinig vertrouwen in leek te hebben. Soms was dit direct terug te voeren naar gebrek aan gegevens, soms leek het dat er zich ontwikkelingen voordeden die niet in het model werden opgenomen. Het model heeft in die gevallen terecht een ruime betrouwbaarheidsmarge om de prognoses aangegeven. Dit valt als een duidelijk pluspunt van het instrument te zien.

Het instrument kan nog op een aantal technische punten worden verbeterd. Het zou bijvoorbeeld beter aangepast kunnen worden aan de analyse van kleine aantallen slachtoffers of ongevallen. Ook zou een implementatie die in het geheel niet afhankelijk is van de begintoestand, op den duur de voorkeur hebben. Een dergelijke verbetering zou samen kunnen gaan met de toevoeging van uitgebreidere statistische toetsen.

Inhoud

1.	<i>Inleiding</i>	11
2.	<i>De verkeersveiligheidsmonitor</i>	12
2.1.	Gebruiksdoeleinden	12
2.2.	Globale opzet	12
2.3.	Vereisten aan het instrument	13
2.4.	Typen invoergegevens	14
2.5.	Typen uitvoergegevens: resultaten	15
2.6.	Theoretische voorbeelden	16
2.6.1.	Voorbeeld 1	16
2.6.2.	Een nog eenvoudiger model: het gemiddelde	18
2.7.	Wegingsmethode voor observaties en voorspellingen	19
2.8.	Verwijzing nadere informatie	24
3.	<i>Vergelijking met alternatieve methoden</i>	25
3.1.	Algemeen perspectief	25
3.2.	Specifieke vergelijkingen	25
3.2.1.	Inleiding	25
3.2.2.	Kwartaalanalyse	26
3.2.3.	Macromodellen	26
3.2.4.	AR(D)MA en andere klassieke modellen	27
4.	<i>Voorbeelden van toepassing</i>	29
4.1.	Toepassing gericht op prognoses	29
4.1.1.	Geval 1: alle leeftijden en alle wijzen van vervoer	31
4.1.2.	Geval 2: 0- tot 14-jarigen in de personenauto	34
4.1.3.	Geval 3: 25- tot 49-jarige personenautoinzittenden	37
4.1.4.	Geval 4: 15- tot 24-jarige bromfietzers	38
4.2.	Registratie AVV/BG- versus LMR-ziekenhuisgewonden	39
4.3.	Invloed mobiliteit snelverkeer op aantallen slachtoffers langzaam verkeer	41
5.	<i>Conclusies</i>	46
	Literatuur	48
	Bijlage 1 Modelbeschrijving	49
	Bijlage 2 Afbeeldingen 1 t/m 24	57

1. Inleiding

Dit rapport beschrijft een prototype van een instrument - *de verkeersveiligheidsmonitor* - om de ontwikkeling van de verkeers(on)veiligheid te volgen. Een van de uitgangspunten was een instrument te ontwikkelen waarbij de verkeersonveiligheid in de context van relevante omstandigheidsvariabelen (in het bijzonder, maar niet uitsluitend, de mobiliteitsontwikkeling) gevolgd kan worden.

Het doel van dit onderzoek was om enerzijds een werkend en praktisch toepasbaar systeem te ontwikkelen en te testen, anderzijds ruimte te laten voor latere specificaties en aanpassingen. Deze zouden moeten worden gedaan op basis van te verkrijgen praktijkervaring, om tenslotte te komen tot een -specificatie voor een- definitief systeem.

De opbouw van dit document is zo, dat eerst in hoofdstuk 2 een beschrijving wordt gegeven voor welke doeleinden de verkeersveiligheidsmonitor kan worden toegepast. In ditzelfde hoofdstuk volgt een beschrijving van het instrument, en wordt een voorbeeld gegeven aan de hand waarvan de werking hiervan op eenvoudige wijze wordt uiteengezet. De probleembenadering en gebruikte methode van het instrument worden in hoofdstuk 3 vergeleken met alternatieve technieken.

Vervolgens worden in hoofdstuk 4 enkele voorbeelden gegeven van toepassing van het instrument op Nederlandse verkeersonveiligheidsgegevens. Hoofdstuk 5 eindigt tenslotte met een aantal conclusies.

2. De verkeersveiligheidsmonitor

2.1. Gebruiksdoeleinden

Het in dit rapport beschreven prototype van de verkeersveiligheidsmonitor is ontwikkeld om de beantwoording van bepaalde beleidsvragen te vereenvoudigen. De beleidsmaker op het gebied van verkeer en verkeersveiligheid ziet zich namelijk regelmatig geconfronteerd met een aantal van de volgende algemene beleidsvragen:

- Hoe ontwikkelt zich de verkeersveiligheid?
- Hoe heeft de verkeersveiligheid zich in het verleden ontwikkeld?
- Welke trends zijn te ontdekken in die ontwikkeling?
- Hoe verhouden deze ontwikkelingen zich ten opzichte van 'de taakstelling'?
- Welke deelaspecten in deze ontwikkelingen vereisen aanvullende verkeersveiligheidsmaatregelen?
- Hoe groot is het effect van reeds genomen maatregelen?

Een belangrijk aspect bij bovenstaande vragen is dat het steeds gaat om zich in de tijd ontwikkelende gegevens.

Het prototype is in principe bruikbaar bij onderzoek ten behoeve van de beantwoording van alle bovenstaande vragen, maar bezit nog geen praktische, eenvoudige en allesomvattende gebruikersomgeving. Deze aanpassingen zouden later - zo nodig - alsnog uitgevoerd kunnen worden. Prioriteit is gegeven aan de inhoudelijke functionaliteit en het onderzoeken van de bruikbaarheid van het instrument.

2.2. Globale opzet

Het instrument is ontwikkeld voor het analyseren van verschillende soorten, in de tijd geordende gegevens, en van de samenhang hiertussen. Hierbij wordt gebruik gemaakt van de aanname dat, indien ontwikkelingen in de verkeersveiligheid enige samenhang vertonen met elkaar of andere waarneembare ontwikkelingen, deze samenhang gebruikt kan worden om het model beter te laten passen, en zodoende bijvoorbeeld prognoses te verbeteren.

Ook kan het instrument - op verschillende wijzen - gebruikt worden om veranderingen in een samenhang op te sporen. Dit laatste kan gebruikt worden om veranderingen in de ontwikkeling van de verkeersveiligheid bloot te leggen.

Voor het bovenstaande is het noodzakelijk dat het instrument verschillende ontwikkelingen gelijktijdig kan analyseren. Hiervoor moet het instrument de mate waarin het op de verschillende ontwikkelingen past onderling afwegen. Daarvoor is voor iedere ontwikkeling een - minimaal relatieve - maat voor zijn meeton nauwkeurigheid noodzakelijk. Een gevolg hiervan is, dat het instrument goed met (veranderingen in) dit soort gegevens om kan gaan.

2.3. Vereisten aan het instrument

Een van de uitgangspunten voor de ontwikkeling van het instrument is geweest, dat het in staat moet zijn om de ontwikkeling in de tijd van een aantal grootheden in onderlinge samenhang te analyseren, en op basis van een dergelijke analyse prognoses te genereren.

Het instrument dient de volgende functies te kunnen vervullen:

- a) Het kan gebruikt worden om de ontwikkeling van een grootheid te ontbinden in:
 - voor beleid relevante invloeden; het is in staat de samenhang tussen beleidsmatig controleerbare ontwikkelingen en de ontwikkelingen in de onderzochte grootheid te analyseren en tot op zekere hoogte te kwantificeren;
 - minder controleerbare, maar wel te specificeren invloeden, zoals economie, seizoenen, weersomstandigheden;
 - toevalsfactoren.

De algemene gedachte is dat een ontwikkeling wordt ontbonden in een som van componenten die een duidelijke en liefst interpreteerbare betekenis hebben, bijvoorbeeld een component voor de mobiliteitsontwikkeling, één of meer componenten voor de economische ontwikkelingen, een component voor een seizoenseffect en een component voor het effect van één (of meerdere) te onderzoeken maatregelen. Sommige van deze componenten kunnen worden gebruikt voor meerdere te onderzoeken grootheden. Tenslotte wordt variatie in de ontwikkeling van de te onderzoeken grootheden, die niet met behulp van betekenisvolle componenten kan worden verklaard, aan één of meer toevalscomponenten toegeschreven, welke vervolgens aan het model worden toegevoegd.

- b) Het effect van de betekenisvolle invloeden op de te bestuderen grootheden wordt idealiter in een getal uitgedrukt, bijvoorbeeld een percentage reductie. Het effect kan echter ook een ontwikkeling blijken te hebben, bijvoorbeeld 10% effect bij invoering, oplopend tot 20% na een maand of twee, en daarna dalend tot ongeveer 15%. Dergelijke effecten of gevolgen kunnen van tevoren gemodelleerd worden of achteraf geschat worden.
- c) Het instrument is in principe in staat op basis van het verleden een zinvolle voorspelling te maken van de ontwikkeling in de (al of niet) nabije toekomst. Hierbij speelt een maat voor de onzekerheid in de voorspellingen een belangrijke rol.
- d) Het instrument moet zoveel mogelijk reeds beschikbare informatie in het model kunnen toepassen. Doordat voor de beantwoording van urgente beleidsvragen over het algemeen slechts een korte reeks gegevens en weinig empirische informatie beschikbaar is, zal een model nauwelijks eenduidig geselecteerd kunnen worden indien er meerdere alternatieve modellen beschikbaar zijn. Een voorbeeld hiervan zijn 'black-box'-methoden waar, binnen vooraf gespecificeerde grenzen, een model uit een (ruime) collectie alternatieve modellen wordt gekozen op basis van één criterium. 'Autobox'-methoden zijn hier een voorbeeld van. Externe informatiebronnen - in de praktijk voornamelijk ervaringsfeiten uit

vergelijkbare situaties - zullen noodzakelijk zijn om alternatieven van elkaar te kunnen onderscheiden, en daardoor duidelijke resultaten te bereiken. Daarom is het een voordeel wanneer het instrument zo flexibel is dat onbekende aspecten geschat kunnen worden, en bekend veronderstelde aspecten juist expliciet ingebracht kunnen worden. Hierdoor blijven zo min mogelijk bekende gegevens over de verkeersveiligheid ongebruikt, en wordt voorkomen dat het instrument een grotere onzekerheid krijgt dan strikt noodzakelijk is.

De bovenstaande beschrijving van het instrument lijkt te leiden tot een soort 'schaap met vijf poten', een instrument dat voor alles bruikbaar is, maar (of misschien daardoor) voor niets optimaal is. Om dit enigszins te voorkomen is het in dit project uitgewerkte instrument in eerste instantie gebaseerd op een zo goed mogelijk uitwerking van functie a), waarbij zoveel mogelijk kennis van de verkeers(on)veiligheid wordt gebruikt (functie d). Het betreft hier voornamelijk maar niet uitsluitend het specificeren van factoren zoals expositie, risico, seizoenseffect en dergelijke. Op basis van de resultaten *daarvan* kunnen de functies b en/of c worden ingevuld.

Bovenstaande benadering beschrijft dus eerst het (verkeersonveiligheids)-proces (dat wil zeggen onderscheidt de essentiële componenten daarin) en maakt daarna prognoses op basis van *die* resultaten. Een dergelijke benadering heeft voordelen boven een benadering waarbij functie a) wordt uitgevoerd op een bepaalde manier, bijvoorbeeld met een (mogelijk) beter op de data passende techniek, terwijl de functies b) en/of c) worden uitgevoerd met behulp van een andere techniek, die op andere principes en resultaten is gebaseerd.

2.4. Typen invoergegevens

Binnen dit project wordt telkens met in tijd geordende gegevens gewerkt. Dit zijn gegevens die telkens betrekking hebben op een even lange opeenvolgende periode. Hierbij wordt aangenomen dat de omstandigheden waarop de ene observatie betrekking heeft enige overeenkomst vertoont met de omstandigheden waarop een volgende observatie betrekking heeft, zodat kennis over een bepaalde waarneming nuttig gebruikt kan worden voor het maken van een voorspelling van een volgende waarneming.

Een waarneming bestaat dus telkens uit (numerieke) gegevens die betrekking hebben op een bepaalde periode, bijvoorbeeld maart 1995. De opeenvolgende periodes hoeven niet noodzakelijk aaneensluitend te zijn; weekcijfers die alleen betrekking hebben op werkdagen zijn ook toegestaan.

In dit onderzoek en in de verkeersveiligheidspraktijk zal het voornamelijk gaan om aantallen slachtoffers per maand of totale aantallen reizigers-kilometers. Maar ook is het mogelijk dat het gaat om bijvoorbeeld de gemiddelde temperatuur in een maand op een weerstation.

Daarentegen kunnen individuele gebeurtenissen op specifieke tijdstippen niet ingevoerd worden. Dus '1 januari 2003: invoering plan A' kan niet als gegeven worden gebruikt. Hiervoor moet een *dummy variabele* worden gebruikt. Deze heeft bijvoorbeeld de waarde 'nul' voor de observaties vóór 1 januari 2003, en 'één' daarna.

De invoergegevens kunnen heterogeen van aard zijn - verschillende bronnen hebben - en hoeven ook niet steeds voor alle waarnemingen beschikbaar te zijn. Zo zouden bijvoorbeeld gordeldraagpercentages uit verschillende onderzoeken afkomstig kunnen zijn. Het instrument kan rekening houden met de daarbij onontkoombare verschillen tussen de resultaten uit de verschillende onderzoeken. Eventuele ontbrekende periodes tussen deze waarnemingen zou het instrument met behulp van de veronderstelde tijdsafhankelijkheid kunnen 'invullen'. Hiervoor is het wel nodig dat de gegevens steeds voor meerdere tijdstippen beschikbaar zijn, zodat bij de analyses enig gebruik gemaakt kan worden van de ontwikkeling van de gegevens uit het verleden. Dit laatste is een belangrijk aspect, dat deze methode onderscheidt van belangrijke andere technieken die in het verleden zijn gebruikt.

De flexibiliteit jegens ontbrekende observaties betekent natuurlijk niet dat de voorspelling daar niet gevoelig voor is: hoe meer ontbrekende gegevens, hoe meer onzekerheid.

Enigszins voor-de-hand-iggend wordt ook aangenomen dat de gegevens op een zinnige wijze in een numerieke grootheid uit te drukken zijn: aantallen, kilometers en dergelijke. Dit wil zeggen dat de gegevens niet (te) kwalitatief van aard mogen zijn, hoogstens de kwalificaties 'wel aanwezig' of 'niet aanwezig' mogen hebben. Op die manier zijn (wettelijke) maatregelen, en andere plotselinge veranderingen in grootheden, in het model te specificeren (zie ook hierboven). In de ontwikkelde implementatie is het niet mogelijk om op nominaal of ordinaal niveau gemeten grootheden met meer dan twee categorieën in het model in te passen. Wel is het mogelijk om op nominaal niveau gemeten grootheden via 'dummy'-variabelen in het model in te passen (door gebruik te maken van indicatorvariabelen voor bepaalde klassen van een variabele). Dit geldt niet voor ordinale variabelen.

Zoals reeds vermeld is een belangrijk aspect van het nieuwe instrument dat naast de -meestal gebruikte- gegevens in de tijdreeks, zoals aantallen slachtoffers en reizigerskilometers, een reeks met gegevens over de (relatieve) onnauwkeurigheid van diezelfde cijfers noodzakelijk is.

2.5. Typen uitvoergegevens: resultaten

Voor bepaalde vraagstellingen is het zeer nuttig de analyseresultaten op een specifieke manier uit het instrument te verkrijgen. Zo zal het nuttig zijn de ontwikkeling van de verkeersonveiligheid te kunnen ontbinden in een trend, datgene waarin men vaak het meest geïnteresseerd is, en in een eventueel seizoenseffect. Daarnaast kunnen de analyseresultaten bekeken worden op invloeden van exogene grootheden, zoals bijvoorbeeld de (eigen) mobiliteit en eventueel die van de botspartner (de belangrijkste in meer geaggregeerde gevallen), maar ook op economische en weersinvloeden. Indien beschikbaar, kan de invloed van een component onderzocht worden die met een maatregel samenhangt.

Bovendien zouden de werkelijke data van een waargenomen ontwikkeling over een bepaalde (meestal recente) periode vergeleken moeten kunnen worden met een prognose van diezelfde ontwikkeling, gegeven de data van het verleden tot juist voor diezelfde periode. Hierbij valt zowel te denken aan een prognose waarbij de geobserveerde exogene invloeden over die periode als

gegevens worden gebruikt, als een prognose waarbij ook voor de exogene invloeden een prognose wordt uitgevoerd. In de in dit rapport uitgewerkte voorbeelden betekende dit dat er één analyse werd uitgevoerd waarbij de laatste (meestal acht) waarnemingen niet gebruikt werden voor het schatten van het model, maar wel werden vergeleken met de prognoses van het model. Dit kan als een indicatie worden gebruikt van de mate waarin het model op de data past. Men zou aan een dergelijke analyse ook kunnen zien dat ‘de veiligheid beter/slechter heeft uitgedaan dan op basis van het verleden had kunnen worden verwacht’. Een dergelijke ontwikkeling kan ook nog het gevolg zijn van een afwijkende ontwikkeling in de mobiliteit. Vervolgens werd er een analyse uitgevoerd waarbij de (uiteindelijke of misschien volgens een scenario veronderstelde) mobiliteitsgegevens juist wel zijn meegenomen. In een dergelijk geval kan een afwijking in de ontwikkeling van het aantal slachtoffers aan het risico worden toegeschreven.

In het eerste geval kan dus met behulp van een retrospectieve analyse gekeken worden of een gevonden aantal slachtoffers meer of minder is dan het te verwachten aantal, waarbij de uiteindelijk geobserveerde mobiliteit wordt gebruikt.

In het tweede geval kan dus met behulp van een dergelijke aanpak bijvoorbeeld de gevolgen van een bepaalde toekomstige mobiliteitsontwikkeling worden onderzocht.

Tenslotte kan een zuivere prognose worden gegenereerd over de toekomstige periode waarover uiteraard qua gegevens nog niets bekend is.

Alle bovengenoemde soorten resultaten zijn in de voorbeelden toegepast. Andere combinaties van bovenstaande keuzen zijn natuurlijk ook mogelijk.

Andere soorten uitkomsten van het model, waaronder toetsen voor de relevantie van verschillende in het model toegepaste componenten, zijn nog niet direct maar slechts via de ‘log-likelihood’ beschikbaar. Met behulp van onder meer een dergelijke toets kan een model worden geconstrueerd dat slechts uit betekenisvolle factoren bestaat. In een latere versie van het model moet aan dit aspect aandacht besteed worden.

2.6. Theoretische voorbeelden

2.6.1. Voorbeeld 1

Het eerste voorbeeld van dit rapport bestaat uit een theoretische tijdreeks. De reeks heeft de vorm van een zaagtand, eerst lineair stijgend, dan lineair dalend, dan weer lineair stijgend enzovoort. (Zie *Afbeelding 1, Bijlage 2*).

Stel dat deze ontwikkeling in werkelijkheid is waargenomen en dat er geen externe verklarende variabelen (zoals mobiliteitscijfers) zijn. Een onderzoeker wordt gevraagd een voorspelling van het vervolg te maken op basis van deze gegevens. In dat geval zal de eerste vraag zijn of het geobserveerde zaagtand-effect iets is dat op zich zelf staat: ‘echt’ is, een werkelijk doorgaand fenomeen van het geobserveerde is of niet. Een tweede mogelijkheid is namelijk dat de zaagtand het resultaat is van het samenvallen van een aantal los van

elkaar staande invloeden, waaronder wellicht ook toevallige. In het eerste geval zal de zaagtand doorberekend, geëxtrapoleerd moeten worden in de toekomst. In het tweede geval moet bij het vaststellen van de betrouwbaarheid rekening worden gehouden met gelijksoortige invloeden in de toekomst. Anders gezegd, moet men bij het maken van prognoses aan het eind van de geobserveerde periode weer 'omlaag' (in het geval van het aannemen van een 'echte' zaagtand) of moet men juist verder omhoog (als men geen zaagtand-beweging aanneemt).

In dit voorbeeld zou het op empirische gronden verdedigbaar zijn aan te nemen dat de zaagtandbeweging zich voortzet, zeker als er meerdere bewegingen op en neer zijn waargenomen (in verkeersveiligheidsonderzoek bijvoorbeeld als er meerdere jaren een seizoenseffect is waargenomen). Als echter de patronen minder duidelijk zijn en als er in het geval van zaagtandbewegingen slechts weinig perioden van op- en neergang waargenomen zijn, zal er toch een keuze op basis van kennis van het onderzochte proces gemaakt moeten worden. Bij anderhalf jaar kwartaalcijfers doet men dit in de praktijk ook: bijna automatisch gaat men uit van een bestaand seizoenseffect, wat neerkomt op het gebruik van externe informatie (vakkennis). Bij een wat kortere reeks blijkt dat al snel heel belangrijk te zijn. Als van de in het voorbeeld gepresenteerde reeks slechts 25 observaties (1¼ zaagtand) bekend zouden zijn geweest, dan zou men een stijging zien gevolgd door een even grote daling met daarna weer een kleine stijging. Een eerste analyse zou tot de nu volgende resultaten leiden.

Als met behulp van een gewone lineaire regressie een prognose van deze verkorte reeks van 25 observaties wordt gemaakt blijkt deze regressie een dalende ontwikkeling voor te stellen (zie *Afbeelding 2*). Dit is extra opvallend in het licht van het feit dat de meest recente observaties juist weer een stijging laten zien. Nu weegt een gewone lineaire regressie alle observaties even zwaar (ze hebben alle evenveel invloed), hetgeen in dit geval niet redelijk gevonden kan worden.

Indien de meest recente observaties zwaarder wegen in vergelijking met de eerdere zal de prognose van een gewogen model dat daar rekening mee houdt deze laatste observaties getrouwer moeten volgen. Zo'n techniek zal echter ook minder rekening houden met verder in het verleden liggende observaties: in het uiterste geval worden de laatste observaties zo zwaar gewogen dat het effect van de eerdere observaties verwaarloosbaar is, en zal de prognoselijijn precies de laatste vijf observaties volgen (omdat de laatste vijf precies op één lijn liggen) en het model zal bovendien 'denken' dat het precies op de data past en dus zal het 'denken' dat het zeer betrouwbare prognoses oplevert. Dit laatste is uiteraard zeer riskant.

De verkeersveiligheidsmonitor kan eveneens zo worden ingesteld dat er sprake is van een lineaire trend. Het instrument is op die wijze ook toegepast op dezelfde data, resulterend in *Afbeelding 3*. Als we kijken naar de geobserveerde gegevens en de schattingen volgens het instrument, dan zien we dat het deze observaties nauwkeurig volgt, terwijl toch een lineaire trend in het instrument zou zijn ingevoerd. Dit is het gevolg van een tweetal weegfactoren in het model. De modelvoorspelling en de geobserveerde waarde krijgen elk een gewicht (de mate van onzekerheid in de respectievelijke cijfers voorstellend). Naarmate de onzekerheid in het model groter uitvalt ten opzichte

van dat van een observatie, komt de gewogen voorspelling dichter te liggen bij die geobserveerde waarde.

Door deze weging ontstaat er telkens voor iedere observatie een verschil tussen de prognose van het instrument en een observatie. Dit verschil wordt 'intern' opgevangen (via een toestandsvector, zie later). Het instrument vangt deze verschillen op door zo klein mogelijke afwijkingen (de feitelijke verschillen) in de trend of het niveau toe te laten, voor zover de gebruiker heeft gespecificeerd dat dergelijke afwijkingen toegelaten worden.

Het instrument neemt nu aan dat deze verschillen toevallig zijn geweest (dit is vrij standaard in statistische methoden) en bepaalt zijn vertrouwen in de prognoses op basis van deze verschillen, hetgeen weer de gewichten bepaalt.

De schattingsmethode kiest nu volgens het aannemelijkheidsprincipe een optimale balans tussen deze gewichten op een zodanige wijze dat de prognose het meest aannemelijk is.

Indien er geen observaties meer zijn, wordt noch de toestandcomponent aangepast, noch gecorrigeerd met geobserveerde waarden. Het model in dit voorbeeld bestaat dan slechts uit een lineaire trend. Het model stelt de onzekerheid van zijn voorspellingen vast aan de hand van de mate waarin het steeds zijn prognoses moet aanpassen aan de hand van de geobserveerde waarden in het verleden.

Terugkomend op de resultaten van het voorbeeld is het duidelijk dat het instrument een stijgende ontwikkeling verwacht, maar de onzekerheid daarover - aangegeven met het (omvangrijke) grijze gebied - redelijk betrouwbaar heeft geschat. Het model houdt kennelijk rekening met de mogelijkheid van zowel een nieuwe stijging in de ontwikkeling van de reeks als een nieuwe daling, hoewel een gelijksoortige daling als na tien stijgende observaties niet onmiddellijk verwacht wordt. De eerlijkheid gebiedt te stellen dat het nieuwe instrument meer parameters omvat dan de lineaire regressie waarmee is vergeleken, en dus ogenschijnlijk beter kan passen, maar het is wel geschat op dezelfde data.

In wezen is het nieuwe instrument in dit voorbeeld niet essentieel verschillend van de lineaire regressie. Het instrument kan gezien worden als een in de tijd-variërende lineaire regressie: de steilheid van de regressielijn en het snijpunt met de (verticale) y-as zijn tijdsafhankelijk geworden: aangepast bij de weging. In het huidige voorbeeld kan dit natuurlijk ook met de hand: in een oogopslag kan men zien dat de grafiek eerst tien stappen omhoog gaat, dan tien stappen omlaag, enzovoort. Het bijzondere is nu hoe met de hierbij horende variatie in de parameterwaarden rekening gehouden wordt bij het opstellen van de prognoses. Dit aangeven van de onzekerheid in de voorspellingen is een belangrijk aspect dat naast de weging van de gegevens het nieuwe instrument doet verschillen van de gewone lineaire regressie.

2.6.2. Een nog eenvoudiger model: het gemiddelde

Voor het verkrijgen van een enigszins gedetailleerde beschrijving van het model wordt aan de hand van Voorbeeld 1 een nog eenvoudiger model gebruikt: het gemiddelde van alle observaties. *Afbeelding 4* laat de

‘prognoses’ met behulp van het gemiddelde zien voor de reeks. *Afbeelding 5* laat hetzelfde zien maar dan met het nieuwe instrument. Net als bij de lineaire regressie wordt ook hier een tijd variërend ‘gemiddelde’ berekend, terwijl dat bij het gewone gemiddelde niet tijdsafhankelijk is. In feite is het berekenen van het gemiddelde gelijk aan het uitvoeren van een lineaire regressie met alleen een snijpunt met de y-as (intercept), zij het dat het op een eenvoudiger manier wordt berekend. Met het nieuwe instrument kan dit snijpunt met de tijd kan variëren, dat wil zeggen: de lijn op een bepaald tijdstip heeft een ander snijpunt met de y-as dan de lijn op een ander tijdstip. In het uiterste geval - anders dan het gebruiken van het algemeen gemiddelde - zou dat betekenen dat er telkens een horizontale lijn wordt gekozen die precies door het laatste punt heen gaat, en niet, zoals bij het gewone gemiddelde om het gemiddelde van de observaties. Hoewel dit helaas niet zo goed zichtbaar is in *Afbeelding 5*, gaat de lijn niet precies door de observaties heen. Er wordt op een speciale manier een middenweg gekozen: er worden wel telkens schattingen gedaan die zeer dichtbij de geobserveerde waarden liggen. Dit komt omdat het model met enkel een constante factor de gegevens slecht beschrijft. Hierdoor is de onzekerheid in de modelvoorspellingen zeer groot en wordt voornamelijk op de observaties vertrouwd. Hierin is de methode die in het nieuwe instrument wordt gebruikt overigens niet uniek: er bestaan meer methoden om een dergelijke keuze te maken.

2.7. Wegingsmethode voor observaties en voorspellingen

De methode voor het kiezen van de middenweg/gewichten die hier gebruikt wordt is een variant op het Kalman-filter (voor de huidige context, de systeemtheorie, zie bijvoorbeeld Harvey, 1989). De methode komt op het volgende neer.

Stel we hebben een schatting van de waarde van de reeks op tijdstip t : een gemiddelde en een (relatieve) ‘betrouwbaarheidsmaat’ (zeg: een gemiddelde plus marges). Deze schatting komt tot stand door een functie toe te passen op een (eindig dimensionale) *toestandsvector*. Deze toestandsvector bestaat uit interpreteerbare componenten, en beschrijft alle relevante componenten van het te onderzoeken (verkeersveiligheids)proces. Zo is de vector in het eerste voorbeeld tweedimensionaal, hij bestaat uit twee componenten: een niveau-component en een steilheidscomponent. In het eenvoudiger vervolg van het voorbeeld bij het ‘gemiddelde’ model bestaat de toestandsvector slechts uit één component: de niveaucomponent. In de volgende voorbeelden wordt de toestandsvector uitgebreid met seizoenscomponenten en andere niveau- en steilheidscomponenten plus nog enkele speciale componenten. De onderzoeker is vrij in het samenstellen van de componenten. In feite bestaat de schatting van de waarde van de reeks op tijdstip t dus eigenlijk uit de schatting van een toestandsvector.

Uitgegaan wordt dus van een schatting van de waarde van de reeks op tijdstip t plus een (on)betrouwbaarheidsmaat. Vervolgens wordt dit gegeven vergeleken met de geobserveerde waarde voor tijdstip t . In de praktijk verschillen deze twee waarden natuurlijk. Nu wordt de definitieve schatting voor tijdstip t vastgesteld door een gewogen gemiddelde te nemen van de schatting en de observatie, die elk een eigen maat voor hun onbetrouwbaarheid hebben. Hierbij zal de nieuwe waarde dichter bij de waarde met de grootste (relatieve) betrouwbaarheid liggen dan bij de andere waarde, omdat het model meer

vertrouwen in de juistheid van de waarde met de grootste (relatieve) betrouwbaarheid heeft. Als er een schatting van de observatie op tijdstip t wordt gemaakt met een ruimere marge dan van de observatie zelf, dan zal het model vooral uitgaan van de observatie. Het omgekeerde gebeurt ook: bijvoorbeeld als de observatie 'missing' is, zal het model zelfs volledig uitgaan van de schatting van de waarde van de reeks op tijdstip t . Als de marges precies gelijk zijn, dan zal het model de waarden precies middelen.

Vervolgens wordt (dit gaat in werkelijkheid direct) de nieuwe schatting terugerekend naar een nieuwe schatting van de toestand(svector) en wordt daarvan ook de (relatieve) betrouwbaarheidsmaat berekend. Daarna kan weer een prognose voor de toestand, en daarmee de nieuwe waarde van de reeks op het tijdstip $t+1$ plus de bijbehorende (relatieve) betrouwbaarheidsmaat worden berekend, en wordt de cyclus herhaald.

Blijft over wat voor de eerste observatie moet worden gebruikt, daarvoor is in het algemeen geen verleden beschikbaar om een prognose op te baseren. Voor de toestandsvector van de eerste observatie wordt in het nieuwe instrument een prognose van nul (0-vector) gebruikt, met een betrouwbaarheidsmaat die (verwaarloosbaar) klein is ten opzichte van de betrouwbaarheidsmaat van de eerste observatie, bijvoorbeeld resulterend in een marge van nul plus of min 1000. In dit voorbeeld betekent dit dat de definitieve schatting voor de eerste observatie praktisch gelijk is aan de eerste observatie. In ingewikkelder gevallen zoals bij latere voorbeelden zal het een (klein) aantal observaties duren (meestal vier in die gevallen met een seizoenseffect voor vier kwartalen) totdat een schatting gevonden wordt die verwaarloosbaar afhankelijk is van de arbitraire keuze van de keuze van de begintoestand. Overigens is deze gang van zaken niet verschillend van de gang van zaken bij alternatieve modellen, zoals bij onder andere AR(I)MA technieken, waar bijvoorbeeld door het nemen van verschillen (de I in ARIMA) de eerste observaties vervallen. Dit geldt ook voor het 'moving average' - (MA-) en autoregressieve (AR-) gedeelte.

Samenvattend verkrijgen we de volgende werkwijze:

- 1) Het instrument gaat ervan uit de toestand van een proces te kunnen beschrijven met behulp van een toestandsvector. Verondersteld wordt dat de (meest belangrijke) informatie uit het verleden over de observaties in de toekomst is gevat in deze toestand. De informatie in de toestandsvector zal voornamelijk bestaan uit waarden van bepaalde componenten plus hun onzekerheid: meestal de trend-componenten, de relatieve invloed van het seizoen, het (slachtoffer- of ongevals-)risico, werkelijke mobiliteitscijfers en effecten van maatregelen. Deze informatie noemen we dus de 'toestand' van het systeem of het verkeersveiligheidsproces. Het instrument gaat uit van alle beschikbare informatie over de begintoestand/uitgangspositie van de reeks. In de praktijk zal dat 'weinig of niets' zijn, en zal worden uitgegaan van een zo onzeker mogelijke begintoestand. Dit hoeft echter niet altijd het geval te zijn, en dan is het mogelijk gebruik te maken van de wèl bestaande informatie.
- 2) Uitgaande van de vorige (uitgangs)toestand wordt een prognose gegeven van de - volgens het model gespecificeerde - nieuwe toestand, plus een onbetrouwbaarheidsmaat daarvan. Deze onbetrouwbaarheidsmaat is deels

het gevolg van de onzekerheid in de vorige toestand, en deels het gevolg van de interne onzekerheid van het model (meestal als parameter te schatten). Dit laatste is het gedeelte dat 'leert' hoe betrouwbaar de prognoses zijn.

- 3) Uit de voorspelling van de toestand wordt een voorspelling van de bijbehorende observatie berekend, bijvoorbeeld de mobiliteit. Deze wordt vergeleken met de werkelijke observatie, en aan de hand van deze vergelijking wordt een nieuwe (verbeterde) schatting van de toestand van het systeem berekend. Als er toevallig geen informatie is geobserveerd, bijvoorbeeld als er geen waarnemingen in de OVG-steekproef zitten, dan zal de toestand (op dat punt) dus niet worden aangepast.

Opmerkingen:

- Doordat de toestand wordt aangepast aan de hand van geobserveerde gegevens plus een betrouwbaarheid daarvan, kan het model zeer goed met gegevens van verschillende kwaliteit omgaan. Zo zal het geen problemen hebben met de vrij plotselinge toename van de omvang, en dus de betrouwbaarheid van de OVG-steekproef, zoals zich vanaf 1994 heeft voorgedaan.
- Ook kan het omgaan met het af en toe ontbreken van (meerdere) observaties.
- Ook kan het model op natuurlijke wijze omgaan met verschillende waarnemingen van 'het zelfde', zoals LMR-gegevens en AVV/BG-gegevens. In feite kunnen de observaties op hun beurt weer schattingen zijn op basis van verschillende gegevens: deze moeten als observaties 'waarnemingen' opleveren plus bijbehorende 'varianties'. Op deze wijze kan bijvoorbeeld het aandeel van de AVV/BG-gewonden worden opgehoogd met behulp van verschillende onderzoeksresultaten betreffende dit aandeel. Deze resultaten hoeven dus niet op dezelfde wijze verkregen te zijn, ze moeten alleen hetzelfde meten. Dit kan zowel 'parallel' gebeuren, door tegelijk zowel AVV/BG- als LMR-gegevens te analyseren, als 'in serie': elkaar in de tijd opvolgend. Dit laatste is te gebruiken als verschillende onderzoeken aan elkaar 'geknoopt' moeten worden. Dit gaat het beste met een overlappende periode, zodat het ene gegeven op termijn door het andere gegeven kan worden vervangen. De verschillende onderzoeken moeten dan wel voldoende lang 'lopen'.
- In de toestand kunnen verschillende *soorten* componenten worden opgenomen. Deze componenten kunnen van belang zijn voor verschillende geobserveerde grootheden. Zo kunnen er componenten voor de mobiliteit worden opgenomen. Deze zijn zowel van belang voor de observaties van de mobiliteit zelf, als voor de observaties van ongevals- en/of slachtoffergegevens. Al deze gegevens zullen van invloed zijn op de schattingen van alle componenten. Zo geldt voor de mobiliteitscomponent dat eventuele (grote) meetfouten in de mobiliteitscijfers kunnen worden gecompenseerd door informatie over de slachtoffercijfers. In het later te geven voorbeeld over de bromfietzers van 15 tot 24 jaar (geval 4, § 4.1.4.) komt dit duidelijk naar voren.

Het verschil tussen de hierboven genoemde *soorten* componenten en de onderstaande *typen* componenten is meer taalkundig, het eerste heeft te maken met de samenhang met de ‘buitenwereld’, het tweede heeft te maken met de modeigenschappen.

- Er kunnen verschillende *typen componenten* worden geschat. In principe is de onderzoeker vrij specifieke componenten te ontwerpen en aan het model toe te voegen. Er bestaan echter een aantal vrij ‘standaard’ componenten; hieronder volgen een aantal hiervan.

‘Level’-componenten. Dit zijn componenten waarvan de waarde op het volgende tijdstip (eventueel in verwachting) dezelfde waarde heeft als het huidige tijdstip. Deze componenten worden gebruikt om het niveau van het een of ander aan te geven. Een goed voorbeeld is de registratiegraad van iets of het aandeel van iets. Indien gewenst kan de onderzoeker toelaten dat de waarde van de component iets met de tijd kan veranderen, dan blijft de waarde van de component in verwachting het zelfde. Dit wordt in samenhang met het Kalman-filter geschat. Indien het niveau een vaste waarde heeft, dan neemt men aan dat de onzekerheid (variantie) in de prognose van die component gelijk is aan nul. Indien de component met de tijd iets kan veranderen, dan krijgt de variantie van de component een positieve waarde, die bijna altijd geschat wordt. In dat laatste geval volgt de component een ‘random walk’ en wordt ook wel onder die naam gebruikt. Het bovengenoemde ‘gemiddelde-model’ (in § 2.6.2) bestaat uit één ‘random walk’-component. Een klassiek voorbeeld van een ‘random walk’ is de totale hoeveelheid winst bij een reeks kansspelen: de hoeveelheid winst blijft gelijk totdat er weer een nieuw spel wordt gespeeld. In principe zou men kunnen stellen dat na ieder spel de ‘random walk’ weer opnieuw begint, maar dan met nieuwe beginwaarde(n). De ‘random walk’-component is tijdsafhankelijk en draagt zich over in de tijd en is dus wezenlijk verschillend van het door toeval een keertje een beetje hoger of lager uitvallen van een observatie.

‘Level’- en ‘drift’-componenten. In essentie zijn drift-componenten ook weer level-componenten, alleen het gebruik ervan verschilt. De waarde van een drift-component wordt telkens bij één of meer level-componenten opgeteld. Deze laatste worden dan trend-componenten genoemd. Zie de normale lineaire regressie van voorbeeld 1. Als nu - vereenvoudigd - de waarde van de drift-component telkens d is, en de beginwaarde van de trend gelijk is aan a , dan is op tijdstip 1 de waarde van de trend $a+d$ en die van de drift nog steeds d . Op tijdstip 2 en verder blijft de waarde van de drift d , terwijl de waarde voor de trend op tijdstip 2 gelijk is aan $a + 2d$, op tijdstip 3: $a + 3d$ en op tijdstip t : $a + dt$. De trend is dus lineair. Het ‘level’ van de ‘drift’ is dus de steilheid van de trend. In vergelijking met een gewone lineaire regressie functioneert ‘ d ’ dus als slope-parameter en ‘ a ’ als intercept. Het enigszins van waarde veranderen van een drift-component ‘ d ’ kan zich dus manifesteren in bijvoorbeeld een afvlakking van een stijgende tendens. Met het enigszins van waarde veranderen van a ligt het ingewikkelder. Op een dergelijke wijze verkrijgt men een trend die soms op een - tijdelijk zelfde - hoger niveau komt te liggen en soms een andere - eveneens tijdelijk een zelfde - helling krijgt. In deze zin is het model te vergelijken met een regressietechniek met in de tijd-variërende coëfficiënten.

seizoenscomponenten (bijvoorbeeld gebruikt bij de analyses van kwartaal-cijfers). Dit zijn tijdelijke (relatieve) afwijkingen waarvan de periode van terugkeer van de component op het zelfde niveau na een van tevoren vastgestelde periode plaats vindt. Bij kwartaalcomponenten is dat na vier tijdstippen, bij maandcijfers na twaalf observaties. De som van de seizoenscomponenten is telkens gelijk aan nul, tenzij ook deze component enigszins in waarde mag veranderen. In dat geval is de som gelijk aan een getal dat in verwachting gelijk is aan nul en - zoals alle afwijkingen - op een zelfde manier is verkregen als de afwijkingen in de level- en drift-componenten.

'cyclische' componenten. Dit zijn eveneens periodieke effecten waarvan de vorm in grote mate vaststaat maar de golflengte - de lengte tot terugkeer - juist niet. Dit is in tegenstelling tot seizoenscomponenten waar de golflengte juist van tevoren vaststaat. Eventueel kan de sterkte (amplitude) van de component uitsterven of juist toenemen. De amplitude en de golflengte worden door een parameter geschat welke dus niet tijd variërend is (zoals bij de trend). Wel kan het faseverloop van de cyclus met de tijd afwijken van een deterministische ontwikkeling.

In het algemeen kunnen de volgende *componenteigenschappen* worden onderscheiden:

- De componenten kunnen *tijdelijk* van aard zijn, bijvoorbeeld effectief zijn *tot* een bepaalde periode. In de komende voorbeelden van hoofdstuk 4 is een dergelijke tijdelijke component gebruikt om een correctie toe te passen voor de afwezigheid in het OVG van 0- tot 11-jarigen in de jaren tot 1994. Ook kan een tijdelijke component juist *vanaf* een bepaald moment effectief zijn bij invoering van een nieuwe maatregel (in het geval van een interventieanalyse).
- De *vorm* van een component kan *tijd(in)variant* zijn. Dit wil zeggen dat het effect van bijvoorbeeld een interventie niet (of juist wel) verandert in de tijd. In het algemeen zal het meer realistisch blijken in een dergelijk geval een in de tijd variërend effect te modelleren. Zo zal het aandeel van de 0- tot 11-jarigen in het OVG niet constant blijken over de jaren. Ook kan bij een interventie, bijvoorbeeld door geleidelijke acceptatie, het effect met de tijd veranderen. Dit tijdsafhankelijke effect is overigens niet altijd goed te schatten, bijvoorbeeld omdat het niet te onderscheiden is van de trend. In een dergelijk geval zal men genoodzaakt blijken een invariant effect te schatten.
In praktijk kan het invariant zijn worden ingesteld, door de parameter van het model die bij de tijdsafhankelijkheid hoort op nul te zetten en te houden. Deze parameter zit in het te schatten gedeelte van de interne onzekerheid. Op deze wijze kunnen zuivere 0-1-effecten worden gemodelleerd. Ook kan een vaste drift of trend of een vast seizoenseffect worden gemodelleerd.

2.8. Verwijzing nadere informatie

In *Bijlage 1* wordt een uitvoeriger, technischer en preciezer beschrijving van het instrument gegeven. Vergelijkingen met alternatieve methodes worden gedaan in hoofdstuk 3.

Wat betreft informatie uit de literatuur kan als belangrijke bron het hoofdstuk "Structural Time Series Models" van Harvey & Shepard in "Handbook of Statistics Vol 11, 1993" gelden. In Harvey & Shepard (1993) wordt in een bondig overzicht van de theorie van structurele tijdreeksmodellen gegeven. Hierin worden ook recentere ontwikkelingen beschreven zoals het diffuus Kalman-filter (bijvoorbeeld de Jong, 1988) en de 'exact score' van Koopman & Shepard (1992). Van deze laatste is dankbaar gebruik gemaakt bij de implementatie in dit rapport. In het uitgebreide 'standaardwerk' Harvey (1989) wordt deze 'exact score' slechts terloops, en in het compactere Kendall & Ord (1990) helemaal niet behandeld. De ontwikkeling van het diffuus Kalman-filter - een oplossing voor de arbitraire keuze voor de begintoeestand door de variantie feitelijk op oneindig te zetten - is nog in volle gang (zie bijvoorbeeld Koopman, 1997). Daarnaast doen zich ontwikkelingen voor in de sfeer van simulatieschatters waarmee gedeeltelijk normaalverdeelde systemen kunnen worden geschat.

3. Vergelijking met alternatieve methoden

3.1. Algemeen perspectief

De in het instrument gebruikte statistische methode kan worden gezien als een gewone regressiemethode met een aantal slim gekozen aanpassingen en uitbreidingen. De methode is uitermate flexibel, zodat het beschrijven van de verschillen tussen deze methode en een aantal eenvoudiger alternatieven niet altijd zinnig is. Zoals uit § 2.6.2. blijkt is de methode in zijn primitiefste vorm te gebruiken als een omslachtige wijze om het gemiddelde van een rij getallen te berekenen. Als zich echter een plotselinge verandering in het gemiddelde van deze reeks voordoet, kan door middel van het update-mechanisme een dergelijk breekpunt redelijk goed worden aangegeven, en kunnen nadien nog realistische prognoses worden gegenereerd. In de onzekerheid van de prognoses wordt dan wel rekening gehouden met het zich (kennelijk onverklaard) voordoen van plotselinge niveauwisselingen.

De methode is tot het accommoderen van dergelijke veranderingen in staat omdat het met in de tijd-variërende coëfficiënten om kan gaan. Dit betekent als het ware dat (de realisatie van) het ‘model’ voor de eerdere observaties kan verschillen van het model dat voor de latere observaties wordt gebruikt. Het instrument verschilt op dit punt essentieel van de gewone, klassieke regressiemodellen. Er bestaan wel meerdere andere technieken die ook deze eigenschappen hebben; het hier gebruikte model heeft op dit punt dus overeenkomsten met die modellen, waar het echter op andere punten weer van verschilt.

De methode van het instrument gaat expliciet om met onzekerheid in de gebruikte variabelen. In tegenstelling tot gewone regressiemodellen (niet errors-in-variables-modellen) wordt in dit model expliciet gebruikgemaakt van de onzekerheid in de gebruikte variabelen (de onafhankelijke variabelen, de regressie variabelen). In de macroscopische modellen, waar met behulp van -overigens niet-lineaire- regressieanalyse slachtofferratio's werden geanalyseerd (bijvoorbeeld Oppe et al, 1988), wordt voor de noemer bij de risico berekening de mobiliteit als exact gegeven gebruikt.

Ook gaat het model niet per se uit van gelijke betrouwbaarheid per observatie (homoscedasticiteit). Hierin verschilt het model van de meeste tijdreeks-analysetechnieken. De technieken die zijn gebruikt bij de macroscopische modellen zijn op een aantal daaraan gerelateerde punten beter.

3.2. Specifieke vergelijkingen

3.2.1. Inleiding

Zoals enigszins uit het voorgaande en uit de voorbeelden in het volgende hoofdstuk kan worden opgemaakt, is het nieuwe instrument ten dele werkelijk verschillend van de meeste alternatieven. De verschillen kunnen worden onderverdeeld in twee hoofdgroepen:

- 1) verschillen behorende bij de modelspecificatie;
- 2) verschillen behorende bij de statistische methodiek..

De mate waarin de verschillen zich voordoen is afhankelijk van de te vergelijken methodiek. In dit hoofdstuk zullen de verschillen en overeenkomsten worden besproken tussen de methode van het nieuwe instrument en de methode zoals die bij de SWOV-kwartaalanalyses en -macromodellen is toegepast. Ook zal een vergelijking gemaakt worden met de belangrijkste alternatieve klasse tijdreeksanalyse: de AR(I)MA technieken.

3.2.2. *Kwartaalanalyse*

De kwartaalanalyse, zoals tot nog toe uitgevoerd op de SWOV, is net als het nieuwe instrument bedoeld voor een snel beschikbare, korte-termijnprognose van onveiligheidsgegevens. Bij de kwartaalanalyse wordt door de laatste vijf jaar een lineaire trend gefit met een multiplicatieve seizoenscorrectie. De oudere observaties worden weggelaten. Gezien de technische mogelijkheden destijds is dit een goede keuze geweest. Het nieuwe instrument zal niet beperkt hoeven worden tot de laatste vijf jaren, omdat het een belangrijke trendverandering in de laatste jaren zelf zal oppakken. Als verbetering op de oude techniek, zal het in zijn betrouwbaarheidsmarges rekening houden met de veranderingen in deze vroegere trend, die zich in de toekomst weer kunnen voordoen. De kwartaalanalyse verwaarloosde alle oudere informatie. Als zich dat voordoet houdt de kwartaalanalyse geen rekening met het feit dat in het verleden zich een andere tendens voor heeft gedaan. Het nieuwe instrument daarentegen, zal het eventueel variëren van de trend meenemen in de onzekerheid van zijn prognoses. Populair gesteld zal het rekening houden met gelijksoortige variaties in de trend. Ontwikkelingen waarin de trend langere tijd gelijk blijft zal het nieuwe model als betrouwbaarder zien dan ontwikkelingen die minder gelijkmatig verlopen.

Daarnaast kan men bij de nieuwe methode bijvoorbeeld kiezen expositiematen in het nieuwe instrument toe te passen. Hierdoor kan de ontwikkeling beter verklaard worden.

De belangrijkste verschillen met de kwartaalanalyses zijn dus: 1) rekening houden met afwijkingen in het verleden, en 2) meer informatie kunnen gebruiken voor het verkrijgen van de schattingen. Het lijkt verstandig de kwartaalanalyses in het vervolg uit te voeren met behulp van de nieuwe methode.

3.2.3. *Macromodellen*

De nieuwe methode als schattingsmethode zou, zij het met enige uitbreiding en aanpassing, toegepast kunnen worden op reeds bestaande macroscopische modellen van de SWOV. Bij deze modellen wordt net als bij de meeste implementaties van het nieuwe instrument - vereenvoudigend - aangenomen dat het aantal doden het product is van mobiliteit en een risico dat een functie van de tijd is:

$$d(t) = \text{mob}(t) * r(t)$$

$Mob(t)$ is geobserveerd, $r(t)$ is een functie van t (tijd). In de functie kunnen enige parameters zijn verwerkt.

Belangrijk is dat $r(t)$ - en de bijbehorende parameters - in feite op basis van een langere periode wordt geschat en op basis van een sterke modelassumptie: meestal een exponentiële risicoafname. Deze keuze is belangrijk omdat bij de macroscopische modellen vooral naar langere-termijnprognoses wordt gekeken. In dat geval is men minder geïnteresseerd in plaatselijke (korte-termijn)afwijkingen van het model maar veel meer in de 'grote lijn'. Het nieuwe instrument is daar juist niet op gericht.

Wel is het zo dat gedeelten van de hier benutte technieken in de macroscopische modellen gebruikt zouden kunnen worden. Zo zou bijvoorbeeld een 'Extended Kalman-filter' in combinatie met het gebruik van differentievergelijkingen in plaats van 'gesloten' functiedefinities op macroscopische modellen toegepast kunnen worden. Bijvoorbeeld zou de gesloten functiedefinitie $f(t) = \exp(a t + b)$ met behulp van de differentievergelijking $f(t) = f(t-1) + p$, met p de afgeleide van $f(t)$ in $t-1$ in het model gebruikt kunnen worden. Een belangrijk aspect van het nieuwe instrument is dus de benadering van het probleem. Deze is meer elementair, men gaat uit van de prognose van tijdstip t naar $t+1$ anders dan het schatten van de algemene lijn zoals bijvoorbeeld bij de macromodellen is gedaan. Bovendien dwingt de methode van het nieuwe instrument het kwantificeren van de onzekerheid in de gegevens af. Bij de andere methoden gebeurt dat veel minder.

Op het punt van het log-lineaire model en de vormen van de ontwikkelingen (exponentieel, logistisch) is het nieuwe instrument beperkter dan de modellen die in de macroscopische modellen zijn toegepast.

De verschillen tussen de macroscopische modellen en het nieuwe instrument zijn aldus veelvuldiger dan de verschillen met de kwartaalanalyses. Het belangrijkste verschil is toch het gevolg van de onderliggende toepassing: lange-termijnontwikkelingen tegenover korte-termijnontwikkelingen. Omdat een gedeelte van de technologie van het nieuwe instrument rechtstreeks in de macromodellen kan worden ingezet, is het niet zinnig de verschillen in implementatie als wezenlijk te onderscheiden. Het verschil komt dan in feite neer op de vraag in welke mate je de laatste observaties weegt: schat je de langere-termijnprognoses op basis van het gemiddelde (macromodellen) of juist met een zekere nadruk op de latere observaties (techniek nieuw instrument). Het is niet uitgesloten dat uiteindelijk toch voor het laatste moet worden gekozen.

3.2.4. *AR(I)MA en andere klassieke modellen*

Uit tabel 1 van Harvey & Shepard (1993) kan worden geleerd dat de prognoses van het model van het nieuwe instrument met trend en seizoens-effect gelijksoortig zijn aan die van bijvoorbeeld het Holt-Wintersmodel. Kendall & Ord (1990) geven op hun pagina 147 als voorbeeld hoe men een ARIMA(2,0,1)-model in de vorm van het huidige instrument kan gieten. De ARIMA-vorm is echter een algemenere dan die hier is gebruikt. Op enkele uitzonderingen na zijn alle ARIMA-modellen in een hier gebruikte (toestandruimte)-vorm te gieten (Akaike, 1974). Het verschil tussen het gestructu-

reerde model van het nieuwe instrument en de specificatie van de ARIMA-modellen is dat een tijdreeks ontbonden wordt in een aantal interpreteerbare componenten (er een structuur in wordt aangelegd), en niet in een aantal abstracte grootheden. Zo wordt het zogenaamde stationair maken (het gemiddeld gelijk aan nul maken) van een tijdreeks in de ARIMA-context gedaan, door het werken met verschillen tussen observaties op tijdstip t en $t-1$ enzovoorts. In het nieuwe instrument wordt dit gedaan door het toevoegen van een drift-component aan de toestandsvector. In het eerste geval wordt in feite aangenomen dat de drift dan overal het zelfde is (wat overigens bij de hier gebruikte voorbeelden bijna overal het geval bleek); in het nieuwe instrument mag de drift met de tijd veranderen. Het nieuwe instrument heeft dus ook algemener kanten. Tenslotte gaat het nieuwe instrument op een zeer doorzichtige manier om met verschillen in de betrouwbaarheid van individuele gegevens.

4. Voorbeelden van toepassing

4.1. Toepassing gericht op prognoses

In het nu volgende hoofdstuk worden een aantal voorbeelden uitgewerkt aan de hand van gedisaggregeerde gegevens. Hierbij zijn de aantallen reizigerskilometers en aantallen doden en ziekenhuisgewonden verzameld uit bestanden van respectievelijk het onderzoek verplaatsingsgedrag (OVG), de Afdeling Basisgegevens van de AVV (AVV/BG) en de Landelijke Medische Registratie (LMR). Deze aantallen zijn voor elk kwartaal van de periode 1985 tot en met 1997 verzameld, evenals schattingen van hun varianties. De gegevens zijn telkens onderscheiden naar vervoerwijze en leeftijdsklasse. De categorieën slachtoffers per vervoerwijze (passagier en bestuurder samengevoegd) zijn personenauto-inzittenden, motorfietsers, bromfietsers, fietsers, voetgangers, overige vervoerwijzen en alle vervoerwijzen samen. De categorieën slachtoffers per leeftijdsklasse zijn 0- tot 14-jarigen, 15- tot 24-jarigen, 25- tot 49-jarigen, 50- tot 64-jarigen, 65 jaar en ouder, en alle leeftijdsklassen samen. Voor de aantallen slachtoffers is een benadering van de variantie op basis van de Poisson-aanname gebruikt, waarbij de variantie van het aantal is geschat door het geobserveerde aantal. Het gaat hierbij om een dubbele benadering: 1) de variantie wordt geschat op basis van het aantal slachtoffers, niet het aantal ongevallen; dit blijkt in de praktijk eerder een onderschatting van de variantie op te leveren; 2) de gegevens worden verder gebruikt als ware het normaalverdeelde grootheden; dat wil zeggen, er wordt aangenomen dat de onzekerheid in de aantallen slachtoffers een normale verdeling volgen. Op den duur kan op beide punten een betere schatting gebruikt worden. Het meest voor-de-hand-liggende is op den duur een iteratief proces te gebruiken, waarbij aanvankelijk de geobserveerde aantallen slachtoffers worden gebruikt voor de variantieschatting, waarna het model op gelijke wijze als nu wordt geschat. Vervolgens kunnen dan de voorspellingen gebruikt worden als schattingen van de variantie, waarna het proces herhaald wordt totdat er zich geen veranderingen meer voordoen. Deze werkwijze is vergelijkbaar met de *'iteratively reweighted least squares'*-methode. In zeer recente literatuur wordt aandacht besteed aan een dergelijk loslaten van normaliteitsassumpties van fouttermen.

Voor de varianties van de reizigerskilometers zijn schattingen gebruikt aan de hand van de benaderingsmethode volgens het CBS (CBS, 1993). Er is verondersteld dat de ontwikkeling in de bevolkingssamenstelling binnen de periode 1985-1997 verwaarloosd kan worden, voorzover die niet in de ontwikkeling van de reizigerskilometers terug te vinden is of door een eenvoudige trend in het risico kan worden opgevangen.

Verzameling van de gegevens leverde voor iedere combinatie van wijze van vervoer en leeftijdsklasse plus marginalen (alle vervoerwijzen en leeftijden) een tabel op met geobserveerde aantallen slachtoffers en reizigerskilometers per kwartaal, en een gelijkvormige tabel met bijbehorende betrouwbaarheidsmaten (in feite de variantie). Deze tabellen zijn als invoer voor het instrument gebruikt.

Voor iedere combinatie is geprobeerd de laatste acht observaties (de jaren

1996 en 1997) te voorspellen uit de eerdere observaties (tot en met 1995). Daarmee kan een indruk worden gekregen van de mate waarin de prognoses uit het model realistisch zijn. Daarna is ongeveer eenzelfde analyse uitgevoerd, gebruikmakend van gegevens van de in werkelijkheid geobserveerde reizigerskilometers voor de jaren 1996 en 1997. Hiermee kan op een zinvolle wijze de verwachte verkeersonveiligheidsontwikkeling in termen van slachtofferaantallen worden vergeleken met de geobserveerde ontwikkeling, zonder dat fouten in de modelvoorspelling van de mobiliteit een rol spelen. Tenslotte is een prognose voor de komende jaren (1998, 1999) uitgevoerd, met behulp van de gegevens tot en met 1997.

Er is verondersteld dat - afgezien van een (multiplicatief) seizoenseffect - de risico-ontwikkelingen exponentieel dalend zijn geweest (en blijven voor 1998 en 1999) en dat de ontwikkeling van de mobiliteit - eveneens afgezien van een multiplicatief seizoenseffect - goed met een exponentiële ontwikkeling beschreven kan worden in de periode 1985-1999. Hoe langer de periode is waarin de toekomstige ontwikkeling moet worden voorspeld, hoe kritischer deze aanname is. Het zal daarom aan te raden zijn niet al te lange prognoses te maken aan de hand van deze aannames. In het geval waarbij langere-termijnprognoses gewenst zijn kan men beter, zoals in voor dat doel ontwikkelde prognosemethoden is gedaan, extra aannames over de vorm van de ontwikkeling van de mobiliteit maken (anders dan de aanname van exponentiële groei) en, indien naar leeftijd wordt gedisaggregeerd, zal men rekening moeten houden met de (te verwachten) ontwikkeling in de leeftijdsopbouw van de bevolking. Dat is in de hier gebruikte modellen niet gebeurd. Zoals gebruikelijk wordt aangenomen dat het te verwachten aantal slachtoffers gelijk is aan het product van een expositiemaat (in dit geval reizigerskilometers) en een risicomat.

Met de aanname van exponentiële ontwikkelingen voor deze ontwikkelingen wordt een log-lineair model voor-de-hand-liggend. Meer specifiek is het volgende algemene (log-lineaire) model aangenomen:

$$\begin{aligned}d(t) &= v(t) + rd(t) \\z(t) &= v(t) + rz(t)\end{aligned}$$

waarbij $v(t)$ de logaritme van het werkelijke aantal reizigerskilometers in het kwartaal t voorstelt, $d(t)$ de logaritme van het verwachte aantal doden in het kwartaal t en $z(t)$ de logaritme van het verwachte aantal ziekenhuisgewonden. De grootheden $rd(t)$ en $rz(t)$ stellen de logaritme van het (theoretische)risico voor de doden respectievelijk ziekenhuisgewonden voor. Behorend bij de grootheden $d(t)$, $z(t)$ en $v(t)$ bestaan in principe observaties $D(t)$, $Z(t)$ en $V(t)$ uit respectievelijk AVV/BG-, LMR- en OVG-gegevens, zij het dat bij sommige leeftijdscombinaties er een trendbreuk in de reizigerskilometers optreedt bij de overgang van de OVG-steekproef van 1993 naar 1994, waarin de jongste leeftijdsklassen aan de steekproef zijn toegevoegd. Dit levert:

$$\begin{aligned}\log(V(t)) &\sim v(t) + c(t) \\ \log(D(t)) &\sim d(t) \\ \log(Z(t)) &\sim z(t)\end{aligned}$$

waarbij $c(t)$ de trend van de jongsten in het OVG aangeeft. Door de log-transformatie worden de componenten $v(t)$, $rd(t)$ en $rz(t)$ verondersteld een in

de tijd-variërende lineaire trend te volgen plus een in de tijd variërend seizoenseffect. Vanaf 1994 moet $c(t)$ gelijk zijn aan nul; de tijdstippen daarvoor heeft het een te schatten waarde, die eventueel ook tijd-variërend kan zijn (maar in de praktijk moeilijk van de trend te onderscheiden zal blijken). De ontwikkeling van de verhouding van het aantal 12-14-jarigen tot het totaal van 0-14-jarigen geven aanleiding tot het toevoegen van een dergelijke tijdsafhankelijke component. De componenten $v(t)$, $rd(t)$ en $rz(t)$ zijn dus in de tijd (onderling) gerelateerd. Hierin verschilt de huidige opzet met 'gewone' log-lineaire modellen.

Uiteindelijk zijn voor iedere combinatie maximaal drie modellen - verschillend in de specificatie van $c(t)$ - geschat: in ieder geval een model zonder correctie voor het afwezig zijn van reizigerskilometers van de kinderen onder 12 in de periode 1985-1993. In de meeste combinaties is dat ook niet nodig. In de andere gevallen - de modellen voor alle leeftijden en voor de leeftijdsklasse 0-14 jaar - is zowel een model met een tijd-variërende correctie als een model zonder tijd-variërende correctie daaraan toegevoegd. Dit levert bij elkaar 7 (6 plus totaal) wijzen van vervoer en 6 leeftijdsklassen (5 plus totaal) waarvoor het ongecorrigeerde model is gefit, plus 14 modellen waarvoor beide uitgebreide modellen worden gefit. Totaal levert dit $42+28=70$ modellen op met ieder iets minder dan 20 (totaal meer dan 1000) afbeeldingen. Om deze reden wordt in de volgende paragrafen alleen de relatief interessante combinaties beschreven.

4.1.1. *Geval 1: alle leeftijden en alle wijzen van vervoer*

Het blijkt dat de fit van het model met tijd-variërende correctie nauwelijks verschilt van de fit van het model zonder enige correctie. Dit was te verwachten gezien de betrekkelijk kleine bijdrage van de allerkleinsten in de totale mobiliteit. Deze analyse kan dus voor het gemak beperkt blijven tot één (het eenvoudigste) model.

De *Afbeeldingen 6a* tot *6c* laten de prognoses van het model voor de kwartalen van de jaren 1996 en 1997 zien. De reizigerskilometers (*Afbeelding 6a*) worden systematisch overschat door het systeem. Zie ook onderstaande prognoses in *Tabel 1*.

Het lijkt dat de trend misschien niet exponentieel toeneemt in weerwil van de aannames. De aanname van het exponentiële model zou verkeerd kunnen zijn. Iets dergelijks geldt ook voor de ontwikkeling van het aantal doden, maar niet voor de ziekenhuisgewonden (*Afbeeldingen 6b* en *6c*). Het is in principe mogelijk dat de overschatting van de doden het gevolg is van de overschatting van de reizigerskilometers.

De *Afbeeldingen 7a* (doden) en *7b* (ziekenhuisgewonden) geven het beeld gebruikmakend van geobserveerde reizigerskilometers in 1996 en 1997. De prognose in *Afbeelding 7a* ligt iets boven de geobserveerde aantallen doden. Dit betekent dat het aantal overleden slachtoffers (volgens dit model) de laatste twee jaren iets minder is geweest dan het aantal dat volgens het model het meest aannemelijk zou zijn geweest. Het is echter niet zo dat (op dit moment) een trendbreuk waargenomen wordt, want de geobserveerde aantallen liggen binnen de marges van de verwachte aantallen.

Kwartaal	Geobserveerd	Prognose	95%	
			ondergrens	bovengrens
Reizigerskilometers (miljard kms)				
1996/1	43,75	44,66	42,62	46,81
1996/2	46,59	49,70	46,83	52,75
1996/3	45,09	46,70	43,55	50,08
1996/4	47,03	48,07	44,49	51,93
1997/1	45,41	46,30	42,31	50,67
1997/2	48,94	51,52	46,69	56,86
1997/3	47,18	48,42	43,53	53,84
1997/4	47,81	49,83	44,54	55,74
Doden (aantal)				
1996/1	242	295,74	254,47	343,70
1996/2	325	340,96	292,54	397,38
1996/3	308	325,15	277,43	381,07
1996/4	305	335,78	285,90	394,36
1997/1	238	293,44	238,91	360,41
1997/2	298	338,31	274,80	416,49
1997/3	338	322,62	260,78	399,11
1997/4	289	333,17	268,78	412,97
Ziekenhuisgewonden (aantal)				
1996/1	3.781	4.151,94	3.732,83	4.618,11
1996/2	5.505	5.398,55	4.830,04	6.033,97
1996/3	5.599	5.730,43	5.087,88	6.454,12
1996/4	4.770	4.851,97	4.288,24	5.489,79
1997/1	4.178	4.150,86	3.525,18	4.887,58
1997/2	5.689	5.397,14	4.565,95	6.379,65
1997/3	5.947	5.728,94	4.816,86	6.813,72
1997/4	4.296	4.850,71	4.062,81	5.791,40

Tabel 1. *Prognoses en marges alle leeftijden en alle wijzen van vervoer.*

De *Afbeeldingen 8a* (reizigerskilometers), *8b* (doden) en *8c* (ziekenhuisgewonden) laten een daling zien in het verwachte aantal doden en niet in het verwachte aantal gewonden. De stijging van de mobiliteit zet nog steeds door. De trendbreuk tussen 1993 en 1994 in de observaties van de mobiliteit is wel zichtbaar, dit kan ten gevolge van het toevoegen van de kinderen van 0-11 jaar aan de OVG-steekproef zijn, maar dit kan ook toeval zijn geweest. De seizoensgecorrigeerde ontwikkeling van de reizigerskilometers in *Afbeelding 9a* laat de toename in het begin van 1994 wel zien, maar er blijken zich meer van dit soort toenames voor te doen. Het is mogelijk dat er een meerjarig periodiek effect bestaat (een zogenaamde cycle). De periodiciteit is echter niet heel systematisch - zowel de golflengte als het maximum variëren (zie aanwezigheid van een duidelijk maximum ongeveer in 1992) - zodat een cycle op deze data ook niet duidelijk geschat kan worden, en deze variatie voorlopig aan het toeval is toegeschreven. Dit leidt dan ook tot de betrekkelijk

grote marges in de prognoses van de mobiliteit (zie ook *Afbeelding 8a*). Wellicht is het model niet voldoende gespecificeerd voor het voorspellen van de mobiliteit.

Afbeelding 9b laat het seizoenseffect van de reizigerskilometers (uit *Afbeelding 8a*) zien. De structuur van deze ontwikkeling laat eigenlijk alleen een afnemende amplitude van het effect zien. De verhouding tussen de seizoenen lijkt zich nauwelijks te veranderen over deze periode. Dit gegeven, het langzaam afnemen van de amplitude tezamen met de toename in de trend en het multiplicatieve model suggereren een additief model. Als het seizoenseffect additief zou zijn dan zou dat betekenen dat de hoeveelheid reizigerskilometers, die minder in de winter wordt verreisd dan in de zomer, over de jaren ongeveer gelijk blijft ondanks de substantiële toename in de totale hoeveelheid kilometers. Dit lijkt toch wat onwaarschijnlijk. Bovendien lijkt de afname in de amplitude in *Afbeelding 9b* zich vooral aan het begin van de ontwikkeling te concentreren, zo tot 1991.

Afbeelding 9c laat de (eenvoudige) trendontwikkeling van het aantal doden zien. Aan de ontwikkeling van het absolute aantal (zichtbaar in de *Afbeeldingen 7a* en *8b*) valt ook een zekere regelmaat waar te nemen. Daarnaast bestaat een sterk seizoenseffect. De grootte van het seizoenseffect blijkt daar meer te variëren dan voor het verkeer. Er bestaat dus een sterkere relatie tussen de afwijkingen van vier observaties uit elkaar dan die één observatie van elkaar verwijderd zijn, vandaar dat de trend (*Afbeelding 9c*) zeer regelmatig is en het seizoenseffect niet (*Afbeelding 9d*). Dit blijkt uit vergelijking met het seizoenseffect van de reizigerskilometers uit *Afbeelding 9b*.

Afbeelding 9e laat de trendontwikkeling van het aantal ziekenhuisgewonden zien. De daling rond 1990 valt op. Ook lijkt er zich een nieuwe daling voor te doen juist aan het einde van 1997, maar of dat reëel is, is nu nog niet vast te stellen. *Afbeelding 9f* laat het seizoenseffect van de ziekenhuisgewonden zien. Ook hier is variatie aanwezig, maar minder dan bij de doden. Reden van de grotere onzekerheid bij de doden is ook de geringere omvang van de groep.

Samenvattend kan gesteld worden dat:

- correctie voor de kinderen van 0 tot 11 jaar niet nodig is gebleken voor het totaal van alle leeftijden samen;
- er zich een geleidelijke daling aan het voordoen is in het risico voor de doden en dat er zich geen markante variaties hierin voordoen;
- het risico voor de (LMR)-ziekenhuisgewonden zich minder gelijkmatig ontwikkelt in vergelijking met het risico voor de doden. Wellicht is dit het gevolg van het feit dat de aantallen (LMR)-ziekenhuisgewonden dusdanig zijn dat variaties wel worden onderscheiden en bij de doden (nog) niet. Dit verschijnsel doet zich overigens ook voor als AVV/BG-ziekenhuisgewonden worden geanalyseerd in plaats van LMR-ziekenhuisgewonden;
- het seizoenseffect van de slachtoffers sterker varieert dan het seizoenseffect van de reizigerskilometers. Het ligt voor de hand dat de seizoensinvloed op het risico wellicht meer met het weer en de invloed daarvan op de 'modal split' te maken heeft, dan met seizoenswijzigingen in het totale mobiliteitsvolume).

4.1.2. Geval 2: 0- tot 14-jarigen in de personenauto

Tabel 2 laat de prognoses van het aantal reizigerskilometers, doden en (LMR)ziekenhuisgewonden voor de groep van 0- tot 14-jarigen in de personenauto zien, met bijbehorende betrouwbaarheidsmarges. Het relatief grote aantal afwijkingen is gedeeltelijk het gevolg van misspecificatie van het model en van het feit dat de 95%-betrouwbaarheidsgrenzen van het gemiddelde worden gebruikt. De in dit voorbeeld relatief grote onzekerheid in de observaties zelf is niet in de marges inbegrepen.

In dit geval: '0- tot 14-jarigen in de personenauto' blijkt het corrigeren voor de 0- tot 11-jarigen in het OVG wel noodzakelijk. Zoals uit *Afbeelding 14a* volgt, is het corrigeren voor de toevoeging van de 0- tot 11-jarigen noodzaak. Waarschijnlijk heeft het model, waarin dus geen correctiecomponent was opgenomen, de trend goed opgepakt in dit geval, maar het model houdt bij zijn prognoses rekening met een grote onzekerheid. Het model heeft de toename in de steekproef aan het toeval toegeschreven, dat zou nog eens kunnen gebeuren.

In tegenstelling tot de algemene trend blijkt het model de reizigerskilometers voor de 0- tot 14-jarigen niet te overschatten in de jaren 1996 en 1997 (*Afbeelding 10a*). Net als bij het totaal laten de doden een dalende tendens zien, zij het dat deze later is ingezet, zie ook *Afbeelding 13b*. De invloed van de onzekerheid in de prognoses van de reizigerskilometers op de onzekerheid bij de doden is, om begrijpelijke redenen, betrekkelijk groot. Dit is duidelijk te zien aan het verschil tussen *Afbeelding 10b* (prognoses doden) en *11a* (prognoses doden, gegeven de reizigerskilometers). De marges in de laatste afbeelding zijn aanmerkelijk smaller dan in de eerste. Het aantal ziekenhuisgewonden laat eveneens een dalende tendens zien, zowel in absolute zin (*Afbeeldingen 10c, 11b en 12c*) als in relatieve zin via de trendontwikkeling in *Afbeelding 13c*.

Afbeelding 13d laat de correctiefactor voor de afwezigheid van 0- tot 11-jarigen zien. Het getal E tot de macht van dit getal (dus $\text{Exp}(x)$) is de factor waarmee de 'werkelijke' reizigerskilometers moeten worden vermenigvuldigd om tot de geobserveerde kilometers te komen. Uiteraard is het aantal reizigerskilometers met 0- tot 11-jarigen groter dan het aantal reizigerskilometers zonder 0- tot 11-jarigen, zodat deze factor dus kleiner dan één zou moeten zijn. E tot een negatieve macht levert een getal kleiner dan 1 op, hetgeen wil zeggen dat het aandeel van de geobserveerde reizigerskilometers in het werkelijke totaal aanvankelijk sterk daalt, daarna wat langzamer daalt tot ongeveer 1989-1990, en daarna heel licht stijgt tot het jaar 1994, waarin geen correctie meer nodig is.

Kwartaal	Geobserveerd	Prognose	95%	
			ondergrens	bovengrens
Reizigerskilometers (miljard kms)				
1996/1	3,15	2,77	2,51	3,04
1996/2	3,33	3,50	3,17	3,86
1996/3	3,38	3,27	2,97	3,61
1996/4	3,22	3,17	2,86	3,51
1997/1	3,21	2,75	2,39	3,16
1997/2	3,66	3,48	3,02	4,01
1997/3	3,90	3,25	2,82	3,75
1997/4	3,29	3,15	2,72	3,65
Doden (aantal)				
1996/1	3	3,00	1,60	5,63
1996/2	6	3,01	1,51	5,98
1996/3	7	2,70	1,26	5,80
1996/4	0	2,47	1,08	5,64
1997/1	3	2,31	0,90	5,94
1997/2	4	2,31	0,83	6,46
1997/3	5	2,07	0,67	6,40
1997/4	1	1,90	0,57	6,35
Ziekenhuisgewonden (aantal)				
1996/1	55	47,82	40,46	56,52
1996/2	47	58,68	50,04	68,82
1996/3	50	66,38	56,83	77,53
1996/4	67	54,34	46,28	63,80
1997/1	50	46,31	38,36	55,91
1997/2	64	56,83	47,43	68,10
1997/3	63	64,28	53,85	76,73
1997/4	50	52,62	43,88	63,11

Tabel 2. *Prognoses reizigerskilometers, doden en (LMR)ziekenhuisgewonden voor de groep 0- tot 14-jarigen in de personenauto met bijbehorende betrouwbaarheidsmarges.*

In Tabel 3 is een tabel met bevolkingscijfers voor de groep 0- tot 14-jarigen opgenomen. De percentages van de groep 12- tot 14-jarigen van het geheel van 0- tot 14-jarigen voor de jaren 1985 tot en met 1993 zijn: 24,44%, 23,42%, 22,29%, 21,17%, 20,52%, 19,99%, 19,74%, 19,52% en 19,62%. Deze ontwikkeling komt redelijk overeen met de gevonden correctie in *Afbeelding 13d*. Dit betekent dat de bevolkingsomvang van 0- tot 14-jarigen de relatieve afname voor een groot deel kan verklaren.

leeftijd:	0-4	5-9	10-11	12-14	Totaal t/m 14
Jaar:	Aantal Rij %				
1984	876 29.90%	904 30.85%	419 14.31%	731 24.94%	2930
1985	874 30.67%	889 31.18%	391 13.71%	697 24.44%	2850
1986	873 31.30%	890 31.92%	372 13.36%	653 23.42%	2788
1987	880 32.07%	892 32.47%	362 13.17%	612 22.29%	2746
1988	896 32.96%	890 32.75%	357 13.11%	575 21.17%	2719
1989	913 33.69%	885 32.67%	356 13.12%	556 20.52%	2710
1990	927 34.15%	887 32.69%	358 13.18%	543 19.99%	2715
1991	947 34.60%	887 32.40%	363 13.27%	540 19.74%	2738
1992	962 34.79%	896 32.42%	367 13.27%	540 19.52%	2765
1993	972 34.83%	913 32.70%	359 12.85%	548 19.62%	2791
1994	982 34.87%	931 33.05%	352 12.51%	551 19.57%	2816
1995	989 34.83%	945 33.30%	356 12.54%	548 19.32%	2838
1996	981 34.44%	964 33.84%	365 12.80%	538 18.91%	2848
1997	973 34.00%	977 34.15%	375 13.10%	536 18.75%	2861

Tabel 3. *Bevolking (x 1000) en rijpercentages. Bron: BIS-V: CBS - Statistiek van de bevolking*

Indien niet gecorrigeerd wordt voor de bijdrage in de mobiliteit van de 0- tot 11-jarigen is de prognose uiterst onzeker (zie *Afbeeldingen 14a* en, in bepaalde mate *14b*). De trend volgt de data getrouw omdat weinig zekerheid wordt verondersteld van de prognoses en de update dus vooral op de observatie zal zijn gebaseerd. Aardig is ook te zien dan de onzekerheid van het model zo groot is dat binnen twee jaren de toename in 1994 weer ongedaan gemaakt kan zijn.

Tenslotte laat *Afbeelding 14c* zien dat, als geen tijd-variërende correctie-component wordt toegepast, het model een dalende mobiliteit in het begin van de observatie periode verondersteld. Dit is begrijpelijk daar de mobiliteitscijfers aan het begin van de periode licht hoger lagen dan daarna.

De belangrijkste conclusies zijn:

- Er moet voor de verandering in het OVG gecorrigeerd worden,
- de correctie komt redelijk overeen met de ontwikkeling in de relatieve aantallen jongeren 0 tot 11 jaar en 12 tot 14 jaar.

4.1.3. Geval 3: 25- tot 49-jarige personenautoinzittenden

Tabel 4 toont de prognoses en betrouwbaarheidsmarges voor 25- tot 49-jarige personenautoinzittenden. Voor deze combinatie hoeft geen correctie te worden toegepast. Ook in dit geval levert de prognose van de reizigerskilometers in 1996 en 1997 een lichte overschatting op (zie *Afbeelding 15a*). Bovendien blijkt het de amplitude van het seizoenseffect (niet afgebeeld) nauwelijks kleiner te worden in de geobserveerde periode, sterker nog, het seizoenseffect is nauwelijks tijd variërend, geeft het zelfde patroon over de gehele periode. Opvallend is de onzekerheid in de prognoses voor de ziekenhuisgewonden (zie *Afbeeldingen 15c, 16b en 17c*). De verklaring hiervoor moet worden gezocht in enige ‘bobbels’ in de trend voor de ziekenhuisgewonden en een opvallende daling tot 1989 zie *Afbeelding 18c*. Het model heeft deze (nog niet verklaarde effecten) aan het toeval toegeschreven, en heeft aldus een vrij onzekere prognose gegenereerd. Het verklaren van de effecten zal de marges van de prognoses en daarmee de mogelijkheid om met behulp van dit instrument veranderingen in de ontwikkeling aan te tonen sterk verbeteren.

Samenvattend:

- Opvallend is de onzekerheid in de prognoses voor de ziekenhuisgewonden.
- Er is een lichte overschatting van de reizigerskilometers in 1996 en 1997.
- De amplitude van het seizoenseffect blijkt nauwelijks kleiner te worden in de geobserveerde periode. Dit is in tegenstelling tot vorige voorbeelden.

Kwartaal	Geobserveerd	Prognose	95%	
			ondergrens	bovengrens
Reizigerskilometers (miljard kms)				
1996/1	18,56	18,00	17,39	18,64
1996/2	18,82	20,07	19,27	20,90
1996/3	17,63	18,34	17,50	19,23
1996/4	19,30	19,45	18,46	20,50
1997/1	18,12	18,53	17,30	19,85
1997/2	19,12	20,65	19,10	22,34
1997/3	17,94	18,88	17,29	20,62
1997/4	19,25	20,02	18,16	22,06
Dodens (aantal)				
1996/1	63	68,91	60,44	78,58
1996/2	59	66,03	57,27	76,13
1996/3	63	62,21	53,28	72,63
1996/4	69	77,83	66,16	91,56
1997/1	48	70,40	58,51	84,70
1997/2	69	67,45	55,16	82,48
1997/3	58	63,55	51,10	79,03
1997/4	64	79,51	63,08	100,20
Ziekenhuisgewonden (aantal)				
1996/1	550	563,93	511,07	622,27
1996/2	506	535,83	473,15	606,80
1996/3	473	552,34	476,18	640,70
1996/4	621	635,47	535,50	754,11
1997/1	523	578,23	473,50	706,11
1997/2	576	549,41	438,84	687,83
1997/3	513	566,35	441,39	726,68
1997/4	567	651,58	495,64	856,57

Tabel 4. *Prognoses en marges 25- tot 49-jarigen in de personenauto.*

4.1.4. *Geval 4: 15- tot 24-jarige bromfietzers*

De categorie 15- tot 24-jarige bromfietzers domineert de categorie bromfietzers van alle leeftijden. Ook hier is geen correctie voor de categorie kinderen noodzakelijk. De *Afbeeldingen 19a* en *22a* laten een duidelijk dalende ontwikkeling in de reizigerskilometers zien. De eerste drie jaar hebben een duidelijk ander seizoenseffect dan de rest van de jaren, die verder geen opvallende verandering van het seizoenseffect laten zien (*Afbeelding 22b*). Het afwijkende seizoenseffect in het begin van de periode is ook terug te vinden in de slachtofferreeksen, zodat het aannemelijk is dat deze afwijking 'echt' is en zich werkelijk in de verreden kilometers heeft voorgedaan. Dit in tegenstelling tot de toch redelijk frequente - soms forse - afwijkingen die zich elders in de reeks voordoen. Het model heeft 'geconstateerd' dat deze afwijkingen zich niet (of niet in die richting) in de slachtofferaantallen

voordoen, zodat het model deze cijfers min of meer corrigeert. Een van de aantrekkelijke eigenschappen van het instrument is nu dat het verder met de gecorrigeerde gegevens de risico's berekent.

Op het meer analytische vlak bekeken blijkt de mobiliteit een (exponentiële) daling ingezet te hebben, daar de ontwikkeling van de trend lineair dalend is in de 'logaritmische ruimte' (*Afbeelding 22a*). De steilheid van deze daling varieert nauwelijks: de drift-component is vrijwel constant. Opvallend is verder dat er geen sprake van een daling is bij de slachtofferrisico's. (Zie *Afbeeldingen 22c* en *22d*). De aantallen zijn in absolute zin wel gedaald, echter per verreden kilometer is het aantal aan het eind van de jaren tachtig juist toegenomen. Het risico voor de ziekenhuisgewonden lijkt aan het begin van de jaren negentig weer omlaag te gaan, maar dat is slechts van korte duur. Bij de verkeersdoden is deze tendens niet opvallend waarneembaar, waarschijnlijk door de kleine aantallen (de gegevens over alle leeftijden laten een zelfde tendens zien). Een complicatie bij deze fluctuatie in het risico van de ziekenhuisgewonden kan wel zijn dat bij een eventuele invoering van de bromfiets op de rijbaan, een daling in de toekomst van het aantal ziekenhuisgewonden relatief moeilijk aan de maatregel toe te schrijven zal zijn. Het zal noodzakelijk blijken daar specifieke indicatoren voor te definiëren.

Samenvattend:

- Er is geen sprake van een daling van de slachtofferrisico's.
- Er is wel een dalende ontwikkeling in het aantal reizigerskilometers.
- De jaren 1985 tot 1988 hebben een ander seizoeneffect dan de rest van de jaren. Dit zou samen kunnen hangen met het snorfietsgebruik.
- Het model heeft 'geconstateerd' dat afwijkingen in de reizigerskilometers zich niet in de slachtofferaantallen voordoen en het model heeft blijkbaar deze cijfers min of meer gecorrigeerd.

4.2. Registratie AVV/BG- versus LMR-ziekenhuisgewonden

In het nu volgende voorbeeld zijn de AVV/BG-ziekenhuisgewonden aan de te analyseren data toegevoegd. Zo zijn voor dezelfde combinaties van leeftijd en wijze van vervoer zowel reizigerskilometers als AVV/BG-doden, AVV/BG-ziekenhuisgewonden én LMR-ziekenhuisgewonden gemodelleerd.

De basisgedachte van het model is gelijk aan dat in de vorige voorbeelden, waarbij in het begin is uitgegaan van de aanname dat de aantallen AVV/BG-ziekenhuisgewonden een (tijd-variërend) percentage zijn van de aantallen LMR-ziekenhuisgewonden. Dit zal tot een extra onzekerheid leiden in de prognoses - wat niet de directe doelstelling van de analyse is - maar het levert wel inzicht op over de (ontwikkeling van de) verhouding tussen slachtofferaantallen uit beide bronnen, waaruit conclusies over de registratiegraad zouden kunnen worden getrokken.

Het blijkt al snel - zeker voor het totaal - dat er sprake kan zijn van een seizoensgebonden verhouding, waarschijnlijk ten gevolge van de seizoensfluctuaties in (brom)fiets gebruik. Om die reden is verder gewerkt met een eigen seizoeneffect voor de LMR- en AVV/BG-aantallen, die een gemeenschappelijke trend behouden. De verhouding tussen AVV/BG- en LMR-aantallen is in het model aangebracht door een losse trendcomponent,

uitgevoerd met behulp van een 'random-walk'. Dit houdt een niet-systematische daling (of stijging) van de registratiegraad in. Wel houdt dit in dat als de registratiegraad in een bepaald kwartaal lager uitvalt dan in het vorige kwartaal, het waarschijnlijk is dat de registratiegraad het volgende kwartaal ook op het lagere niveau ligt.

Bij het model voor alle leeftijden en alle wijzen van vervoer komen de resultaten uiteraard sterk overeen met Geval 1. *Afbeelding 23b* laat een minder goede voorspelling zien dan *Afbeelding 6c*. *Afbeelding 23c* laat een plotselinge (sterkere) daling van de verhouding tussen AVV/BG-gewonden en LMR-gewonden zien, die aan het einde van 1989 of het begin van 1990 inzet. Deze daling lijkt te stagneren vanaf eind 1992 met een plotselinge toename in 1996. Deze toename is waarschijnlijk niet het gevolg van een plotselinge afname in LMR-cijfers. Met enige fantasie kan men veronderstellen dat de systematische daling die zich aan het begin van de observatieperiode lijkt in te zetten zich na een dip voortzet vanaf 1996/1997. (Het is met dit gegevensmateriaal nooit uit te maken of het niet al in gang was voor die tijd.) Het toevoegen van een drift verbetert echter weinig aan het model. Niet ondenkbaar is een relatie tussen deze afname in de AVV/BG-registratie (op basis van politiegegevens) in verhouding tot LMR-registratie, en de reorganisatie van de (verkeers)politie in ongeveer dezelfde periode.

Afbeelding 23g laat voor alle verkeer de seizoenseffecten zien van LMR- en AVV/BG-ziekenhuisgewonden. Het lijkt hier dat de effecten in wezen dezelfde vorm hebben, maar dat de LMR-gewonden een grotere variatie vertonen. Doordat de effecten moeten optellen tot nul is het niet meteen duidelijk of de registratiegraad van de AVV/BG-gewonden - waarschijnlijk - minder wordt naarmate er meer gewonden vallen. Dit zou, als deze hypothese juist blijkt, tenminste voor een deel veroorzaakt kunnen worden door een overbelastingsprobleem bij de registratie van de ziekenhuisgewonden.

Dezelfde analyses als bovenstaande zijn ook uitgevoerd voor alleen bromfietzers, fietsers en voetgangers. De resultaten van de ontwikkelingen in de registratiegraad staan in de *Afbeeldingen 23d* (bromfietzers) *23e* (fietsers) en *23f* (voetgangers). De seizoenseffecten zijn afgebeeld in de *Afbeeldingen 23h* (bromfietzers), *23i* (fietsers) en *23j* (voetgangers). Opvallend is dat het patroon voor bromfietsen ongeveer hetzelfde is als voor alle vervoerwijzen samen. De seizoenseffecten van AVV/BG- en LMR-gewonden verschillen zelfs nauwelijks van elkaar. Daarentegen laten de fietsers en voetgangers een heel ander beeld zien. Bij de fietsers start de daling in de verhouding tussen AVV/BG- en LMR-gewonden ongeveer gelijktijdig met de rest, echter het afnemen van de daling begint veel later, ongeveer begin 1996. Bij de voetgangers blijkt de daling zelfs eerder te beginnen en is er sprake van een geleidelijk herstel rond 1996. De seizoenseffecten van de voetgangers laten overeenkomsten zien tussen AVV/BG- en LMR-registratie, die van de fietsers veel minder. Voor fietsers laten de AVV/BG-gegevens (gecorrigeerd voor OVG-cijfers) een piek in het najaar zien, terwijl de LMR-gegevens een piek vertonen in de zomer. Dit feit geeft overigens aan dat er wellicht een andere verklaring dan alleen overbelasting een rol speelt. Hierbij valt te denken aan onderbezetting van het politiepersoneel in de zomermaanden. Uiteraard zijn dit soort conclusies aan de hand van een uitgewerkt voorbeeld alleen bedoeld om de werking van de methode te illustreren en moeten deze met enige voorzichtigheid worden benaderd.

Enkele conclusies met gewaagde interpretaties:

- Het blijkt dat er sprake kan zijn van een seizoensgebonden verhouding tussen LMR- en AVV/BG-gewonden. Het lijkt hier dat de seizoenseffecten in wezen dezelfde vorm hebben, maar dat het seizoenseffect bij de LMR-gewonden sterker is. Dit zou misschien voor een deel veroorzaakt kunnen worden door een overbelastingsprobleem bij de registratie van de ziekenhuisgewonden.
- Er bestaat een plotselinge (sterkere) daling in de verhouding tussen AVV/BG-gewonden en LMR-gewonden aan het einde van 1989 of het begin van 1990. Deze daling lijkt te stagneren vanaf eind 1992 en de verhouding neemt weer plotseling toe in 1996. Deze toename is waarschijnlijk niet het gevolg van een plotselinge afname in LMR-cijfers. Mogelijk is er een relatie tussen deze afname in de AVV/BG-registratie (op basis van politiegegevens) ten opzichte van de LMR-registratie, en de reorganisatie van de (verkeers)politie in ongeveer dezelfde periode.
- Het valt op dat het patroon van de ontwikkeling van de relatieve registratiegraad voor bromfietsen ongeveer dezelfde is als voor alle vervoerwijzen samen.

4.3. Invloed mobiliteit snelverkeer op aantallen slachtoffers langzaam verkeer

In de bovenstaande voorbeelden is telkens aangenomen dat de aantallen slachtoffers in een bepaalde combinatie van leeftijdsklasse en wijze van vervoer in essentie proportioneel zijn met de hoeveelheid reizigerskilometers. In het nu volgende voorbeeld is hieraan de totale hoeveelheid reizigerskilometers in personenauto's (alle leeftijden) als benadering voor de hoeveelheid voertuigkilometers snelverkeer (personenautoverkeer) toegevoegd. De analyses zijn beperkt tot motorfietsers, bromfietsers, fietsers en voetgangers, eventueel per leeftijdsklasse. De mate van invloed van het snelverkeer wordt hierbij geschat terwijl de invloed van de eigen mobiliteit verondersteld wordt lineair (met coëfficiënt 1) te zijn.

In het nu gebruikte model zijn geen AVV/BG-gewonden geanalyseerd.

Als eerste voorbeeld is het equivalent van 'Geval 4: 15- tot 24-jarige bromfietsers' uitgewerkt. In de *Afbeeldingen 24a* (eigen reizigerskilometers), *24b* (doden) en *24c* (ziekenhuisgewonden) zijn de resultaten voor deze analyse opgetekend. In *Afbeelding 24b* valt het op dat de marges van de prognoses voor de laatste observaties kleiner zijn dan die in *Afbeelding 19b*. Voor het gemak zijn de marges en de prognoses voor beide modellen en voor doden en ziekenhuisgewonden in de volgende tabellen (*Tabellen 5a t/m 5d*) weergegeven. Daarnaast is de chi-kwadraatwaarde gegeven. De chi-kwadraatwaarde is gelijk aan de som van de geobserveerde minus de geschatte waarde in het kwadraat gedeeld door de geschatte waarde. Dit is niet gelijk aan de waarde die gebruikt bij de schattingsprocedure, waar gewogen is naar de geobserveerde aantallen.

Kwartaal	Geobserveerd	Prognose	95 %		Chi ²
			Ondergrens	Bovengrens	
1996/1	10	12,32	8,48	17,91	0,44
1996/2	18	14,95	10,08	22,17	0,62
1996/3	18	16,08	10,57	24,46	0,23
1996/4	11	15,51	9,94	24,19	1,31
1997/1	7	12,07	7,19	20,27	2,13
1997/2	15	14,65	8,57	25,04	0,01
1997/3	12	15,76	9,02	27,54	0,90
1997/4	12	15,19	8,50	27,15	0,67
				Som	6,30

Tabel 5a. *Prognoses voor het aantal doden in de laatste acht kwartalen (twee jaar, zonder personenautomobiliteit). Zie ook 'Geval 4: 15- tot 24-jarige bromfietsers'.*

Kwartaal	Geobserveerd	Prognose	95 %		Chi ²
			Ondergrens	Bovengrens	
1996/1	10	11,66	8,58	15,85	0,24
1996/2	18	14,28	10,50	19,44	0,97
1996/3	18	15,17	10,99	20,92	0,53
1996/4	11	14,79	10,60	20,65	0,97
1997/1	7	11,30	7,59	16,83	1,63
1997/2	15	13,84	9,27	20,66	0,10
1997/3	12	14,69	9,72	22,21	0,49
1997/4	12	14,33	9,38	21,90	0,38
				Som	5,31

Tabel 5b. *Prognoses voor het aantal doden in de laatste acht kwartalen (twee jaar, met personenautomobiliteit).*

Kwartaal	Geobserveerd	Prognose	95 %		Chi ²
			Ondergrens	Bovengrens	
1996/1	253	312,69	246,67	396,39	11,40
1996/2	403	406,94	310,91	532,62	0,04
1996/3	500	468,40	346,53	633,14	2,13
1996/4	384	420,83	304,70	581,21	3,22
1997/1	314	305,11	204,87	454,39	0,26
1997/2	410	397,07	260,44	605,36	0,42
1997/3	445	457,04	292,39	714,41	0,32
1997/4	398	410,62	258,08	653,32	0,39
				Som	18,17

Tabel 5c. *Prognoses voor het aantal ziekenhuisgewonden in de laatste acht kwartalen (twee jaar, zonder personenautomobiliteit).*

Kwartaal	Geobserveerd	Prognose	95 %		Chi ²
			Ondergrens	Bovengrens	
1996/1	253	306,50	246,44	381,20	9,34
1996/2	403	404,48	318,38	513,88	0,01
1996/3	500	460,81	354,18	599,54	3,33
1996/4	384	416,97	315,25	551,50	2,61
1997/1	314	298,18	209,98	423,41	0,84
1997/2	410	393,50	272,65	567,91	0,69
1997/3	445	448,30	304,83	659,29	0,02
1997/4	398	405,64	272,08	604,77	0,14
				Som	16,98

Tabel 5d. *Prognoses voor het aantal ziekenhuisgewonden in de laatste acht kwartalen (twee jaar, met personenautomobiliteit).*

In de *Tabellen 5a t/m 5d* is duidelijk zichtbaar dat de voorspellingen voor het aantal verkeersdoden meer verbeterd worden dan die voor de ziekenhuisgewonden door de toevoeging van de reizigerskilometers personenauto.

Over het algemeen leidt deze toevoeging tot een (niet altijd even grote) verbetering.

Tabellen 6a en *6b* tonen de chi-kwadraatwaarden voor de prognoses voor verkeersdoden en ziekenhuisgewonden. De chi-kwadraatwaarden van de prognoses met en zonder rekening te houden met het aantal reizigerskilometers in personenauto's worden in deze tabellen vergeleken.

Uit *Tabellen 6a* en *6b* zou geconcludeerd kunnen worden dat het voordeel van het rekening houden met de mobiliteit van de personenauto's het grootst is bij het voorspellen van het aantal doden. In de meeste gevallen (16 uit 24, alle voor alle leeftijden) levert dit een verbetering op. Bij de fiets is dit niet altijd het geval.

Tabel 7 toont de geschatte coëfficiënten die de invloed weergeven van de personenautomobiliteit op de slachtofferaantallen bij het langzaam verkeer. Deze coëfficiënten laten een beeld zien waaruit de gedachte kan ontstaan dat de coëfficiënten in principe per wijze van vervoer zijn gedefinieerd, en waarbij (wellicht naar kwetsbaarheid) de jongere en oudere verkeersdeelnemers meer beïnvloed worden dan de middengroepen.

	Motor		Bromfiets		Fiets		Lopen	
	zonder Personenauto kilometers	met reiziger kilometers	zonder Personenauto kilometers	met reiziger kilometers	zonder Personenauto kilometers	met reiziger kilometers	zonder Personenauto kilometers	met reiziger kilometers
Alle	8,13	9,95	33,41	10,61	10,16	7,95	12,17	10,51
0-14	0,00	0,01	14,03	6,96	4,05	9,09	22,69	9,04
15-24	11,99	11,72	6,31	5,31	8,96	9,04	38,00	30,90
25-49	10,20	13,21	4,57	5,81	14,41	14,79	7,42	7,39
50-64	16,36	15,95	28,31	11,66	6,22	6,31	17,21	18,02
65+	0,30	0,32	12,14	10,94	9,11	9,36	10,17	9,96

Tabel 6a. *Vergelijking tussen χ^2 -waarden zonder en met bijdrage mobiliteit (in reizigerskilometers) van personenauto's: Doden, leeftijdsklassen maal wijze van vervoer.*

	Motor		Bromfiets		Fiets		Lopen	
	zonder Personenauto kilometers	met reiziger kilometers	zonder Personenauto kilometers	met reiziger kilometers	zonder Personenauto kilometers	met reiziger kilometers	zonder Personenauto kilometers	met reiziger kilometers
Alle	28,62	28,77	37,20	37,95	97,74	105,20	28,76	21,65
0-14	4,25	4,29	7,10	7,57	35,13	30,17	71,67	14,05
15-24	14,55	13,72	18,17	16,98	18,86	17,39	8,91	8,20
25-49	42,21	38,22	27,68	27,19	23,73	22,99	11,38	14,35
50-64	39,85	17,80	15,43	16,87	26,77	26,84	16,36	18,79
65+	30,73	49,04	22,06	23,94	71,63	72,36	14,38	15,19

Tabel 6b. *Vergelijking tussen χ^2 -waarden zonder en met bijdrage mobiliteit (in reizigerskilometers) van personenauto's: Ziekenhuisgewonden, leeftijdsklassen maal wijze van vervoer.*

	Motor	Bromfiets	Fiets	Lopen
Alle	1,82	1,93	1,45	1,47
0-14	0,60	2,25	1,38	1,48
15-24	2,00	1,96	1,28	1,40
25-49	1,76	1,78	1,16	1,33
50-64	1,71	1,85	1,54	1,58
65+	1,17	2,02	1,88	1,86

Tabel 7. *Coëfficiënten voor de invloed van de personenautomobiliteit.*

Het model dat rekening houdt met de mobiliteit van het snelverkeer kan op een aantal belangrijke punten worden verbeterd: door het meenemen van echte motorvoertuigkilometers in plaats van personenauto-reizigerskilometers, en wellicht door te beperken tot dodelijke ongevallen of overleden slachtoffers.

Verder moet worden bedacht dat het om alle slachtoffers van langzaam verkeer gaat en niet alleen om slachtoffers van ongevallen tussen langzaam verkeer en motorvoertuigen of personenauto's. Voor deze groep zijn grotere effecten te verwachten.

Samenvatting:

- Door de toevoeging van de reizigerskilometers van personenauto-inzittenden worden de voorspellingen van de aantallen doden onder het langzaam verkeer sterker verbeterd dan die van de aantallen ziekenhuisgewonden. Dit geldt misschien niet voor de slachtoffers onder fietsers.
- De geschatte invloed van reizigerskilometers van personenauto-inzittenden op het aantal slachtoffers onder het langzaam verkeer, lijkt in principe per wijze van (langzaam) vervoer te verschillen. Hierbij lijkt de invloed groter te zijn op de slachtofferaantallen onder de jongere en oudere verkeersdeelnemers, dan de onder de middengroepen.
- Er moet worden bedacht dat het bij de hier gepresenteerde analyses om alle slachtoffers van langzaam verkeer gaat en niet alleen om slachtoffers van ongevallen met motorvoertuigen of personenauto's. Voor deze deelgroepen zijn grotere effecten te verwachten. Bovendien zijn de motorvoertuigmobiliteit in plaats van de reizigerskilometers personenauto gebruikt kunnen worden. Idealiter zou hier de mobiliteit op hogere-ordewegen vanaf gehaald moeten worden.

5. Conclusies

Het ontwikkelde instrument heeft een zeer grote mate van flexibiliteit, waardoor zeer gevarieerde problemen kunnen worden geanalyseerd. Het betreft de gelijktijdige analyse (in samenhang) van diverse ontwikkelingen en de relaties daarin, zoals bijvoorbeeld de ontwikkeling van de mobiliteit en van het aantal AVV/BG- en LMR-slachtoffers. Daarnaast kan eenvoudig rekening worden gehouden met gedeeltelijk ontbrekende gegevens.

Door de modulaire opbouw is het instrument eenvoudig aan te passen aan specifieke wensen. Dit betreft de aard van de invoergegevens, het aantal variabelen en de relaties die daartussen kunnen worden gedefinieerd. In een inputmatrix kunnen de beschikbare gegevens eerst worden gedefinieerd. In een soort ‘design-matrices’ worden vervolgens de modelspecificaties vastgelegd. Hiermee kan worden aangegeven hoe de relaties tussen componenten onderling en componenten en gemeten grootheden liggen. Tenslotte kan worden aangegeven welke analyseresultaten als uitvoer gewenst zijn. De uitvoer van het model wordt in principe gegeven in interpreteerbare grootheden, zoals een algemene trend, een seizoenseffect, een (economische) cyclus, één of meer mobiliteits- en risicocomponenten, eventuele externe factoren als interventies, enzovoort.

Het ontwikkelde instrument is op een groot aantal verkeersveiligheidsproblemen toegepast. Daarbij blijkt dat door de gestructureerde opbouw ervan, aspecten van de ontwikkeling van het verkeer en de verkeersveiligheid zichtbaar kunnen worden gemaakt die niet direct evident zijn. Zo kunnen bijvoorbeeld verschillen in seizoenseffecten en veranderingen daarin worden weergegeven bij verschillende typen ongevallen of slachtoffers. Ook zijn verhoudingen en ontwikkelingen daarvan gemodelleerd. Dit is zowel in het geval van een studie naar de registratiegraad uitgevoerd, als in het geval van de correctie voor het aandeel van de jongste kinderen (0- tot 11-jarigen) als middel om bepaalde, anders niet bruikbare, gegevens toch te kunnen gebruiken. Hierbij kan een correctie worden geschat welke redelijk blijkt overeen te komen met de ontwikkeling in de bevolkingssamenstelling.

Het model kan op natuurlijke wijze gebruikmaken van kennis van de onbetrouwbaarheid van gebruikte gegevens. Dit is bijvoorbeeld nuttig gebleken bij het gebruiken van de OVG-gegevens, met de daarin opgetreden vergroting van de betrouwbaarheid door uitbreiding van de steekproef. Het model blijkt in staat rekening te kunnen houden met een wijziging in de betrouwbaarheid van de gebruikte gegevens: het gaat niet per se uit van gelijkblijvende betrouwbaarheid van de cijfers. In deze zin is de hier gebruikte implementatie een verbetering op de implementatie van de London School of Economics (STAMP) waarin geen rekening kan worden gehouden met verschillende varianties voor verschillende tijdstippen. Dit is vrij gebruikelijk in de economie en heeft ook bepaalde voordelen. Wel kan in STAMP rekening worden gehouden met ontbrekende data en ook wordt een beter systeem voor de begintoestand gebruikt: het diffuus Kalman-filter.

Het gebruikmaken van kennis van de onbetrouwbaarheid van gebruikte gegevens kan in extremum gebruikt worden als bepaalde cellen (systematisch)

ontbreken. Zo kunnen bijvoorbeeld jaarcijfers gebruikt worden in een op kwartaalcijfers gebaseerde analyse. Of, in het algemeen, voor een langere periode geaggregeerde gegevens kunnen gebruikt worden in een analyse waarin kortere aggregatieperioden worden gebruikt. De onzekerheid van ontbrekende gegevens wordt in de huidige implementatie zo groot aangegeven dat de fysiek aangegeven waarde een verwaarloosbaar kleine invloed op de analyse heeft.

Het model leverde soms prognoses waarin het weinig vertrouwen leek te hebben, resulterend in ruime betrouwbaarheidsmarges. Dit is bijvoorbeeld het geval bij de 25- tot 49-jarige personenauto-inzittenden. Het vertrouwen in de voorspelling wordt daarbij blijkbaar enerzijds beperkt door de kwaliteit en onbetrouwbaarheid van de gegevens en anderzijds door beperkingen in de modelstructuur, omdat een goede verklaring (vooralsnog) ontbreekt. Sommige modellen en prognoses zouden aanzienlijk kunnen worden verbeterd als verklaringsfactoren worden toegevoegd aan het model, voor zich plotseling voordoende veranderingen in de ontwikkeling, die nu door het systeem nog als toevallige fluctuaties moeten worden gezien en waarmee het model dus ook in de toekomst rekening moet houden. Wat echter belangrijk is naast de precisie van de directe prognoses van het model, is dat de mate van betrouwbaarheid van de uiteindelijke prognose en de onzekerheid daarin bij het gedefinieerde model goed wordt weergegeven, zodat deze beperkingen van de prognoses goed zichtbaar worden.

De ontwikkeling van het instrument bevindt zich nu op een 'meersprong': er zijn een aantal nuttige richtingen denkbaar, waarin sommige van de technische verbeteringen (nog) niet zijn uitgewerkt en dus veel inspanning kosten om te implementeren. Bovendien blijken veel innovatieve ontwikkelingen eerst voor univariate gevallen te worden ontwikkeld, zodat deze modellen eerst aangepast moeten worden. (Ook dit model is in het verleden toegepast op een univariate reeks in Bos & Bijleveld (1991), een multivariate implementatie was toen niet beschikbaar.) Een belangrijke richting die kan worden ingeslagen is de richting van het diffuus Kalman-filter. Een andere richting, die zeker voor kleine aantallen slachtoffers relevant kan zijn is de aanname van de Poisson-verdeling voor slachtofferaantallen. Het Kalman-filter is onder die omstandigheden niet optimaal; de vraag is echter of een benadering van de Poisson-verdeling met behulp van herhaalde herweging of als vorm van een E(xpectation)M(aximization)-algoritme niet toepasbaar is, mede rekening houdend met het beperkte aantal observaties. Tenslotte kan men zich afvragen of niet afgestapt moet (kunnen) worden van het gebruikte (log)lineaire model. In het voorbeeld voor alle leeftijden en alle wijzen van vervoer werd de mogelijkheid geopperd van een additief seizoenseffect. Vergelijkbare aanpassingen zijn gesuggereerd bij modellen waarbij rekening wordt gehouden met de mobiliteit van botspartners.

Gezien de resultaten is het verstandig te onderzoeken hoe het nieuwe instrument ingezet kan worden bij een nieuwe opzet van de kwartaalanalyses. Ook zou het mogelijk zijn de techniek toe te passen op de macroscopische modellen zoals die in de lange-termijnprognoses zijn gebruikt.

Literatuur

Akaike, H. (1974). *Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes*. In: Annals of the Institute of Statistical Mathematics, Nr: 26, p. 363-387.

Bos, J.M.J. & Bijleveld, F.D. (1991). *Tijdreeksanalyse van het gordeleffect*. R-91-92. SWOV, Leidschendam.

CBS (1993). *De mobiliteit van de Nederlandse bevolking in 1992*. Centraal Bureau voor de Statistiek, Voorburg/Heerlen.

Harvey, A.C. (1989) *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge.

Harvey, A.C. & Shepard, N. (1993). *Structural time series models*. in: Handbook of Statistics Vol 11, Maddala, G.S, Rao, C.R. & Vinod, H.D. (Red.), p. 261-302. Elseviers Science.

Jong, P. de (1988). *The likelihood for a state space model*. In: Biometrika, Vol. 75, Nr. 1, p. 165-169.

Kendall, M. & Ord, J. K. (1990). *Time Series*. Third Edition. Edward Arnold, London.

Koopman, S. J & Shepard, N. (1992). *Exact score for time series models in state space form*. In: Biometrika, Vol. 79, Nr. 4, p. 823-826.

Koopman, S. J. (1997). *Exact initial Kalman filtering and smoothing for nonstationary time series models*. In: Journal of the American Statistical Association, Vol. 92, Nr. 440, p. 1630-1638.

Oppe, S., Koornstra, M.J. & Roszbach, R. (1988). *Macroscopic models for traffic and traffic safety*. In: Traffic safety theory & research methods; Proceedings International Symposium, April 26-28, 1988, Amsterdam, Session 5: Time dependent models. SWOV, Leidschendam.

Technische beschrijving

Het model is gebaseerd op een niet-geobserveerde toestand.

Beginnende bij een initiële toestand (waarover later meer) wordt verondersteld dat de nieuwe toestand rechtstreeks volgt uit de vorige toestand, plus eventueel een stochastische verstoring (random fout). In het eerste geval is de toestand als volgt gedefinieerd:

$$z(t) = z(t-1) \quad (1)$$

en voor de geobserveerde tijdreeks geldt:

$$y(t) = z(t) + e(t) \quad (2)$$

waarbij $e(t)$ een toevallige fout is. Duidelijk is dat de som van de kwadraten van de termen $e(t)$ wordt geminimaliseerd als $z(0)$ (en daarmee iedere $z(t)$) het gemiddelde is van de $y(t)$.

Als nu aangenomen wordt dat er zich een zekere verandering voordoet in de waarden van de $y(t)$, dan kan (1) worden uitgebreid met een toevallige fout $v(t)$:

$$z(t) = z(t-1) + v(t) \quad (3)$$

Een afwijking door middel van $v(t)$ op tijdstip t blijft in alle daar op volgende tijdstippen $t+1$, $t+2$, ... doorwerken terwijl een afwijking $e(t)$ dat niet doet. Dit betekent langduriger schommelingen in het niveau van de $y(t)$ in $v(t)$ worden opgenomen terwijl tijdelijke afwijkingen door $e(t)$ worden gecompenseerd. Op een dergelijke wijze kan een plotseling niveauverschil worden opgespoord. Als er echter sprake is van een trend (bijvoorbeeld als de $v(t)$ gemiddeld niet gelijk zijn aan nul, dan kan dat in het model worden toegevoegd door middel van een drift-component $d(t)$:

$$z(t) = z(t-1) + d(t) \quad (4a)$$

$$d(t) = d(t-1) \quad (4b)$$

Dit model, zonder verstoringstermen, komt neer op een gewone lineaire regressie $y(t) = a + b$ met $a = d(0)$ en $b = z(0)$. Uiteraard kunnen ook hier fouttermen worden toegevoegd om niveauwisselingen of veranderingen in de trend aan te geven in het model.

In zijn algemeenheid kan de toestand als een vector Z worden gezien (in het bovenstaande geval met componenten $z(t)$ en $d(t)$ in Z verwerkt) waarbij:

$$Z(t) = F Z(t-1) + V(t) \quad (5)$$

en

$$Y(t) = H Z(t) + E(t) \quad (6)$$

In het bovenstaande model is F

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

en $Y(t)$ een één-dimensionale vector $y(t)$, tenslotte is H gelijk aan (1 0).

Een kwartaaleffect wordt verkregen door aan de matrix F een submatrix toe te voegen:

$$\left(\begin{array}{cc|ccc} 1 & 1 & & & & 0 \\ 0 & 1 & & & & \\ \hline & & + & - & - & - \\ & & & -1 & -1 & -1 \\ & 0 & & 1 & 0 & 0 \\ \hline & & & 0 & 1 & 0 \end{array} \right)$$

zodat, als $Z(t)=(z_1, z_2, z_3, z_4, z_5)$ dan is $F Z(t) = (z_1+z_2, z_2, -z_3-z_4-z_5, z_3, z_4)$. Dus als z_3 het kwartaaleffect is op tijdstip t, dan is het kwartaaleffect op tijdstip $t-1$ gelijk aan z_4 en $t-2$ gelijk aan z_5 , het kwartaaleffect op tijdstip $t+1$ is nu gelijk aan $-z_3-z_4-z_5$, dus allen samen zijn telkens gelijk aan nul. De matrix H zal nu in het algemeen gelijk zijn aan (1 0 1 0 0). Een cycle wordt gemaakt door een rotatiematrix in de matrix F op te nemen.

Als nu storingsfouten bij de transitie worden toegelaten, dan wordt het Kalman-filter in staat gesteld enige afwijking in de nieuwe $Z(t+1)$ te kiezen (door het gewogen gemiddelde) waardoor $Z(t+1)$ kan verschillen van $F Z(t)$ op gekozen plaatsen. Hierdoor kunnen veranderingen in de componenten worden aangebracht.

Grofweg levert (5) een schatting op voor de toestand $Z(t+1)$ op basis van toestand $Z(t)$ (plus een onzekerheidsmaat). Daarnaast kan uit (6) ook een schatting gemaakt worden van $Z(t)$ (plus een onzekerheidsmaat). Uit deze samen kan een nieuwe schatting worden gedistilleerd (plus een onzekerheidsmaat). Dit wordt gedaan met behulp van (een variant) van het Kalman-filter. Als nu op een bepaald tijdstip geen waarneming van een bepaalde grootheid beschikbaar is dan kan men stellen dat de informatie een oneindige variantie heeft. De waarde is dan niet meer van belang en het model zal alleen afgaan op de prognose uit het vorige tijdstip. Zo kan op natuurlijke wijze omgegaan worden met lege cellen in bijvoorbeeld de OVG-cijfers of het niet systematisch beschikbaar zijn van gegevens.

De huidige implementatie gaat uit van een zeer grote variantie van de onbekende gegevens, te weten van de begintoestand en de ontbrekende gegevens terwijl de tendens is gebruik te gaan maken van een zogenaamd diffuus Kalman-filter dat uit gaat van een oneindige variantie maar dat

analytisch oplost. Zeer recent is 'exact initial Kalman filtering' (Koopman, 1997). Het is wellicht verstandig nog even te wachten met het doen van de volgende stap in deze richting.

Programmabeschrijving

In zijn algemeenheid is het model te flexibel voor een eenvoudige gebruikersomgeving. Om die reden moet een keuze worden gemaakt in de voor de gebruiker beschikbare alternatieven. Daarnaast bestaan er een aantal alternatieve programma's waarin de gebruikersomgeving uitgevoerd kan worden: bijvoorbeeld het World Wide Web. Op een dergelijke wijze kan het instrument op een (qua beheer) eenvoudige wijze beschikbaar worden gesteld aan een (selecte) groep gebruikers. Van een praktische implementatie van een WWW-omgeving is nog afgezien. Daarentegen is een eenvoudige omgeving op *Mathematica*TM geschreven.

Voor de praktische toepassing van het model is een (betrekkelijk) eenvoudige gebruikersomgeving gemaakt, gericht op eenvoudige toepassingen. Het gaat hierbij om het analyseren van een reeks van mobiliteitscijfers, verkeersdoden en gewonden over een tijdreeks. Deze reeks kan eventueel een seizoenseffect hebben. De methode levert alle soorten van de in de voorbeelden gegenereerde afbeeldingen.

Stel nu dat we beschikken over het volgende bestand gegevens: de eerste kolom stelt het jaar voor, de tweede kolom het kwartaal, de derde kolom de reizigerskilometers in dat kwartaal, de vierde kolom (een schatting van) de variante van dat cijfer, de vijfde kolom (bijvoorbeeld het aantal doden) en tenslotte de zesde kolom het aantal gewonden. De kolommen worden door spaties onderscheiden (space-delimited). Een voorbeeld van een bestand:

1985	1	321.41607051	82.89202768	293	2754
1985	2	381.51552282	110.56090236	356	3927
1985	3	347.25676071	106.42293880	400	3927
1985	4	382.33428139	120.37136245	389	3912
1986	1	335.53500006	88.54378031	320	3062
1986	2	412.67923663	140.18670857	389	3940
1986	3	377.64207739	129.90493660	385	3891
1986	4	365.56239929	107.06335338	433	3813
1987	1	349.20567170	91.24183593	303	2726
1987	2	397.35219682	131.20166135	375	3694
1987	3	379.84626359	98.85935259	380	3940
1987	4	389.96358268	110.12723134	427	3606
1988	1	368.10710142	105.21526987	356	3380
1988	2	427.68261373	146.16302024	312	3498
1988	3	388.68952919	134.38239904	326	3352
1988	4	390.82637783	104.93148894	372	3414
1989	1	378.62266866	98.03588151	323	3035
1989	2	419.04236208	124.22421531	363	3681
1989	3	386.10524306	156.51426227	408	3492
1989	4	424.08623470	139.82099163	362	3452
1990	1	397.10690125	138.25801646	374	3209
1990	2	432.29649581	134.64803103	326	3713
1990	3	404.71296775	120.56063452	321	3551
1990	4	401.05624651	105.66261597	355	3184
1991	1	406.94343570	126.00632581	287	2741
1991	2	451.67078441	178.61142941	328	3144
1991	3	412.82624760	155.21740924	381	3279

1991	4	427.83663998	158.27452123	285	2856
1992	1	409.85735242	124.15964272	273	2503
1992	2	442.75742980	154.33715288	325	3328
1992	3	427.06888059	148.31735927	356	2947
1992	4	439.74074436	140.79287145	331	2876
1993	1	408.18844026	150.44479400	266	2586
1993	2	452.27322455	167.58679242	361	3233
1993	3	410.66118645	141.26516582	308	3069
1993	4	434.30004755	155.86424738	317	2674
1994	1	443.79442301	78.17739815	277	2440
1994	2	487.40155952	91.39222550	372	3244
1994	3	453.76682014	75.13659042	321	3119
1994	4	490.38225624	88.58127493	328	2932
1995	1	462.05316972	25.20297529	320	2585
1995	2	506.78090287	32.41510886	332	3100
1995	3	483.59719896	32.14277774	332	3120
1995	4	493.38667482	28.74591223	350	2883
1996	1	464.64837681	30.21717300	242	2339
1996	2	492.33334390	28.00717317	325	3259
1996	3	474.87136287	33.27783480	308	3237
1996	4	494.21618049	31.59209407	305	3131
1997	1	478.62071239	32.14389615	238	2516
1997	2	514.00041706	36.38202279	298	3229
1997	3	493.13736959	32.53838055	338	3162
1997	4	502.49093536	27.31982298	289	2811

Voor dit bestand kan het volgende programma worden gebruikt:

Standaard model ontwikkeling VV

Invoer:

1) Jaar (eerste tijdseenheid)

2 (eventueel) kwartaal of maand, tweede tijdseenheid

2 - 3 expositie

3-4) Variantie expositie

4-5) Aantal doden

5-6) Aantal gewonden

"Selecteer invoer bestand"

"Selecteer uitvoer bestand"

"Job title"

"Prognose lengte"

"Lengte Seizoen"

"Kappa coëfficiënt"

Invoer bestand: M:\ONTWVV\Totaal.dat

Uitvoer bestand: M:\ONTWVV\Totaal.m

Job title: Job: M:\ONTWVV\Totaal.dat-4

Prognose lengte: 4

Seizoen lengte: 4

Kappa: 100

"Zet model op"

"Uitvoeren!"

Uitleg functies van de knoppen:

- De knop "Selecteer invoer bestand" start een procedure

waarmee het bovenstaande invoer bestand moet worden geselecteerd. Eigenlijk wordt hier een 'save as'-procedure mee opgestart. Het bericht 'weet u zeker dat u wat wilt overschrijven' dat *Mathematica* genereert kan gerust worden genegeerd.

De **"Selecteer uitvoer bestand"** -knop wordt gebruikt om een bestand te maken met de technische uitvoerresultaten: niet de afbeeldingen en dergelijke. Het bestand dat hiermee wordt aangegeven wordt, indien het reeds bestaat, wel degelijk overschreven met een nieuwe versie.

De **"Job title"** -knop wordt gebruikt om een titel (voor in de grafieken) op te geven.

De **"Prognose lengte"** -knop wordt gebruikt om op te geven voor welke periode prognoses gemaakt moeten worden. De drie typen analyses uit de voorbeelden worden uitgevoerd:

- Het eerste type is een analyse waarbij de observaties van de prognoselengte worden weggelaten uit de schattingsprocedure, om vervolgens vergeleken te worden met de prognoses. Deze analyse dient om een indruk te krijgen van de betrouwbaarheid van de prognoses.
- Een tweede analyse wordt uitgevoerd waarbij de expositiecijfers wel in de prognoseperiode worden meegenomen. Dit levert prognosecijfers op gegeven een (wellicht afwijkende) mobiliteitsontwikkeling. Deze analyse kan gebruikt worden om de ontwikkeling van de risico's in het laatste prognoselengte-aantal observaties te bestuderen.
- Tenslotte wordt een analyse uitgevoerd waarin alle beschikbare observaties worden gebruikt en een prognoselengte-aantal observaties in de toekomst (eigenlijk: na de laatste observatie) wordt voorspeld.

De **"Lengte Seizoen"** -knop wordt gebruikt om aan te geven of er een tweede tijdseenheid kolom in het bestand aanwezig is. Dit kan een seizoen voorstellen, bijvoorbeeld een kwartaal of maand. Indien de waarde nul wordt gegeven wordt de kolom verondersteld afwezig te zijn. Er zullen dan in totaal vijf kolommen worden ingelezen. Het getal 'één' als argument zal tot problemen leiden, daar een seizoenseffect dat precies één observatie duurt niet geschat kan worden. Hierop wordt door het programma niet gecontroleerd. Voor data per kwartaal moet een 'vier' worden ingevuld, voor data per maand een 'twaalf'.

De **"Kappa coëfficiënt"** -knop wordt gebruikt om de onzekerheid in de begintoestand aan te geven. Hier moet een voldoende groot getal worden ingevuld, zodat de gevolgen van deze keuze verwaarloosbaar zijn. Het getal 1000 wordt bij verstek door het programma zelf ingevuld.

Met de **"Zet model op"** -knop wordt een eerste stap uitgevoerd waarbij de basisinstellingen van het model intern worden vastgesteld. Na deze stap kunnen ingewijden nog bijzondere aanpassingen in het model aanbrengen.

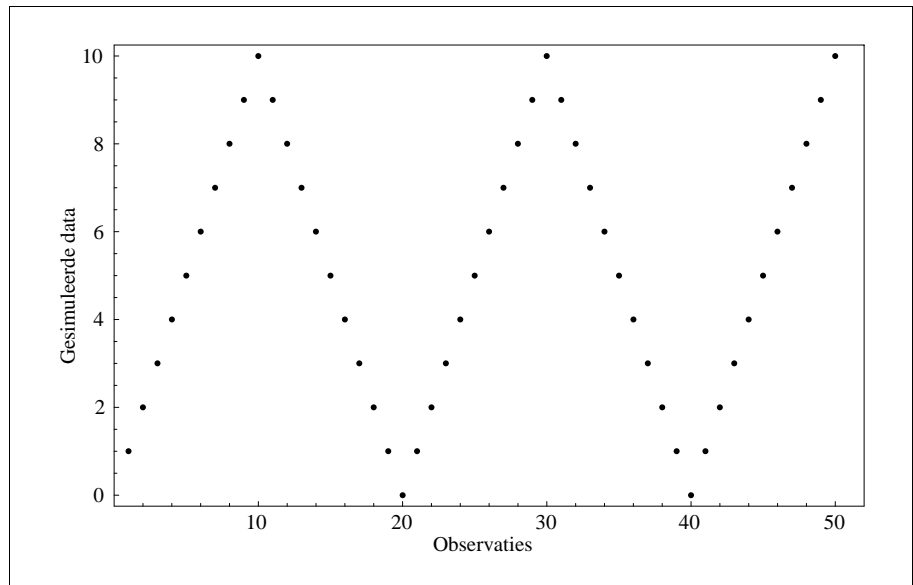
De **"Uitvoeren!"** -knop start de procedure. In een nieuw 'untitled' document wordt gemaakt, waarin de iteratiegeschiedenis en de resultaten worden weggeschreven. Dit document kan desgewenst geprint of bewaard worden.

Bijlage 2

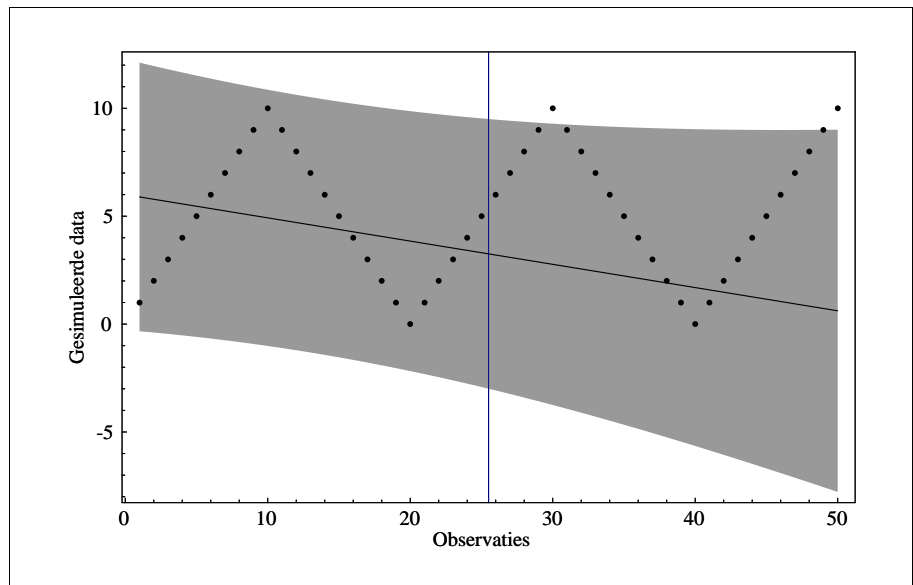
Afbeeldingen 1 t/m 24

In *Afbeelding 1 t/m Afbeelding 24* worden de observaties met stippen aangegeven en de betreffende modelprognoses met een doorgetrokken lijn. De bijbehorende 95%-betrouwbaarheidsintervallen voor het gemiddelde zijn met een grijs tint aangegeven.

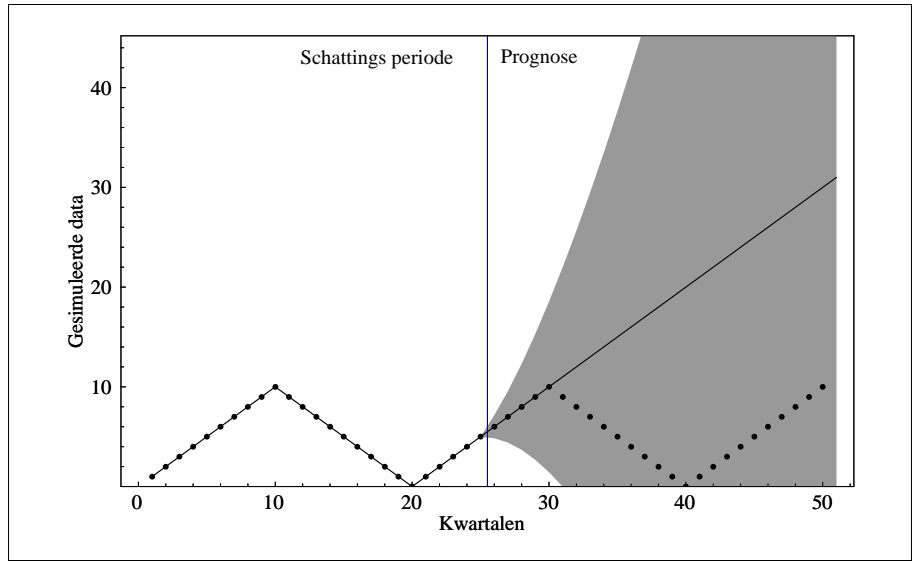
1. *Voorbeeld 1 (Simulatie)*



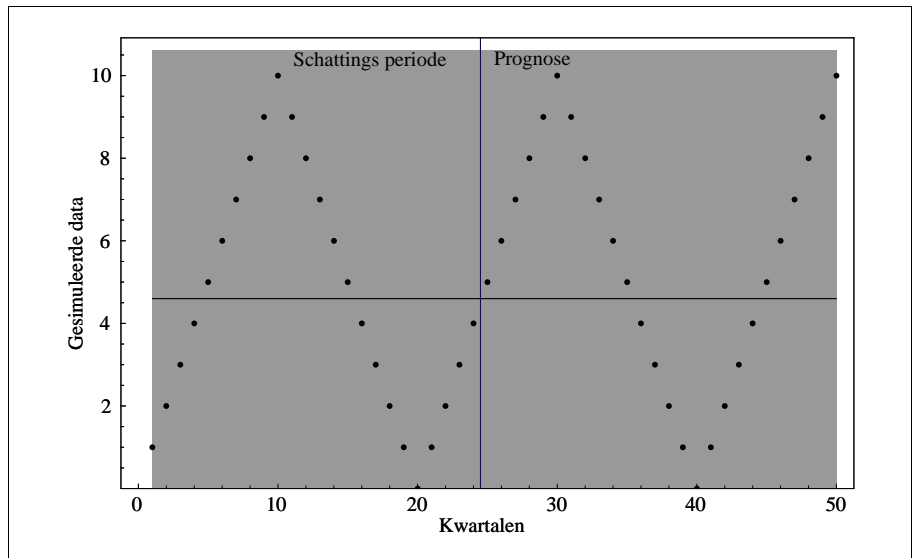
Afbeelding 1. *Gesimuleerde data.*



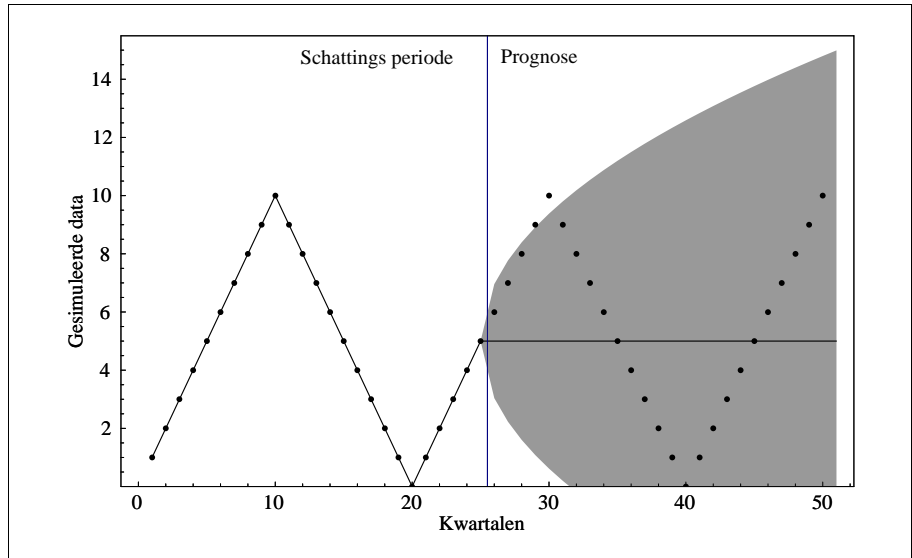
Afbeelding 2. *Lineaire prognose gesimuleerde data.*



Afbeelding 3. *Prognoses gesimuleerde data met behulp van het nieuwe instrument.*

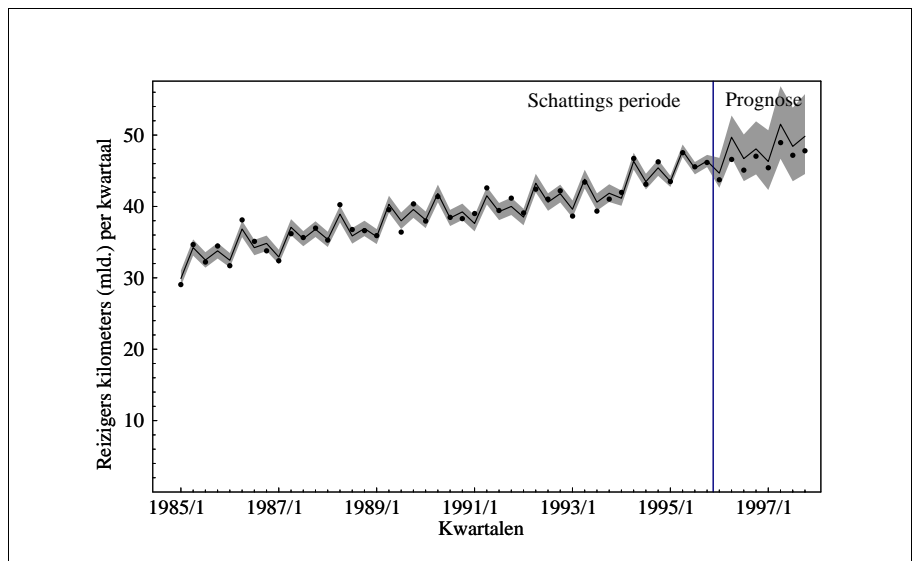


Afbeelding 4. *Prognoses gesimuleerde data met behulp van het gemiddelde.*

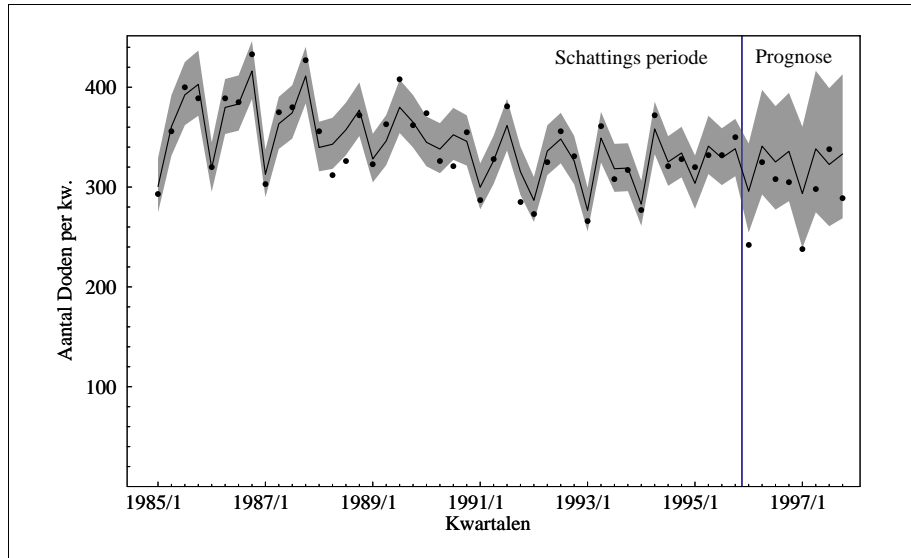


Afbeelding 5. Prognoses gesimuleerde data met behulp van het 'random walk'-model.

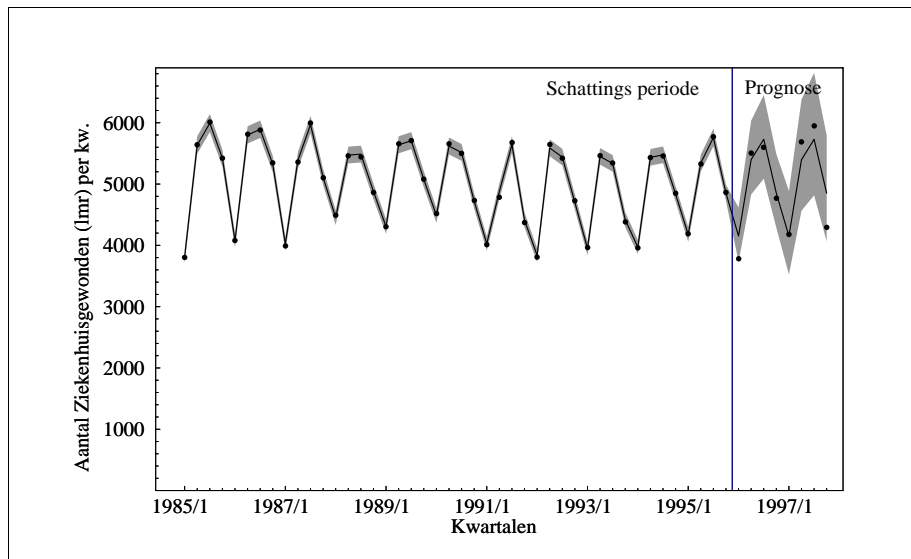
2. Alle leeftijden en alle wijzen van vervoer.



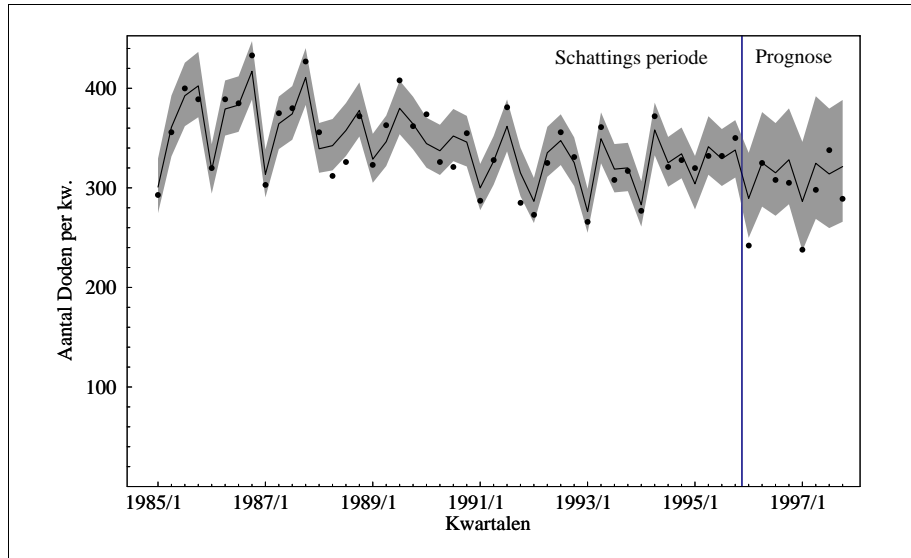
Afbeelding 6a. Schattingen en prognoses (1996 + 1997) mobiliteitscijfers. Alle leeftijden en alle wijzen van vervoer.



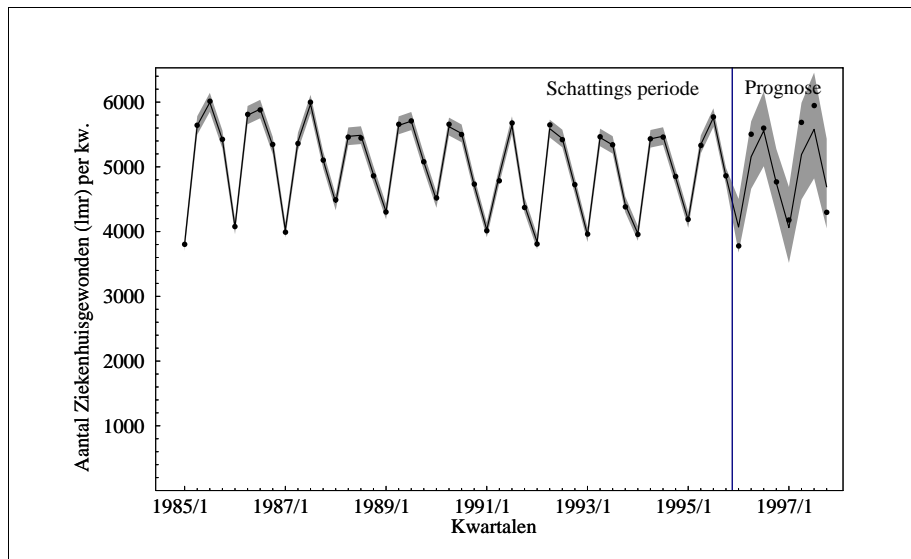
Afbeelding 6b. Schattingen en prognoses (1996 + 1997) verkeersdoden. Alle leeftijden en alle wijzen van vervoer.



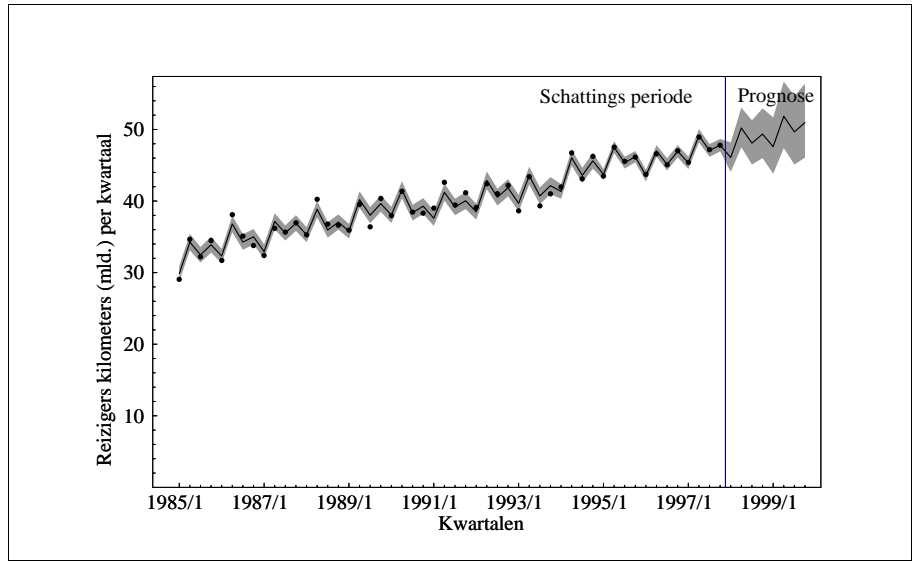
Afbeelding 6c. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden. Alle leeftijden en alle wijzen van vervoer.



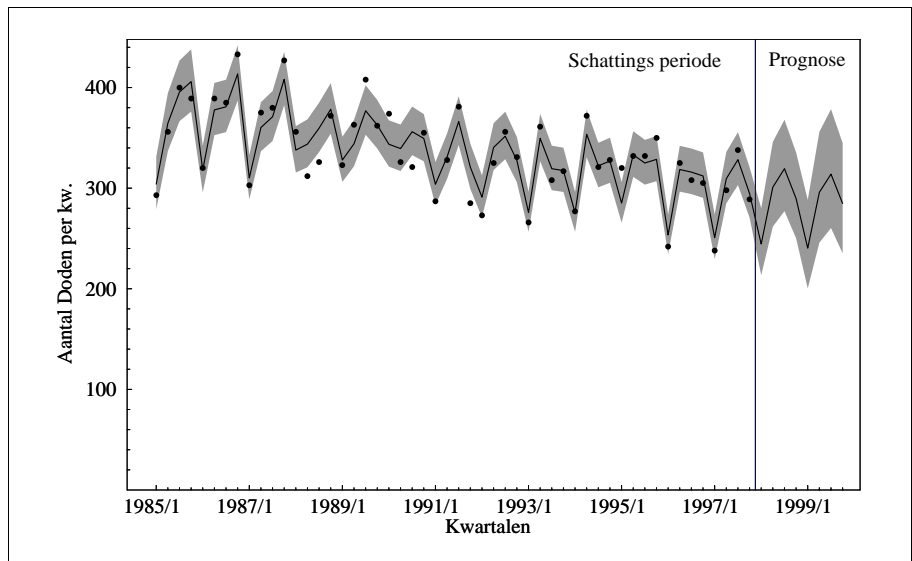
Afbeelding 7a. Schattingen en prognoses (1996 + 1997) doden gegeven mobiliteit. Alle leeftijden en alle wijzen van vervoer.



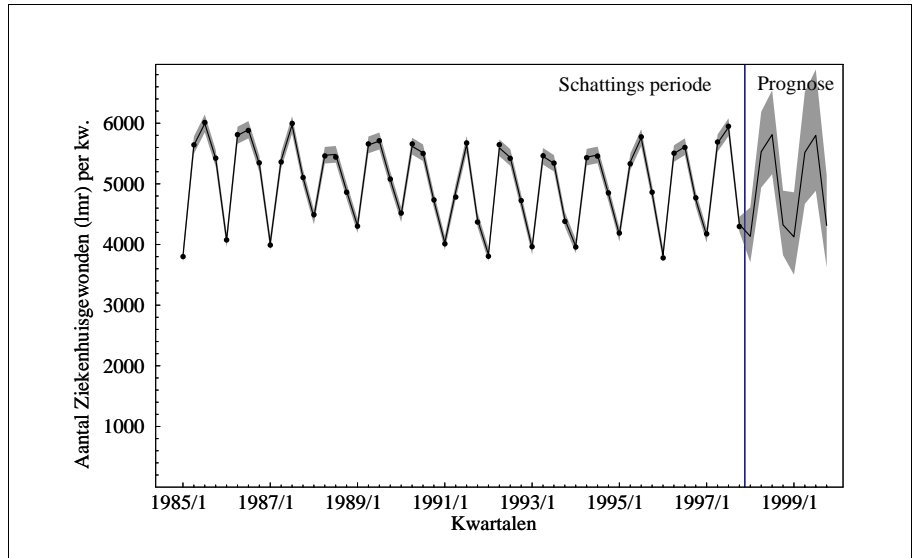
Afbeelding 7b. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden gegeven mobiliteit. Alle leeftijden en alle wijzen van vervoer.



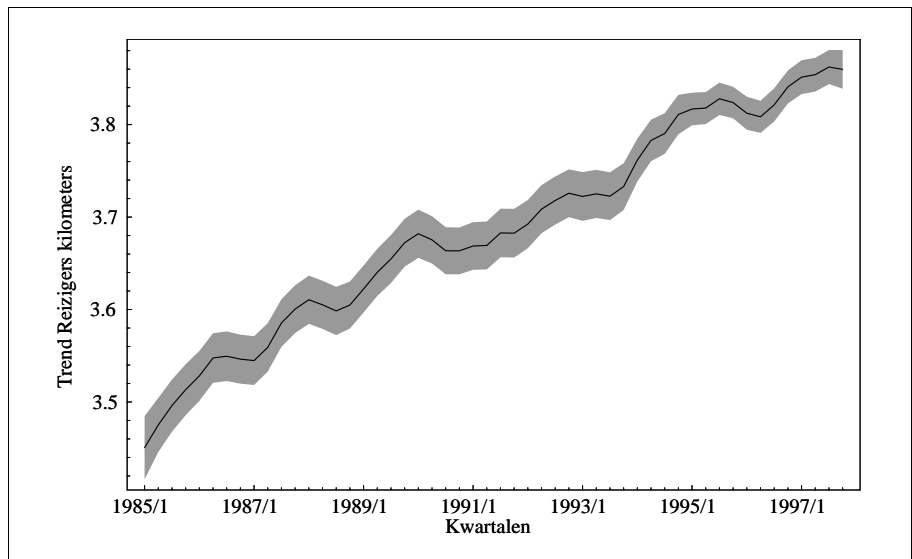
Afbeelding 8a. *Schattingen en prognoses (1998 + 1999) reizigerskilometers. Alle leeftijden en alle wijzen van vervoer.*



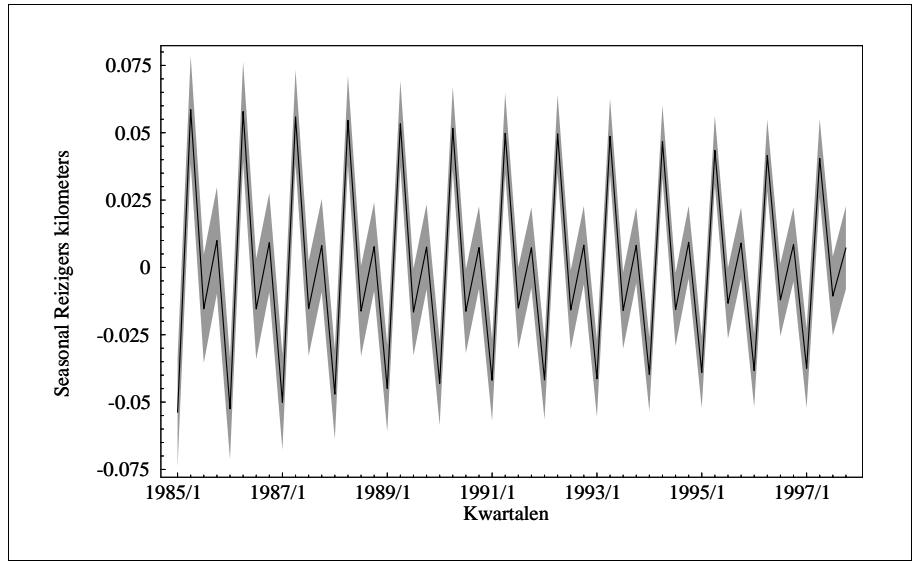
Afbeelding 8b. *Schattingen en prognoses (1998 + 1999) verkeersdoden. Alle leeftijden en alle wijzen van vervoer.*



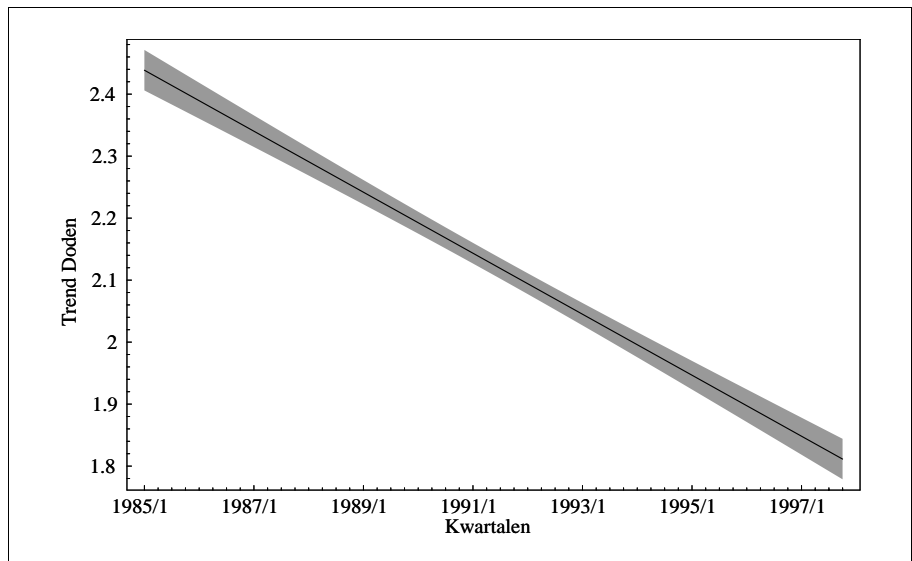
Afbeelding 8c. Schattingen en prognoses (1998 + 1999) ziekenhuisgewonden. Alle leeftijden en alle wijzen van vervoer.



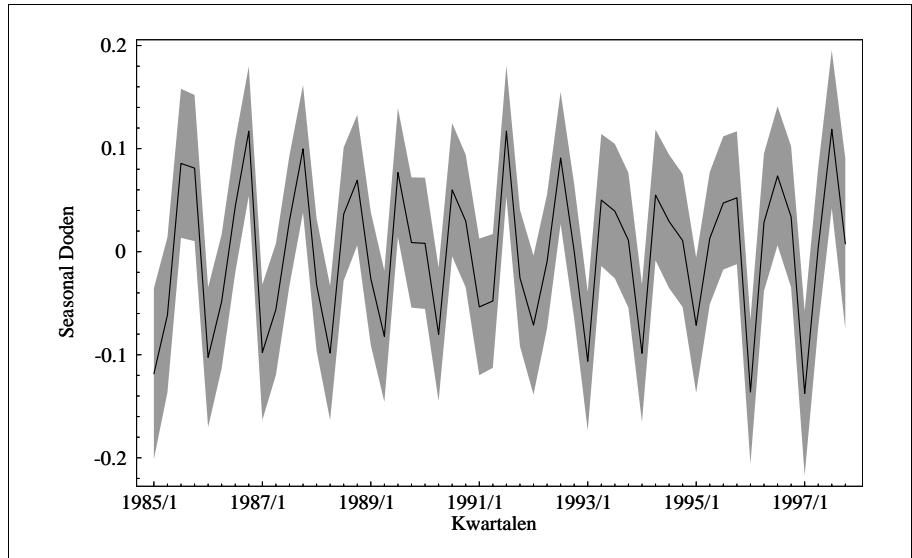
Afbeelding 9a. Trend reizigerskilometers. Alle leeftijden en alle wijzen van vervoer.



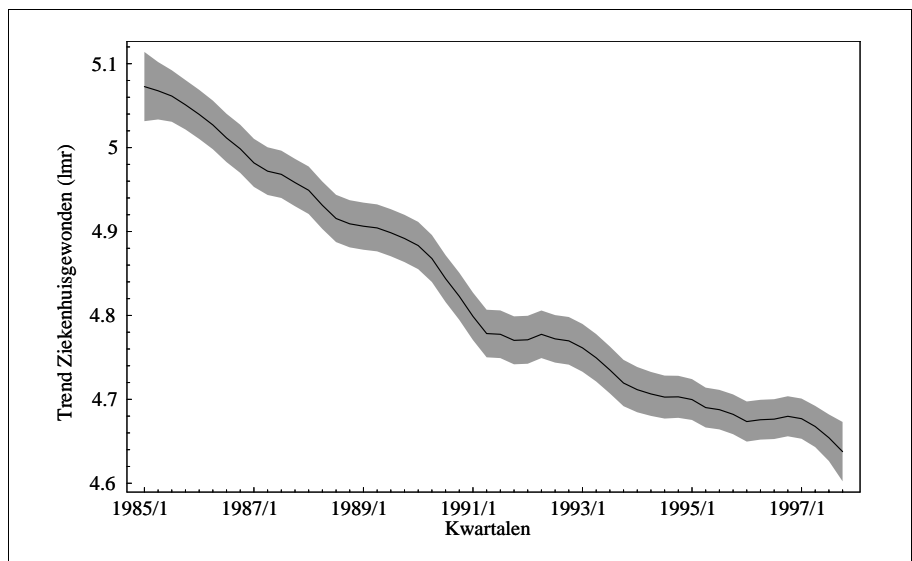
Afbeelding 9b. *Seizoenseffect reizigerskilometers. Alle leeftijden en alle wijzen van vervoer.*



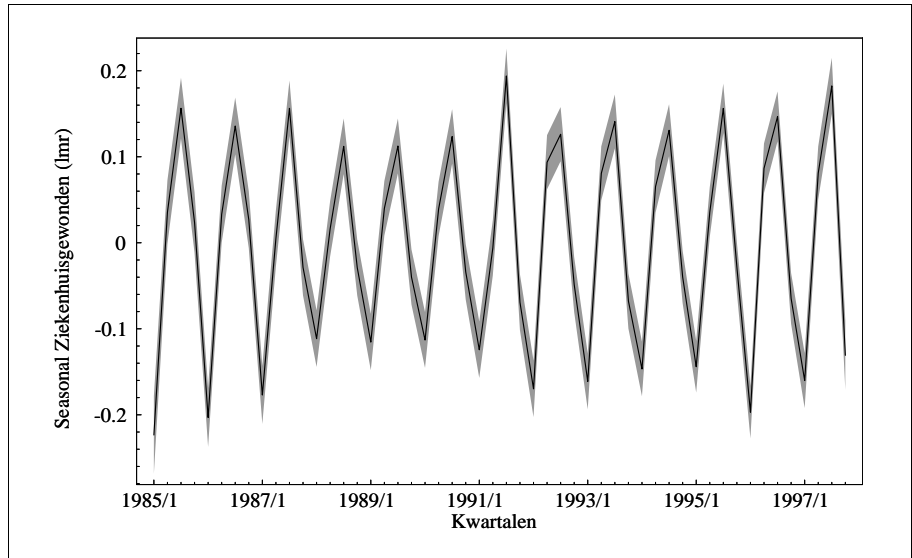
Afbeelding 9c. *Trend verkeersdoden. Alle leeftijden en alle wijzen van vervoer.*



Afbeelding 9d. *Seizoenseffect verkeersdoden. Alle leeftijden en alle wijzen van vervoer.*

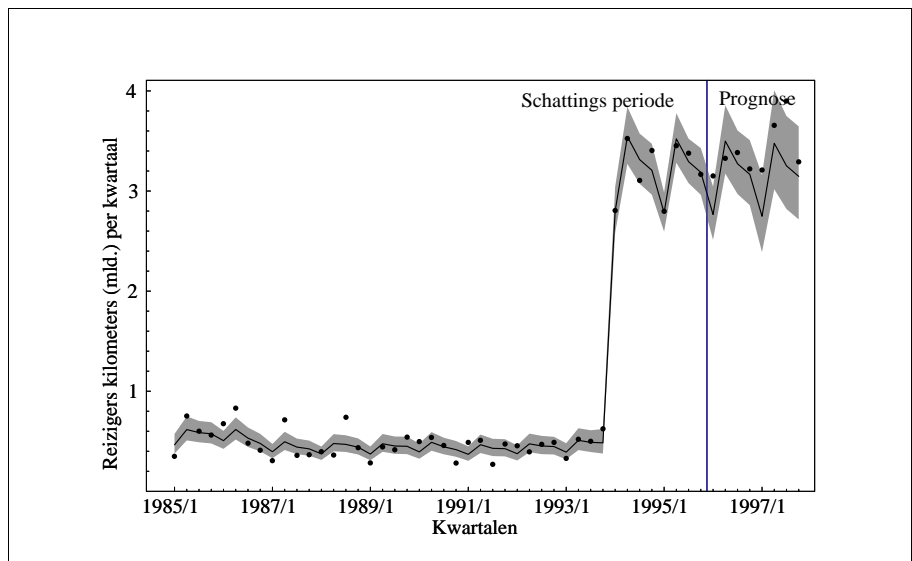


Afbeelding 9e. *Trend ziekenhuisgewonden. Alle leeftijden en alle wijzen van vervoer.*

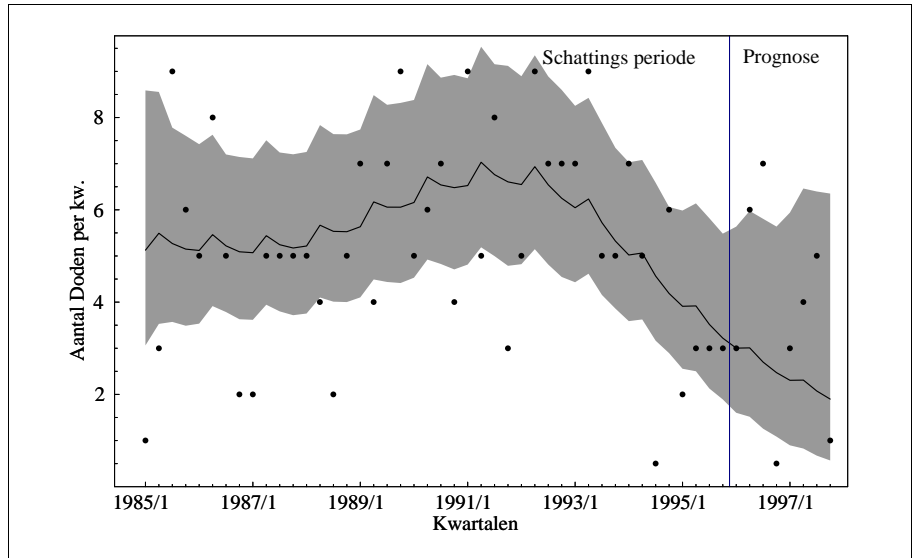


Afbeelding 9f. *Seizoenseffect ziekenhuisgewonden. Alle leeftijden en alle wijzen van vervoer.*

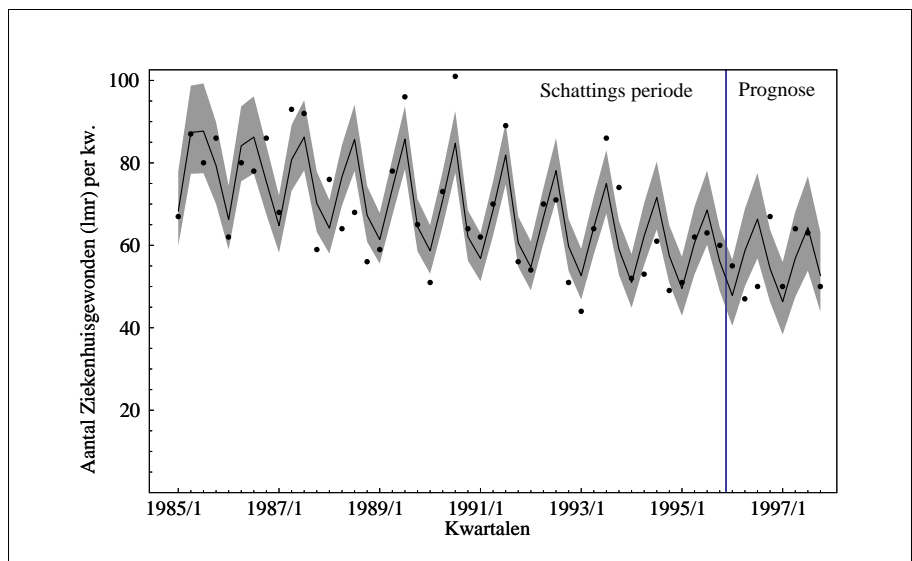
3. *0- tot14-jarige personenauto-inzittenden*



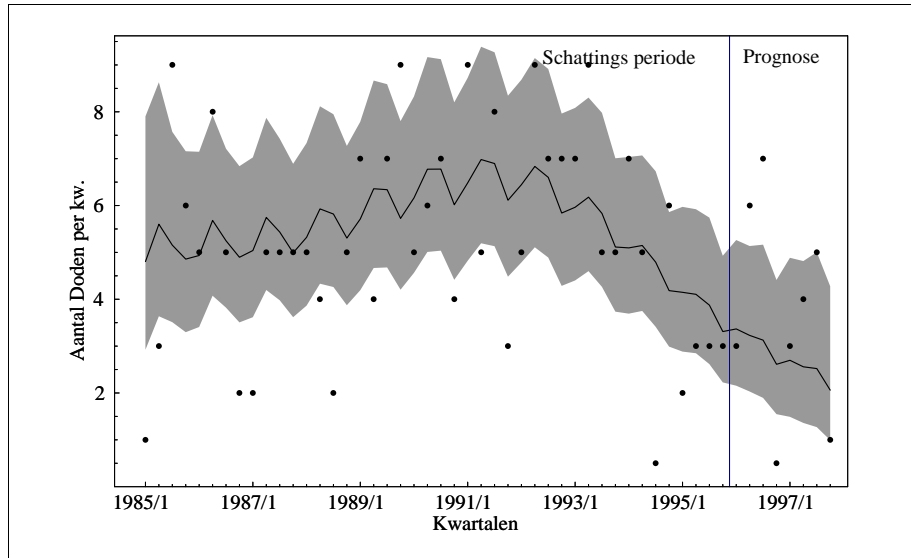
Afbeelding 10a. *Schattingen en prognoses (1996 + 1997) mobiliteitscijfers. 0- tot 14-jarige personenauto-inzittenden.*



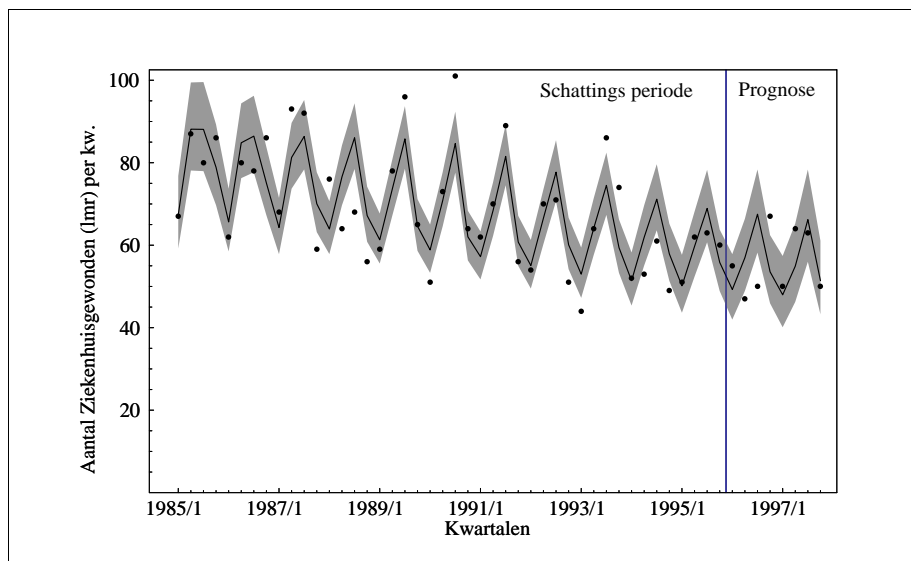
Afbeelding 10b. Schattingen en prognoses (1996 + 1997) verkeersdoden. 0- tot 14-jarige personenautoinzittenden.



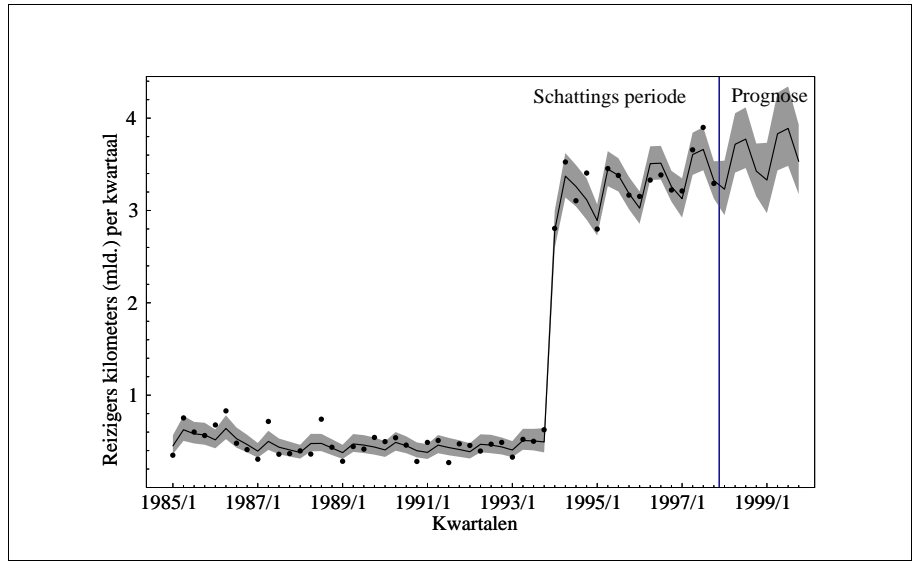
Afbeelding 10c. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden. 0- tot 14-jarige personenauto-inzittenden.



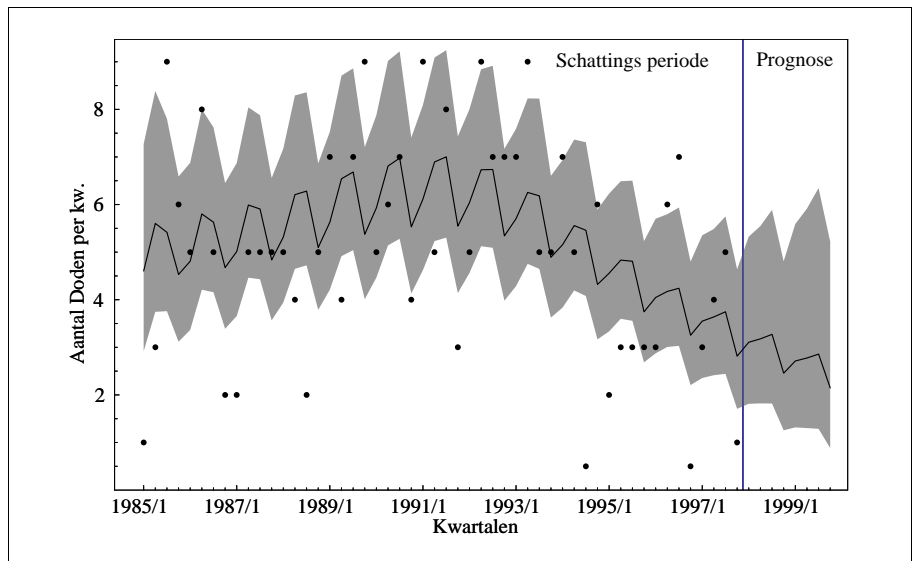
Afbeelding 11a. Schattingen en prognoses (1996 + 1997) verkeersdoden gegeven mobiliteit. 0-14 jarigen personenauto-inzittenden.



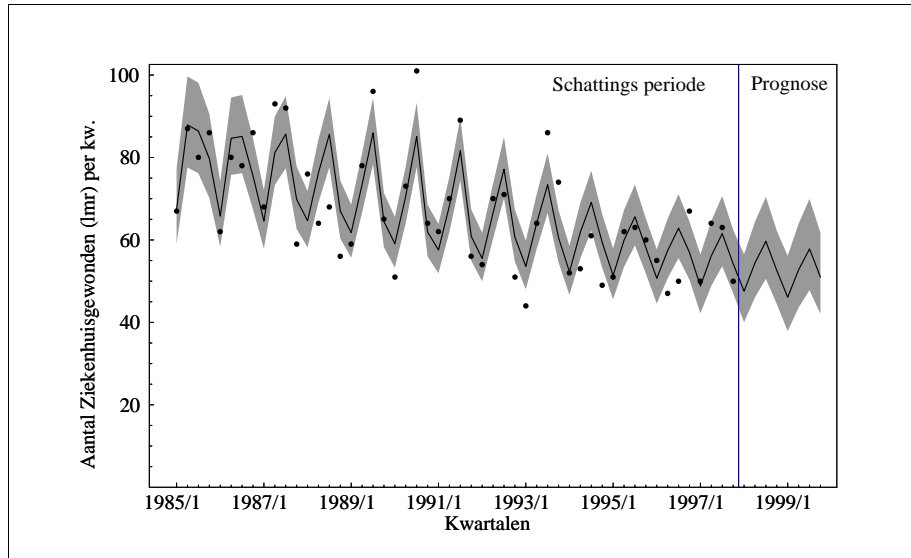
Afbeelding 11b. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden gegeven mobiliteit. 0- tot 14-jarige personenauto-inzittenden.



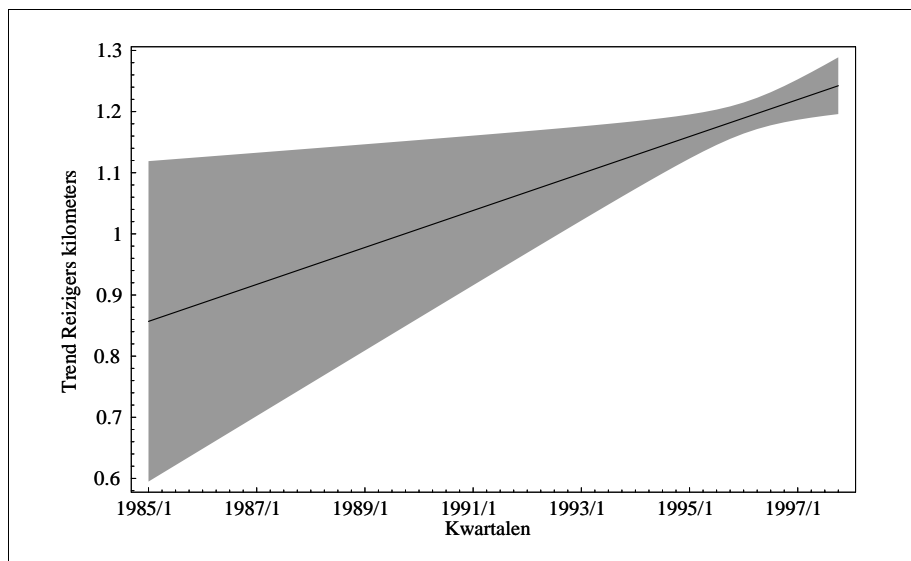
Afbeelding 12a. Schattingen en prognoses (1998 + 1999) reizigerskilometers. 0- tot 14-jarige personenauto-inzittenden.



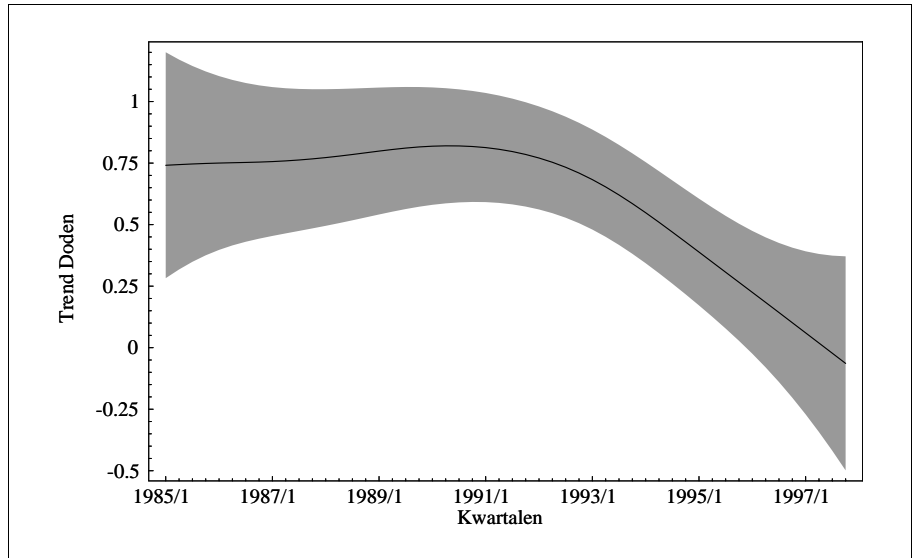
Afbeelding 12b. Schattingen en prognoses (1998 + 1999) verkeersdoden. 0-tot 14-jarige personenauto-inzittenden.



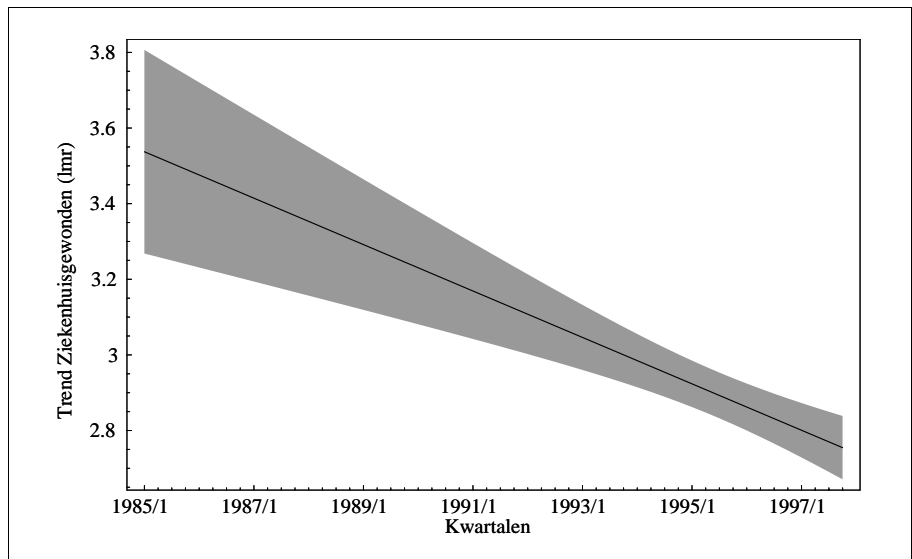
Afbeelding 12c. Schattingen en prognoses (1998 + 1999) ziekenhuisgewonden. 0- tot 14-jarige personenauto-inzittenden.



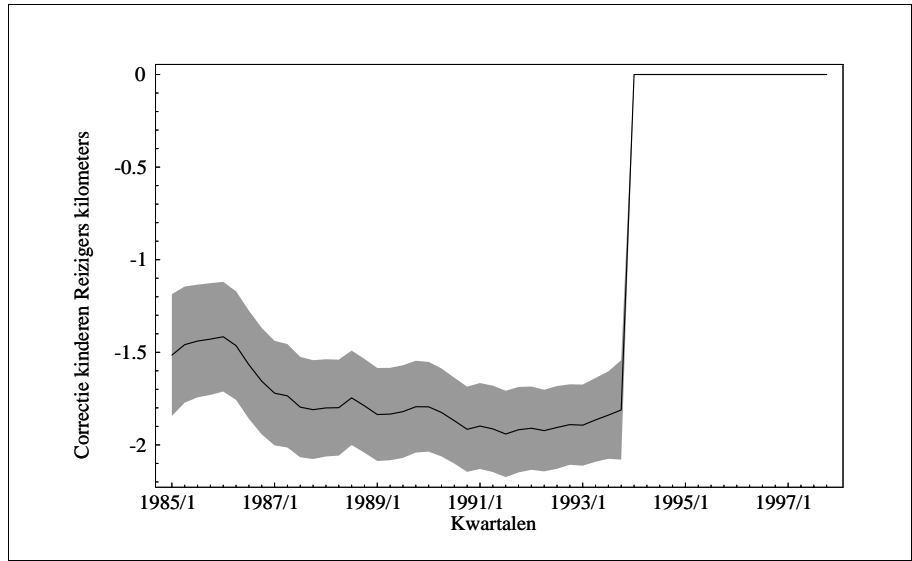
Afbeelding 13a. Trend reizigerskilometers. 0- tot 14-jarige personenauto-inzittenden.



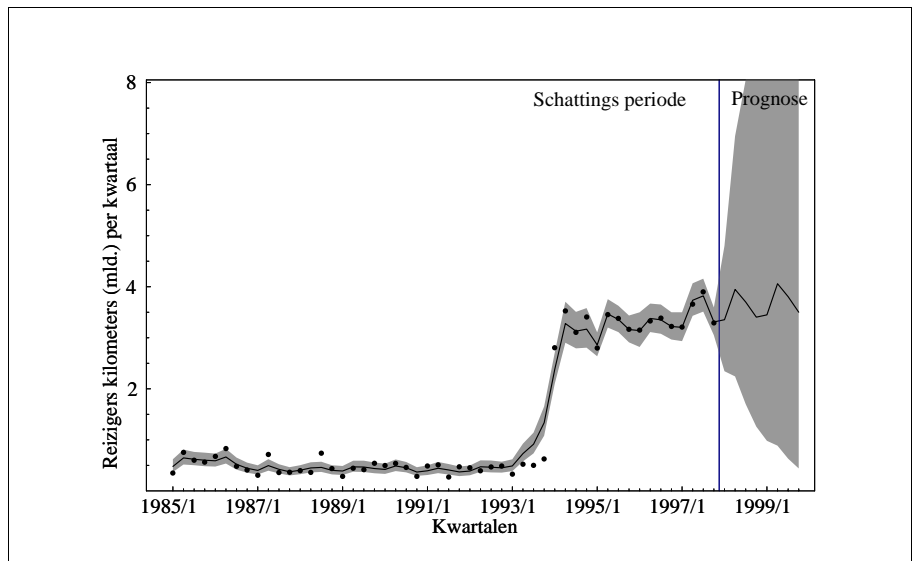
Afbeelding 13b. *Trend verkeersdoden. 0- tot 14-jarige personenauto-inzittenden.*



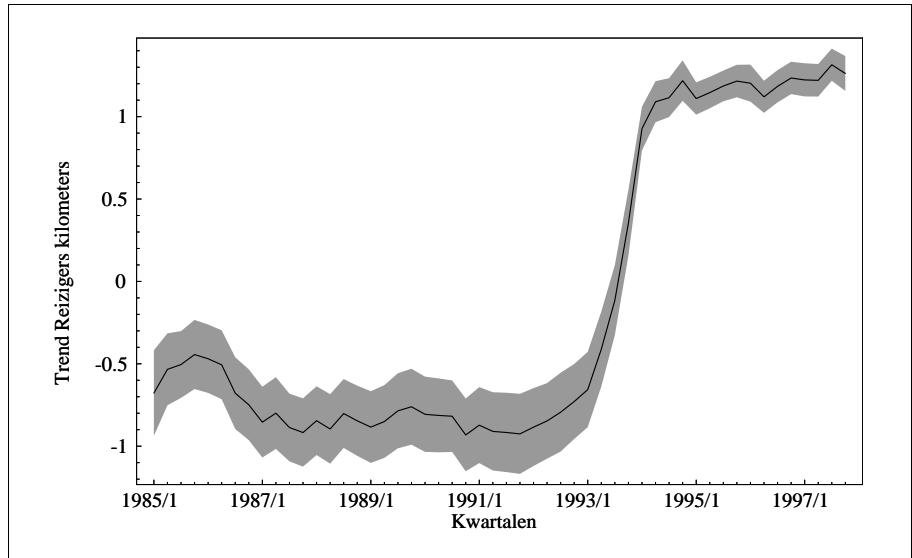
Afbeelding 13c. *Trend ziekenhuisgewonden. 0- tot 14-jarige personenauto-inzittenden.*



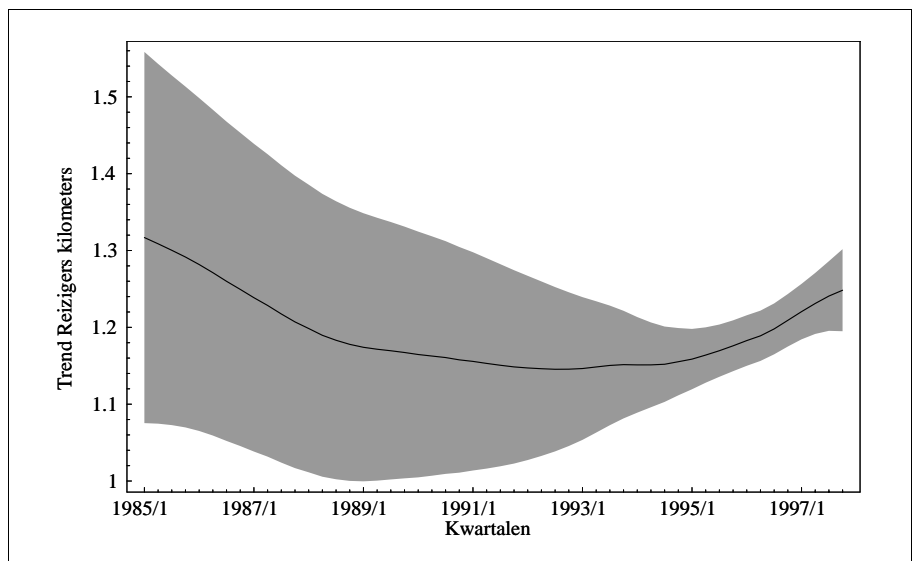
Afbeelding 13d. Correctie 0- tot 12-jarigen. 0- tot 14-jarige personenauto-inzittenden.



Afbeelding 14a. Prognoses reizigerskilometers, zonder correctie 0- tot 12-jarigen. 0- tot 14-jarige personenauto-inzittenden.

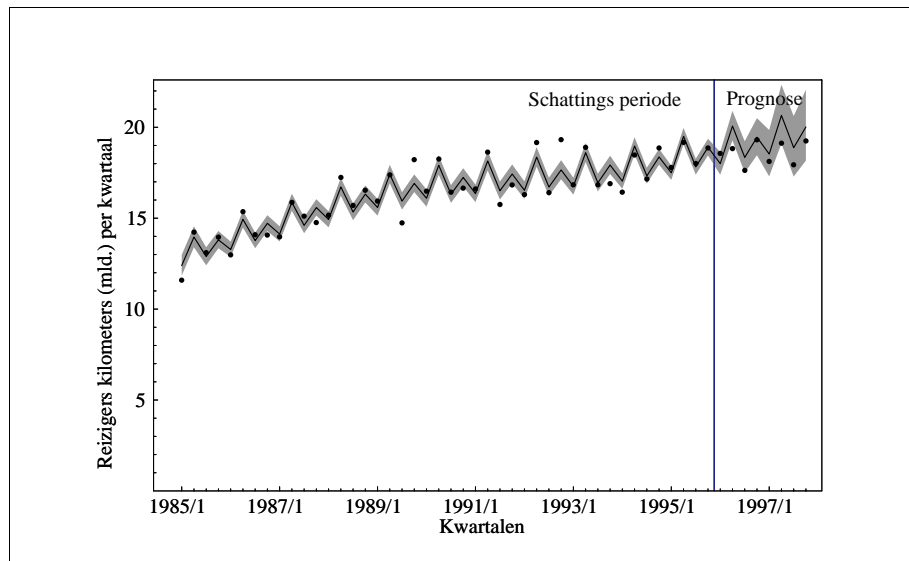


Afbeelding 14b. *Trend reizigerskilometers, zonder correctie 0- tot 12-jarigen. 0- tot 14-jarige personenauto-inzittenden.*

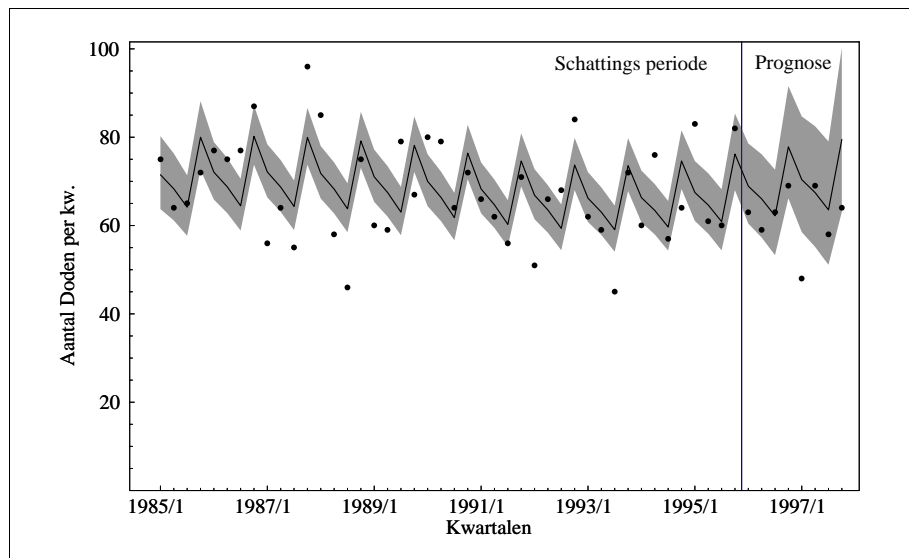


Afbeelding 14c. *Trend reizigerskilometers, met niet-tijd-variërende correctie 0- tot 12-jarigen. 0- tot 14-jarige personenauto-inzittenden.*

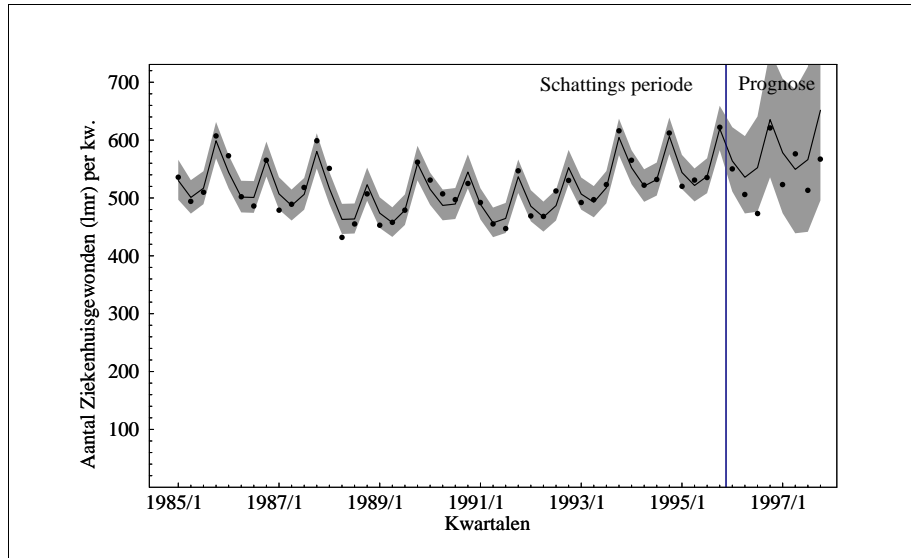
4. 25- tot 49-jarige personenauto-inzittenden



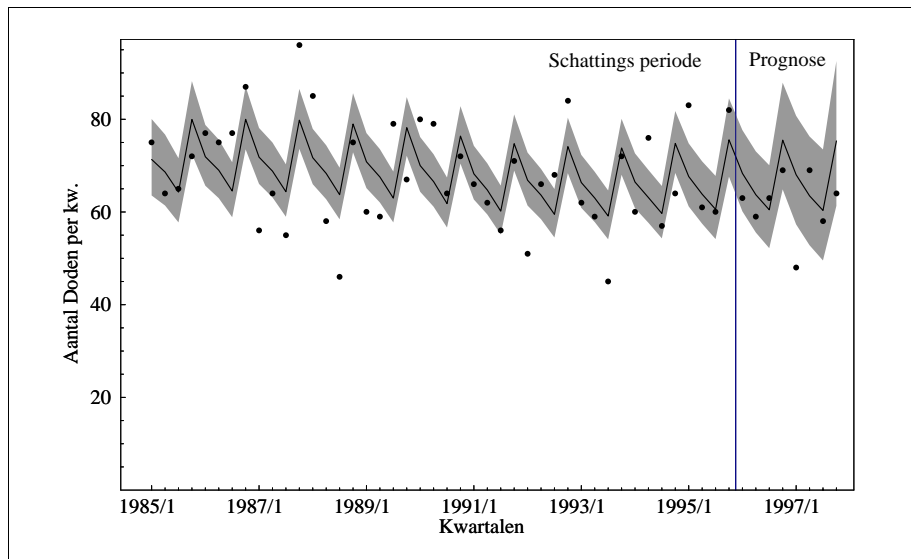
Afbeelding 15a. Schattingen en prognoses (1996 + 1997) mobiliteitscijfers. 25- tot 49-jarige personenauto-inzittenden.



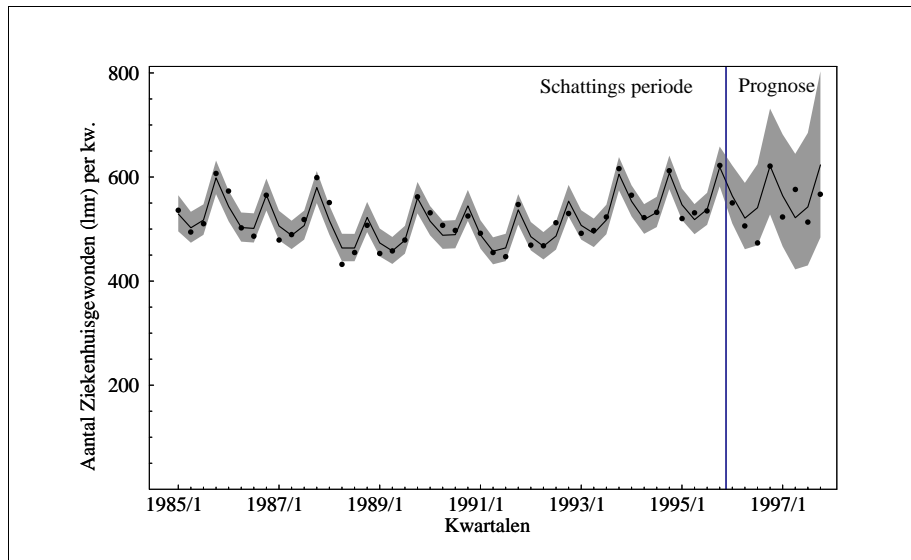
Afbeelding 15b. Schattingen en prognoses (1996 + 1997) verkeersdoden. 25- tot 49-jarige personenauto-inzittenden.



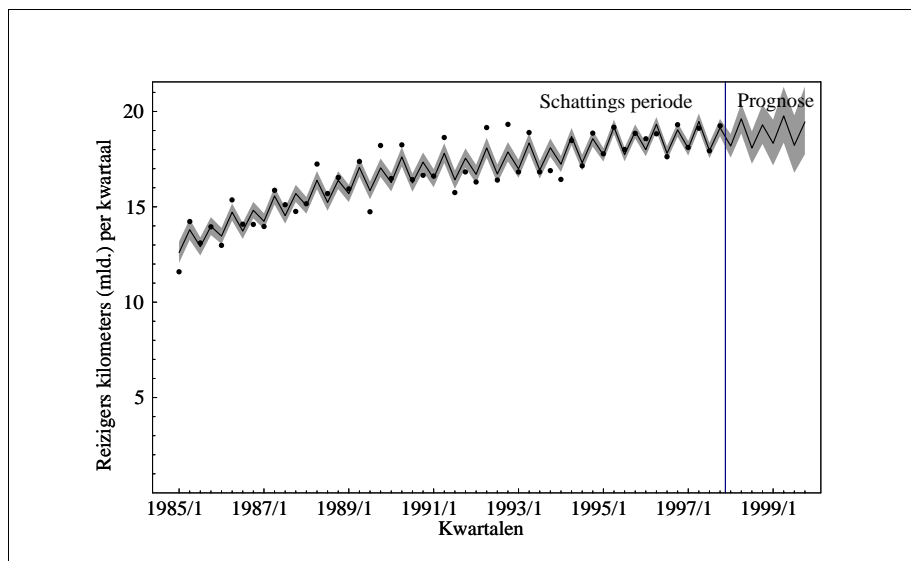
Afbeelding 15c. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden. 25- tot 49-jarige personenauto-inzittenden.



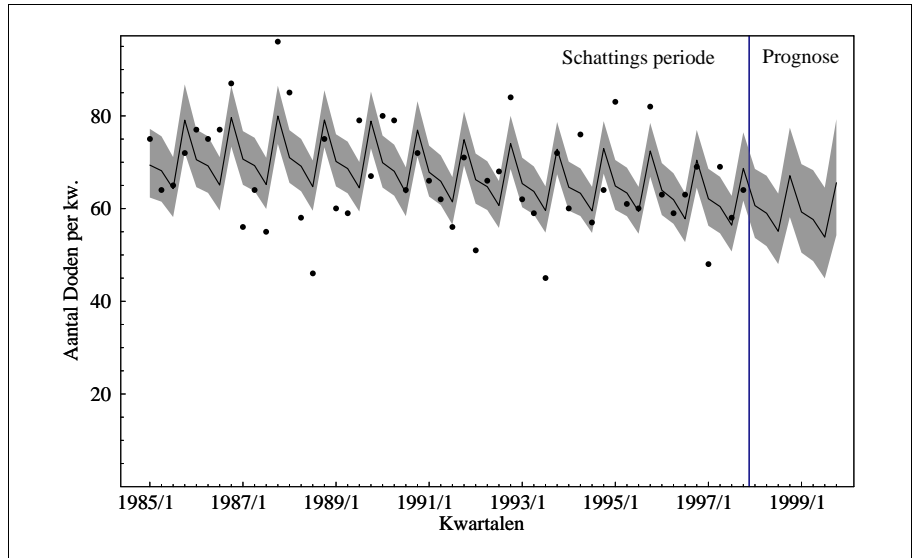
Afbeelding 16a. Schattingen en prognoses (1996 + 1997) verkeersdoden gegeven mobiliteit. 25- tot 49-jarige personenauto-inzittenden.



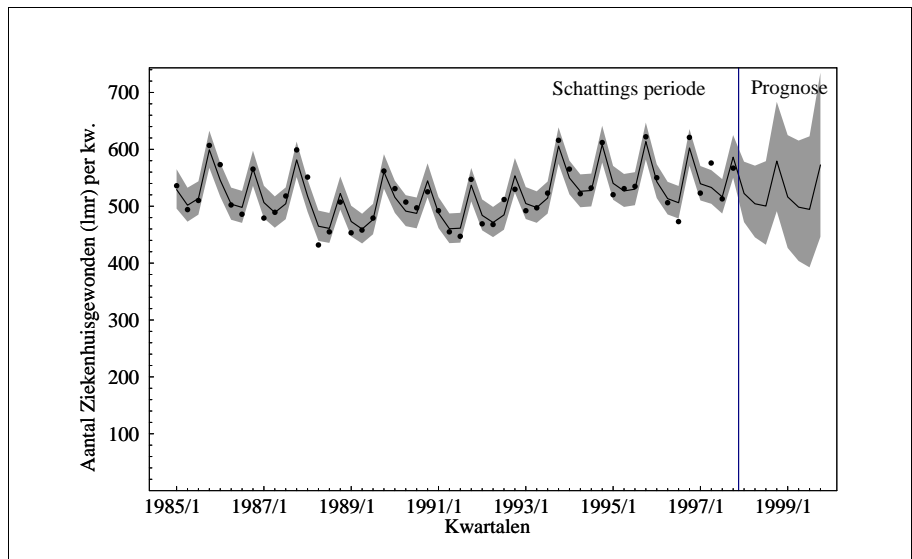
Afbeelding 16b. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden gegeven mobiliteit. 25- tot 49-jarige personenauto-inzittenden.



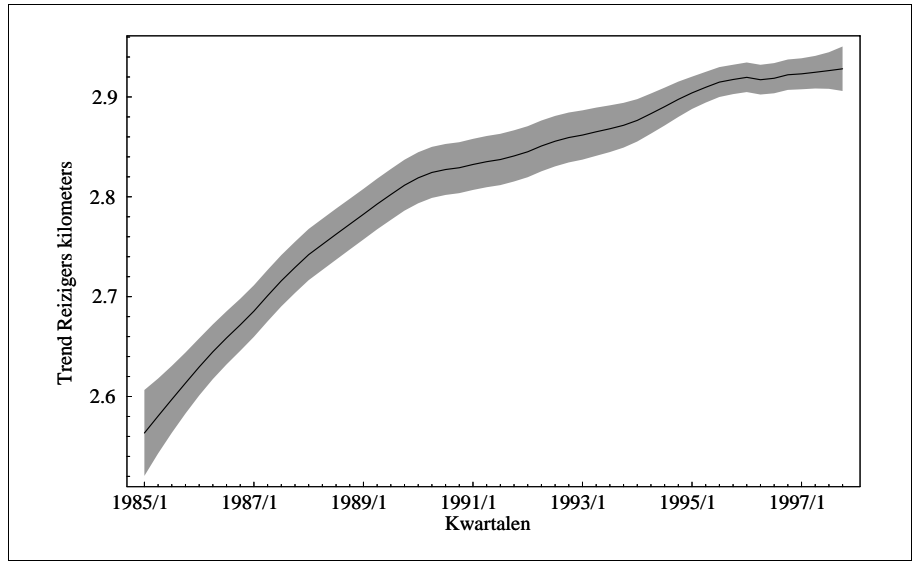
Afbeelding 17a. Schattingen en prognoses (1998 + 1999) reizigerskilometers. 25- tot 49-jarige personenauto-inzittenden.



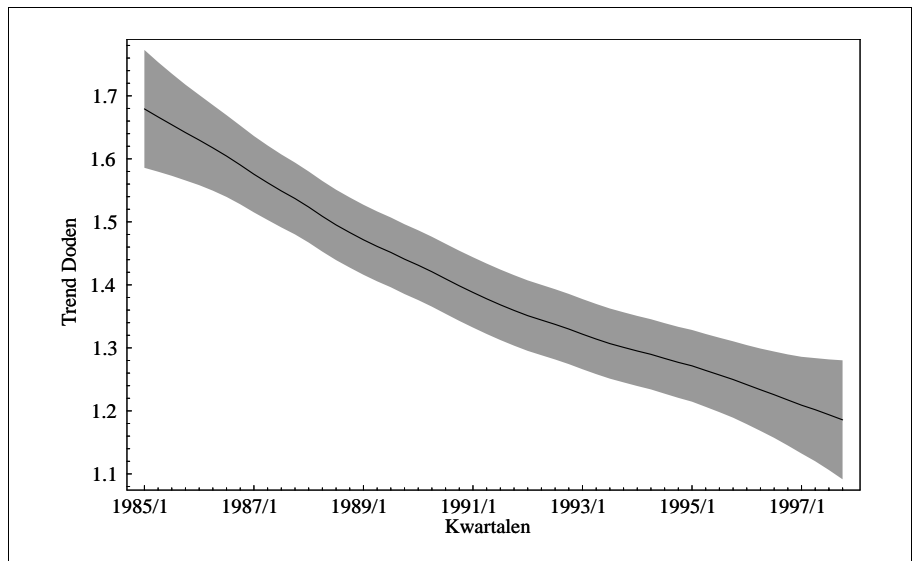
Afbeelding 17b. Schattingen en prognoses (1998 + 1999) verkeersdoden. 25- tot 49-jarige personenauto-inzittenden.



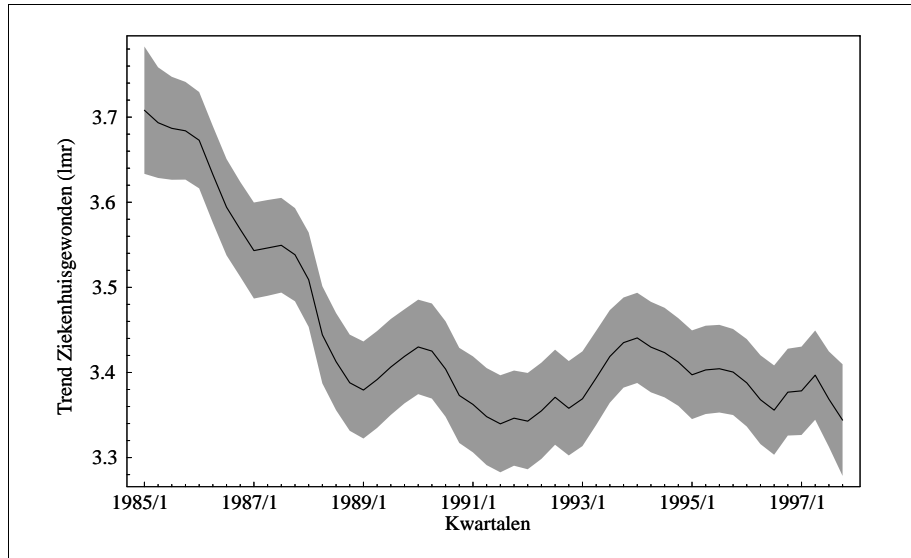
Afbeelding 17c. Schattingen en prognoses (1998 + 1999) ziekenhuisgewonden. 25- tot 49-jarige personenauto-inzittenden.



Afbeelding 18a. *Trend reizigerskilometers. 25- tot 49-jarige personenauto-inzittenden.*

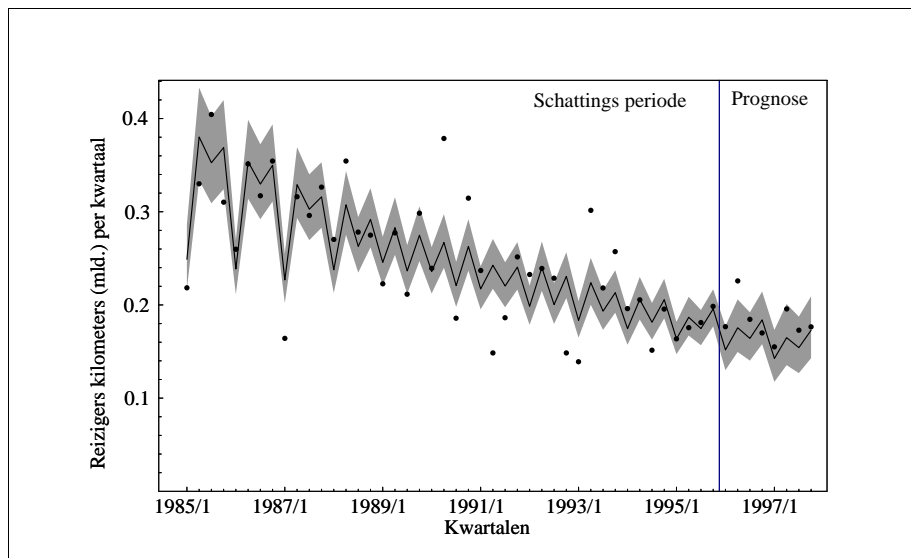


Afbeelding 18b. *Trend verkeersdoden. 25- tot 49-jarige personenauto-inzittenden.*

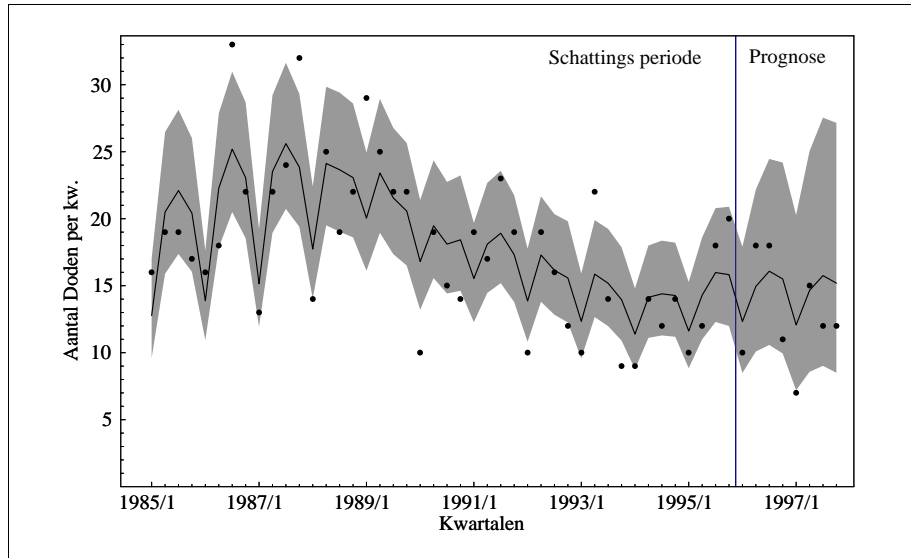


Afbeelding 18c. Trend ziekenhuisgewonden. 25- tot 49-jarige personenauto-inzittenden.

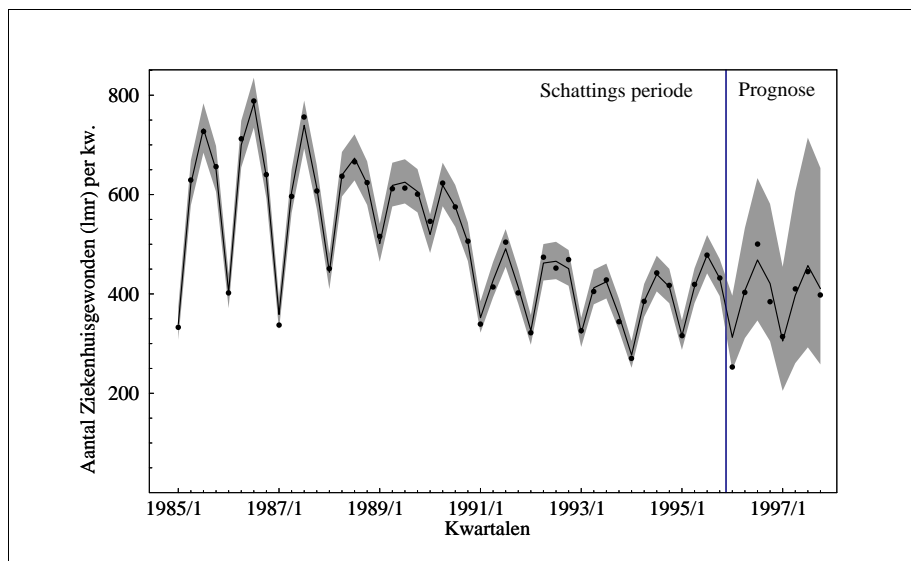
5. 15- tot 24-jarige bromfietzers



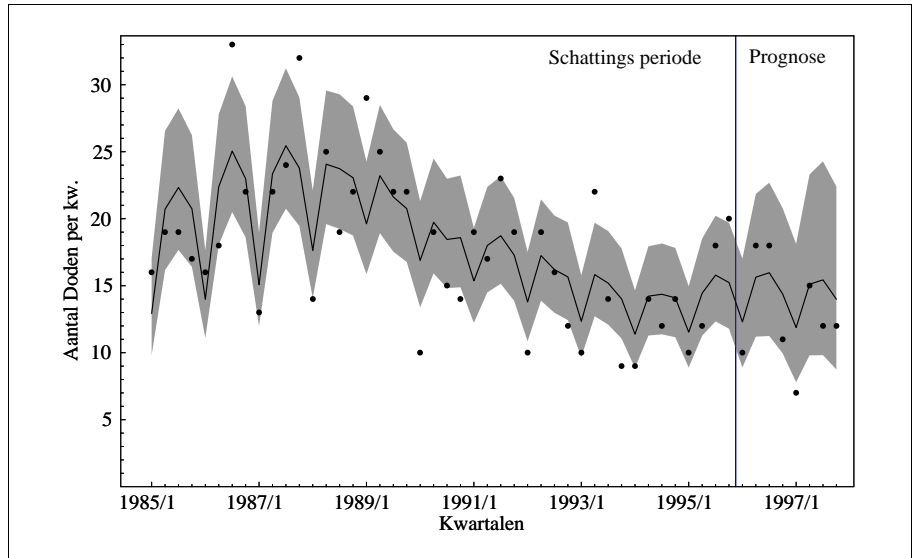
Afbeelding 19a. Schattingen en prognoses (1996 + 1997) mobiliteitscijfers. 15- tot 24-jarige bromfietzers.



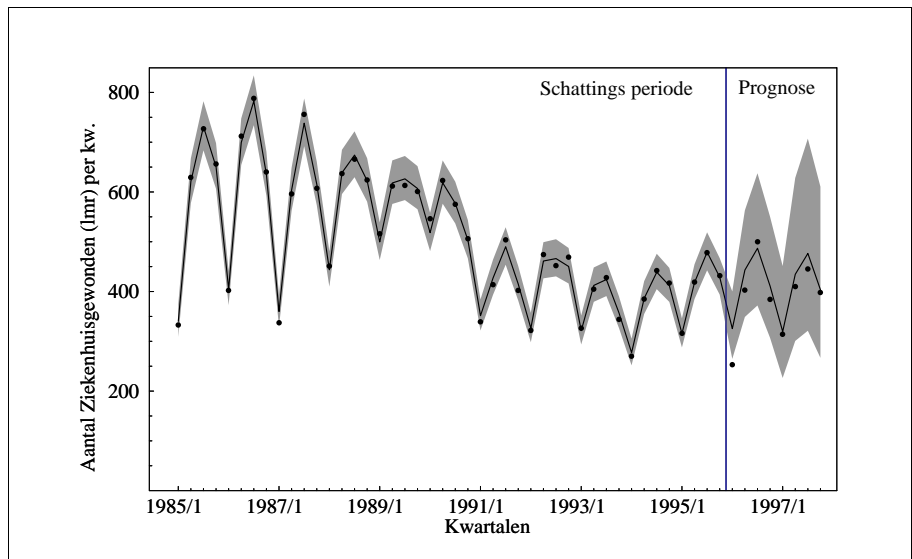
Afbeelding 19b. Schattingen en prognoses (1996 + 1997) verkeersdoden . 15- tot 24-jarige bromfietzers.



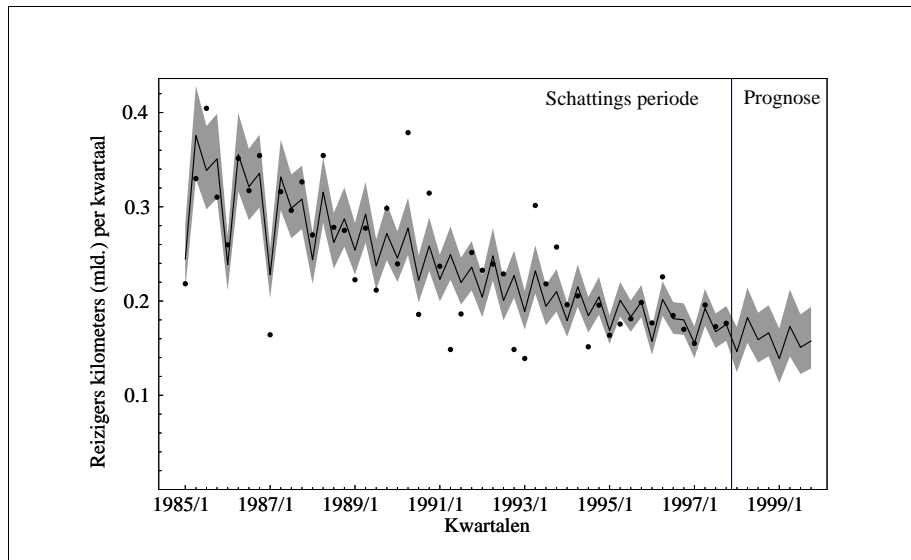
Afbeelding 19c. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden. 15- tot 24-jarige bromfietzers.



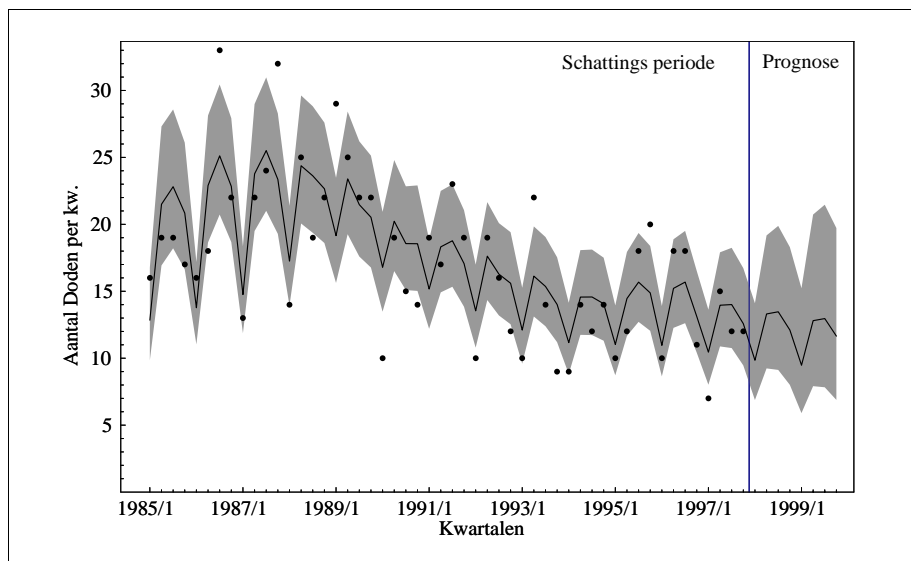
Afbeelding 20a. Schattingen en prognoses (1996 + 1997) doden gegeven mobiliteit. 15- tot 24-jarige bromfietzers.



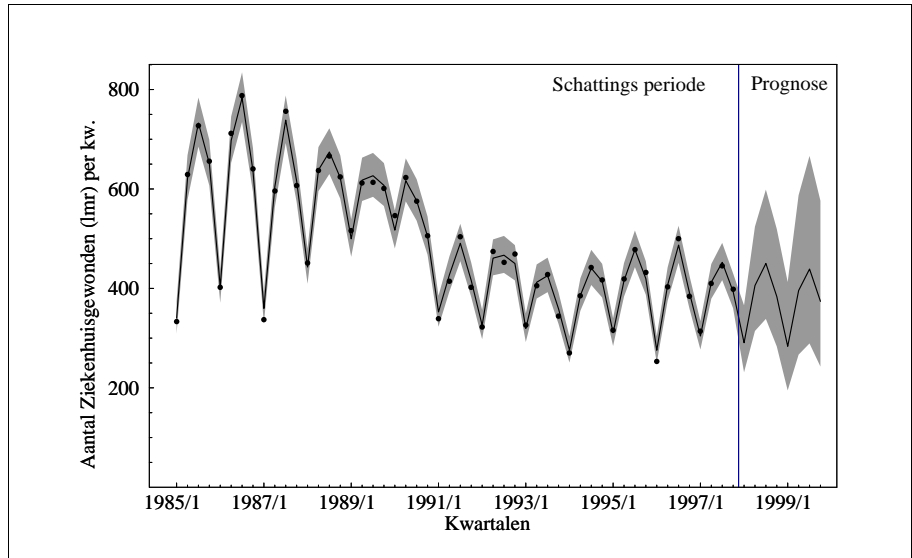
Afbeelding 20b. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden gegeven mobiliteit. 15- tot 24-jarige bromfietzers.



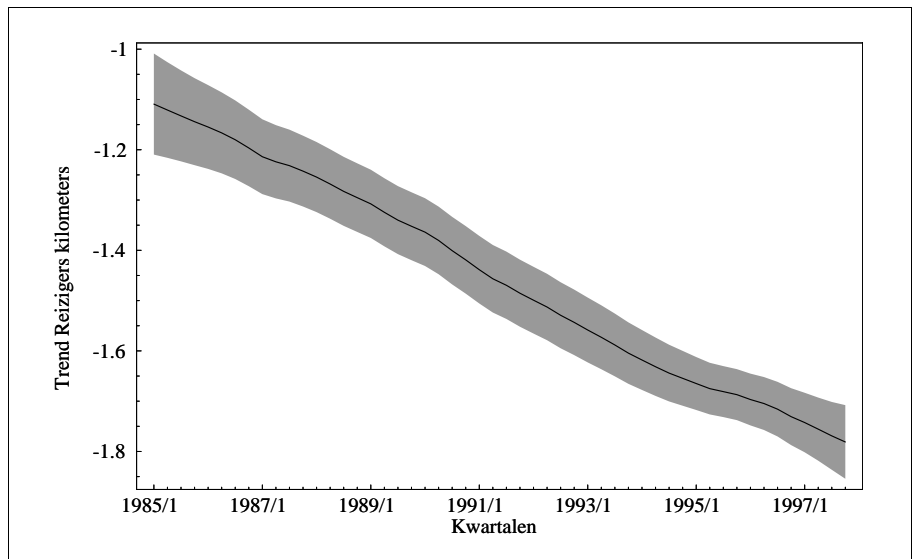
Afbeelding 21a. Schattingen en prognoses (1998 + 1999) reizigerskilometers. 15- tot 24-jarige bromfietzers.



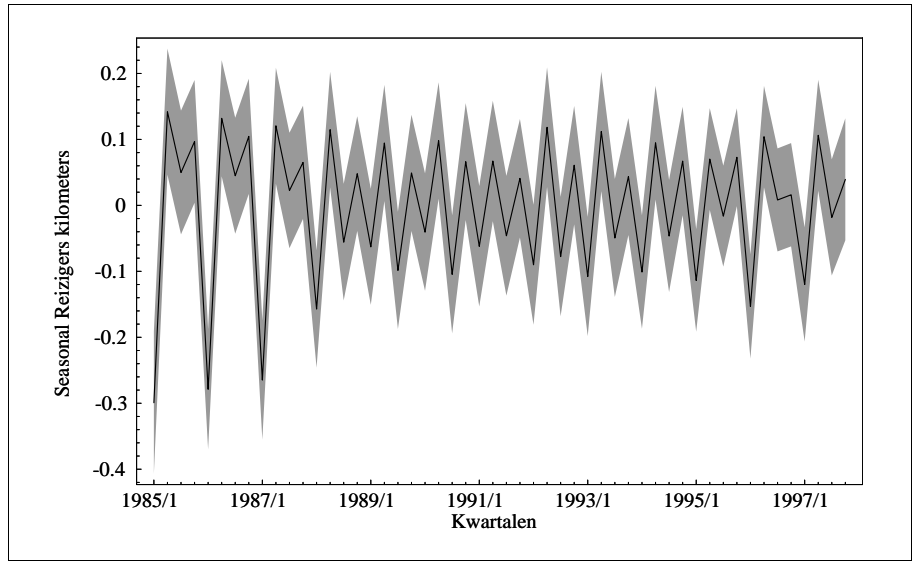
Afbeelding 21b. Schattingen en prognoses (1998 + 1999) verkeersdoden. 15- tot 24-jarige bromfietzers.



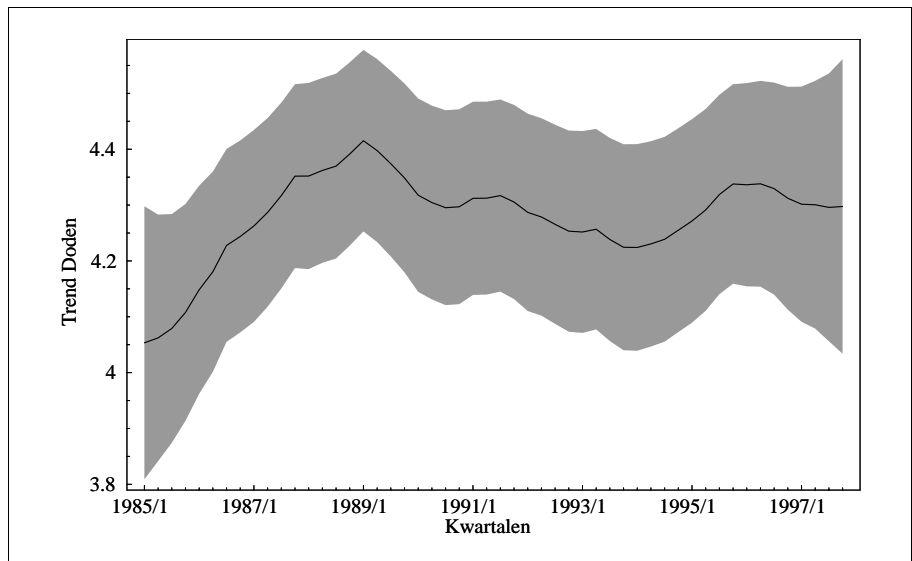
Afbeelding 21c. Schattingen en prognoses (1998 + 1999) ziekenhuisgewonden. 15- tot 24-jarige bromfietzers.



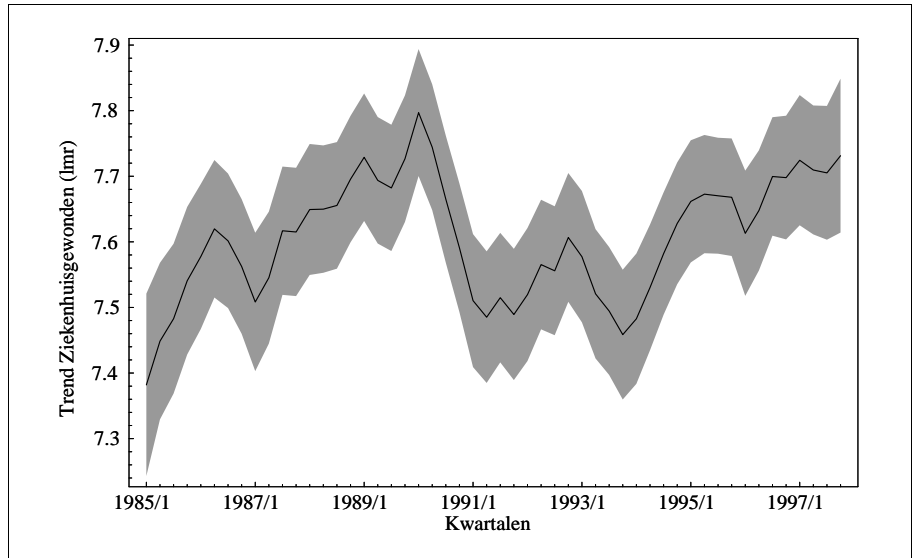
Afbeelding 22a. Trend reizigerskilometers. 15- tot 24-jarige bromfietzers.



Afbeelding 22b. *Seizoenseffect reizigerskilometers. 15- tot 24-jarige bromfietzers.*

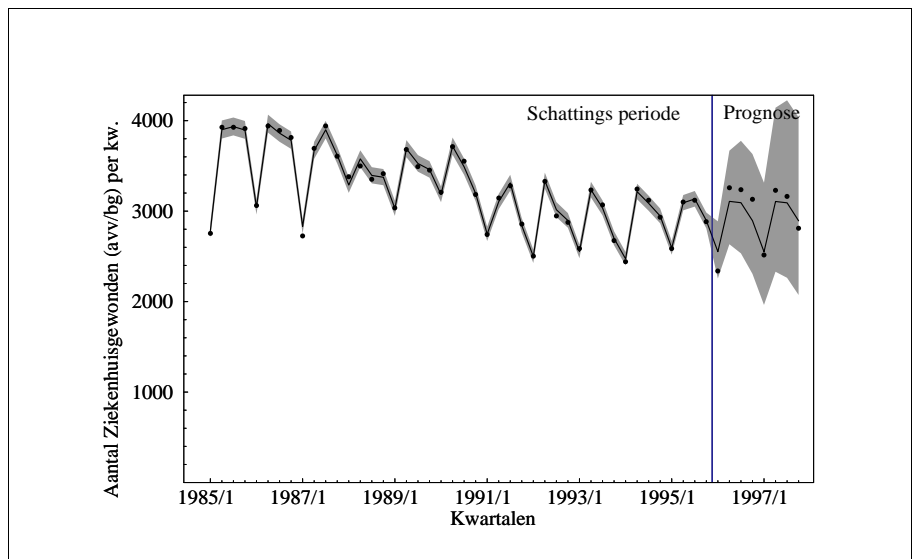


Afbeelding 22c. *Trend verkeersdoden. 15- tot 24-jarige bromfietzers.*

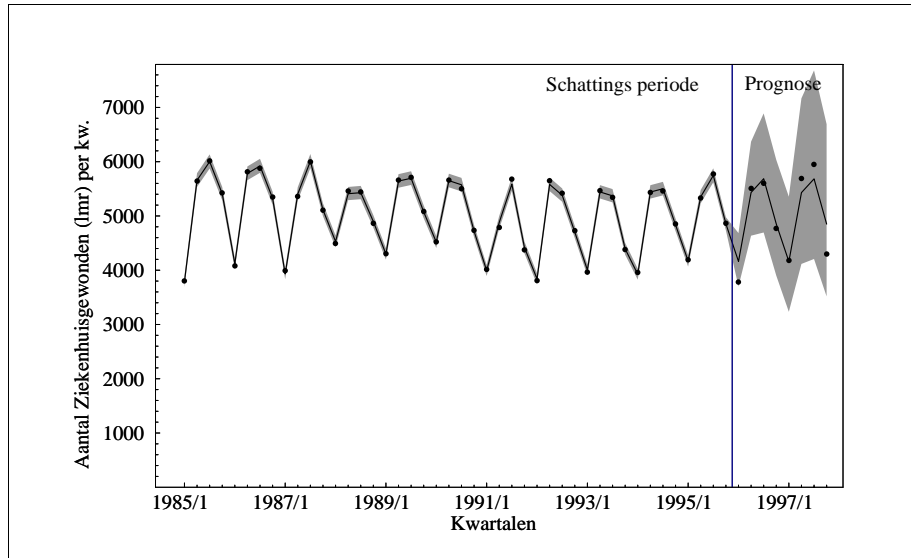


Afbeelding 22d. Trend ziekenhuisgewonden. 15- tot 24-jarige bromfietzers.

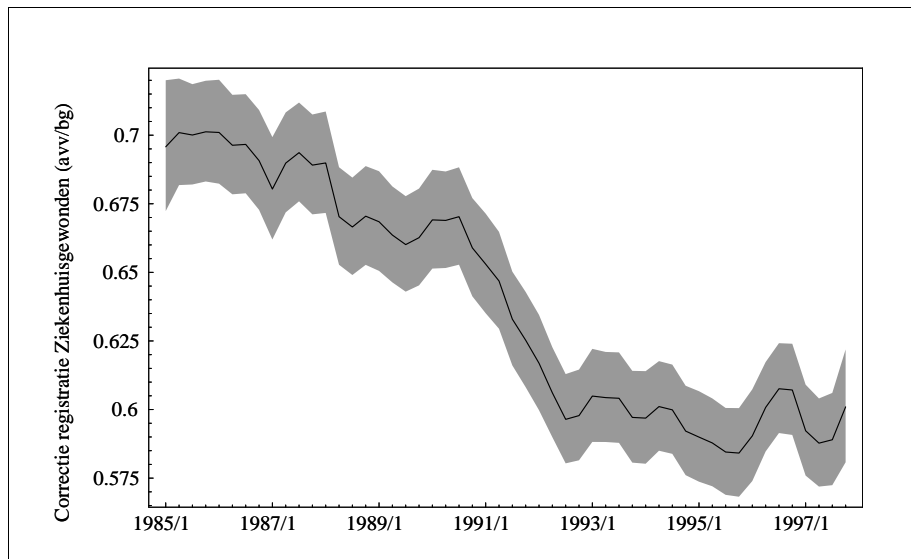
6. *Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.*



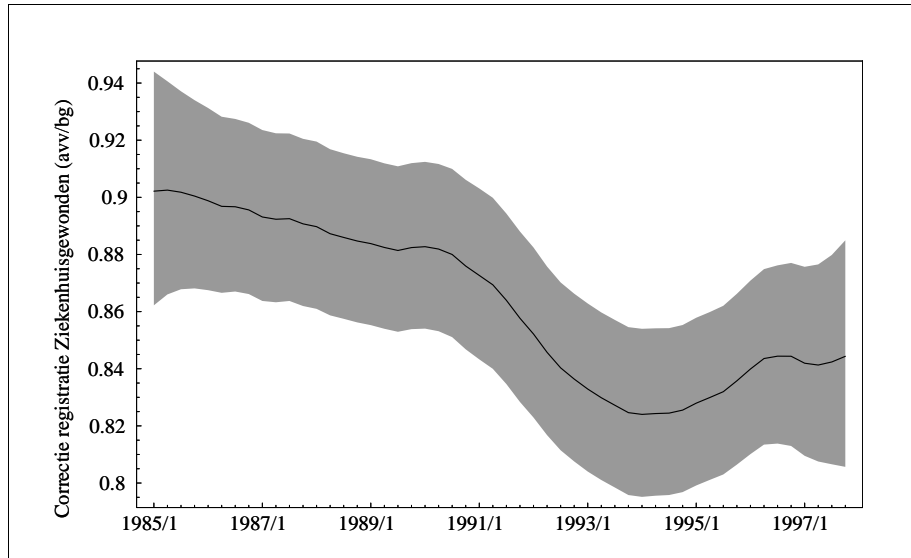
Afbeelding 23a. Schattingen en prognoses (1996 + 1997) AVV/BG-gewonden. Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.



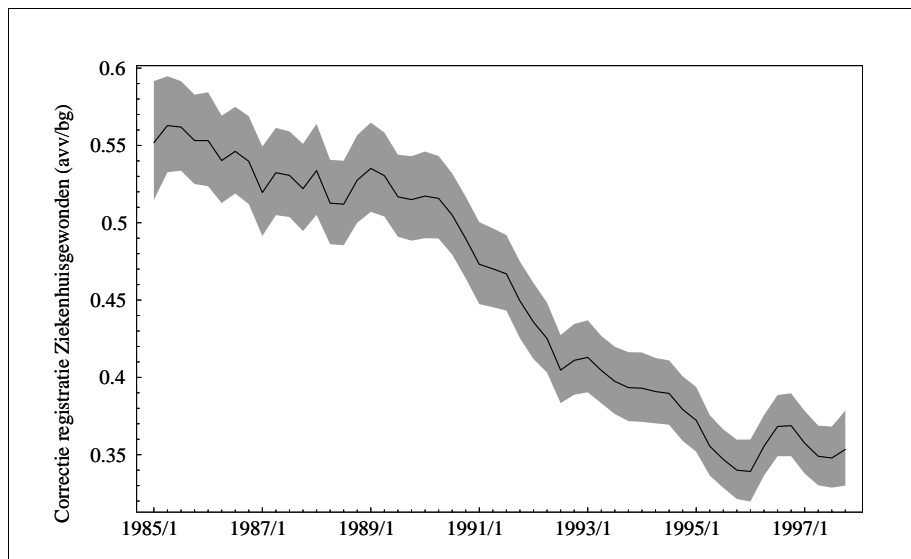
Afbeelding 23b. Schattingen en prognoses (1996 + 1997) LMR-gewonden. Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.



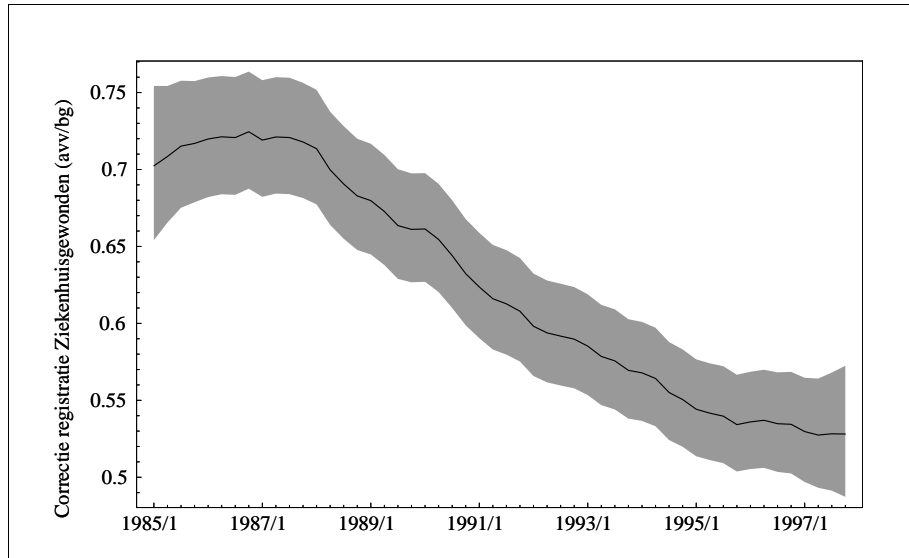
Afbeelding 23c. Verhouding AVV/BG-gewonden t.o.v. LMR-gewonden. Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.



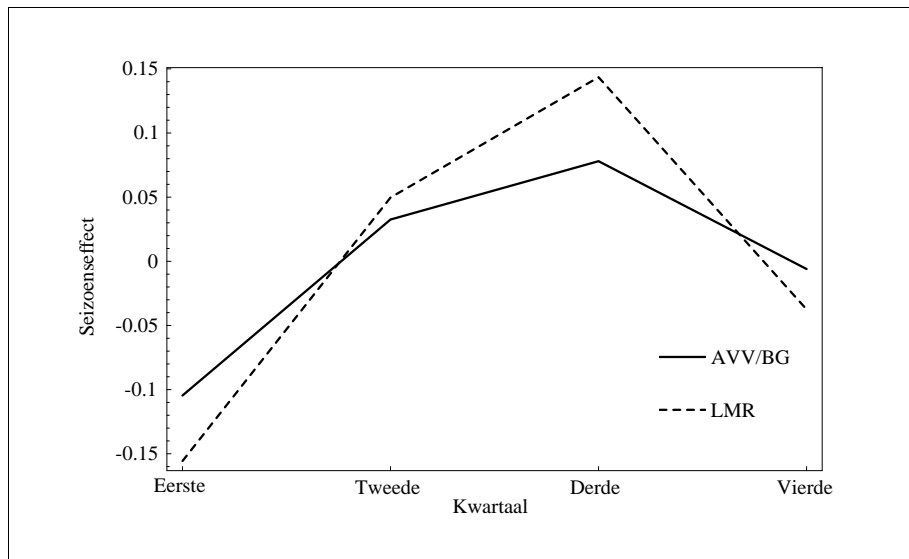
Afbeelding 23d. *Verhouding AVV/BG gewonden t.o.v. LMR-gewonden (bromfietsen). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.*



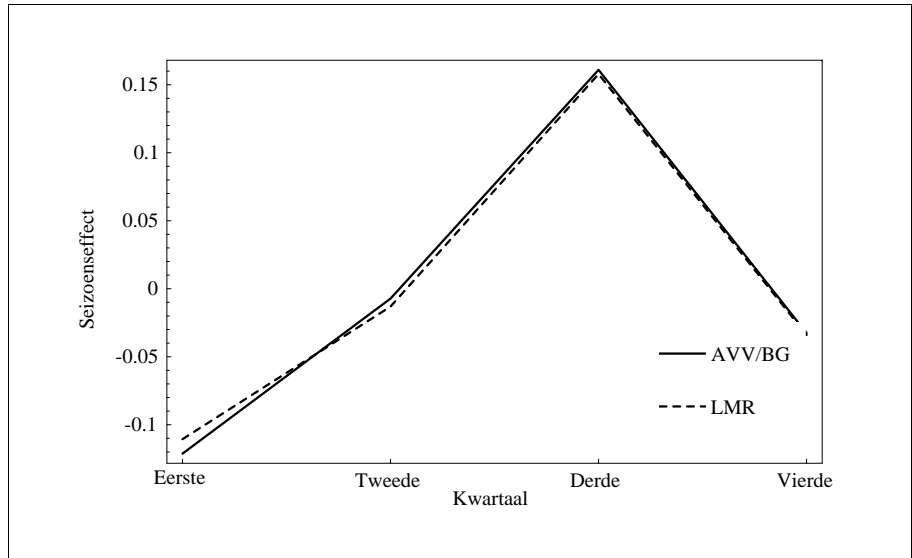
Afbeelding 23e. *Verhouding AVV/BG-gewonden t.o.v. LMR-gewonden (fietsen). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.*



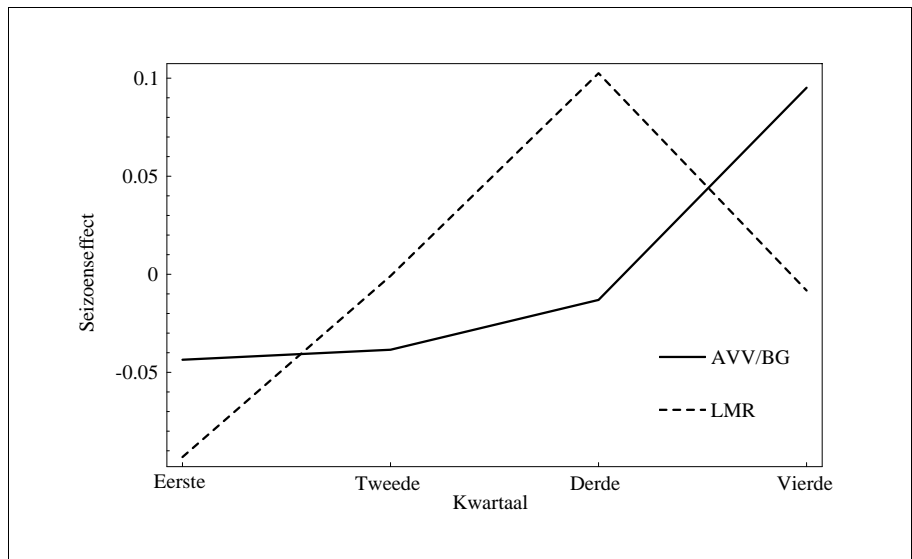
Afbeelding 23f. Verhouding AVV/BG-gewonden t.o.v. LMR-gewonden (lopen). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.



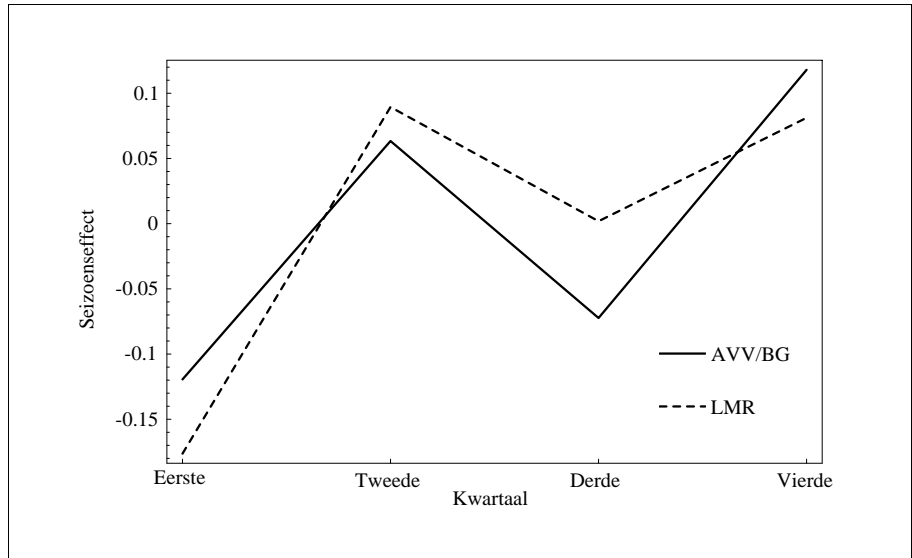
Afbeelding 23g. Seizoenseffecten AVV/BG- en LMR-gewonden (alle verkeer). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.



Afbeelding 23h. Seizoenseffecten AVV/BG- en LMR-gewonden (bromfiets). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.

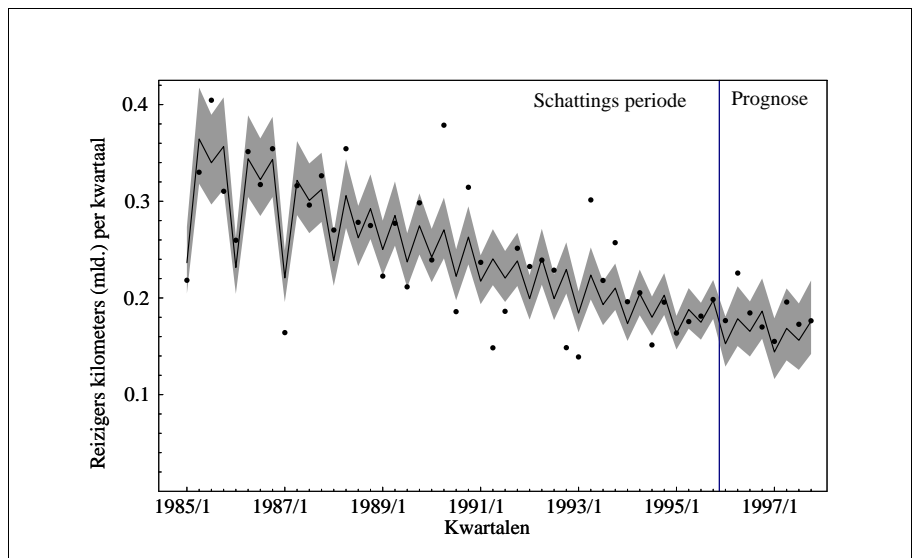


Afbeelding 23i. Seizoenseffecten AVV/BG- en LMR-gewonden (fiets). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.

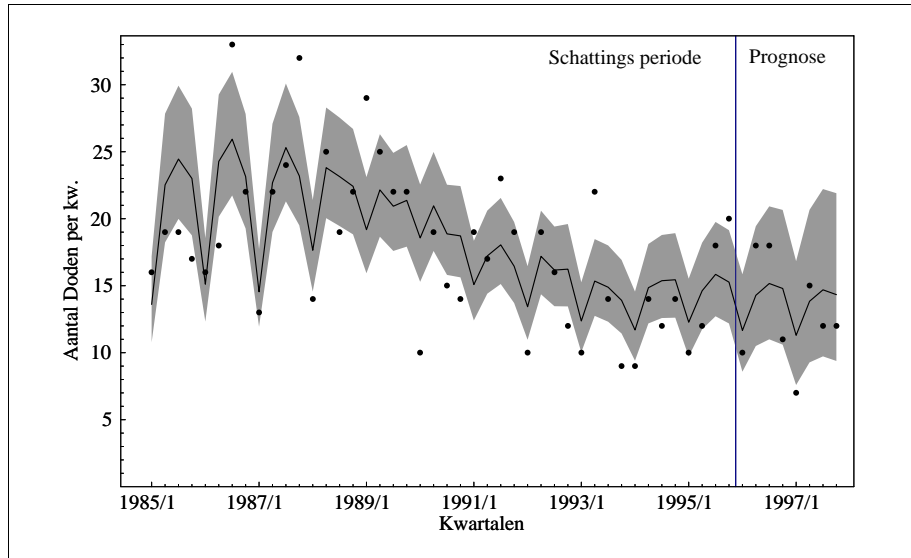


Afbeelding 23j. Seizoenseffecten AVV/BG- en LMR-gewonden (lopen). Alle leeftijden en alle wijzen van vervoer inclusief LMR-gewonden.

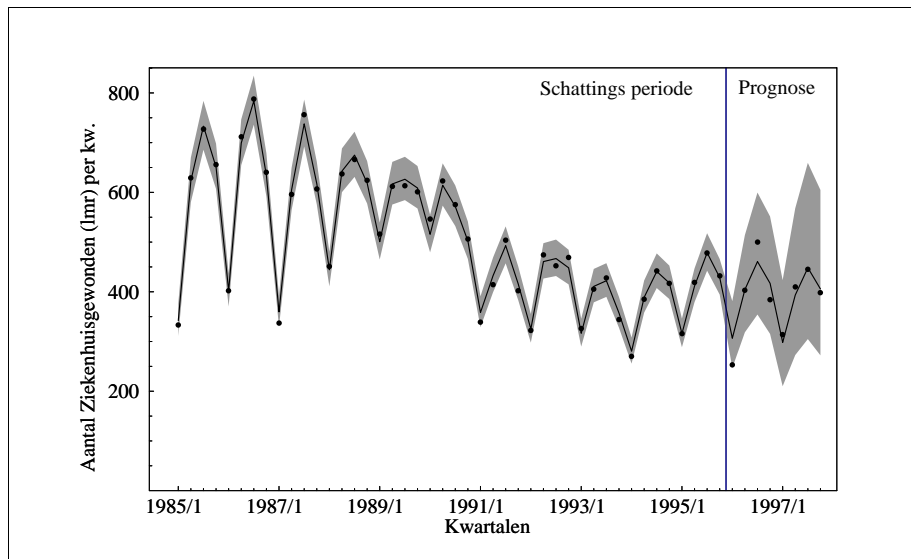
7. 15- tot 24-jarige bromfietzers met reizigerskilometers personenauto's



Afbeelding 24a. Schattingen en prognoses (1996 + 1997) mobiliteitscijfers. 15- tot 24-jarige bromfietzers met reizigerskilometers personenauto's.



Afbeelding 24b. Schattingen en prognoses (1996 + 1997) verkeersdoden. 15- tot 24-jarige bromfietzers met reizigerskilometers personenauto's.



Afbeelding 24c. Schattingen en prognoses (1996 + 1997) ziekenhuisgewonden. 15- tot 24-jarige bromfietzers met reizigerskilometers personenauto's.