

Ontwikkeling van nieuwe modellen ten behoeve van verkeersveiligheidsprognoses

Nieuwe methoden en vergelijking verschillende landen

R-95-73

Drs. F.D. Bijleveld

Leidschendam, 1995

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV

Documentbeschrijving

Rapportnummer: R-95-73
Titel: Ontwikkeling van nieuwe modellen ten behoeve van verkeersveiligheidsprognoses
Ondertitel: Nieuwe methoden en vergelijking verschillende landen
Auteur(s): Drs. F.D. Bijleveld
Onderzoeksmanager: Drs. S. Oppe
Projectnummer SWOV: 74.204
Opdrachtgever: Het onderzoek waarvan dit rapport verslag doet, werd uitgevoerd in het kader van de jaarlijkse doelsubsidie van het Ministerie van Verkeer en Waterstaat aan de SWOV.

Trefwoorden: Traffic, safety, forecast, mathematical model, development, international.
Projectinhoud: Dit rapport behelst een gedetailleerd vergelijkend onderzoek naar verschillen en overeenkomsten tussen de alternatieve risico-modellen en prognose-modellen voor een aantal ontwikkelde landen, ten behoeve van verkeersveiligheidsprognoses.

Aantal pagina's: 116
Prijs: f 35,-
Uitgave: SWOV, Leidschendam, 1995

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV



Stichting
Wetenschappelijk Postbus 1090
Onderzoek 2260 BB Leidschendam
Verkeersveiligheid Duindoorn 32
SWOV telefoon 070-3209323
 telefax 070-3201261

Samenvatting

Dit rapport behelst een gedetailleerd vergelijkend onderzoek naar verschillen en overeenkomsten tussen de alternatieve risico-modellen en prognose-modellen voor een aantal ontwikkelde landen. Er wordt een uiteenzetting gegeven van de gebruikte modellen, met een beknopte achtergrond-beschrijving. Hierbij wordt gepoogd uit gegevens als verkeersprestatie en tijdsontwikkeling, het aantal doden per jaar te verklaren. Er worden geen voorspellingen gedaan. Ook wordt geen moeite gedaan een inhoudelijke verklaring te geven voor de gevonden resultaten.

Een van de doelstellingen is het opzetten van een systeem waarmee het betrekkelijk eenvoudig is om verschillende alternatieve modellen met elkaar te vergelijken. Van deze modellen worden vervolgens een aantal statistische grootheden afgeleid, die op een systematische manier met elkaar vergeleken kunnen worden. Er wordt een nauwkeurige beschrijving van deze grootheden gegeven.

Voor vijf landen (Nederland, Groot-Brittanië, het voormalige West-Duitsland, Japan en de Verenigde Staten) worden zes modellen ontwikkeld en vergeleken. Er blijkt geen algemeen optimaal model voor handen te zijn voor deze landen, gekozen uit de zes gebruikte modellen. Waarschijnlijk zal de klasse gebruikte modellen moeten worden uitgebreid.

Voor Nederland blijkt een logistisch model aanmerkelijk beter te voldoen dan het in BIS-V gebruikte exponentiële model, zodat voorgesteld kan worden dit logistisch model in het vervolg te gebruiken.

Summary

This report contains a detailed comparative study into differences and similarities between alternative risk and prognostic models for a number of developed countries. It is attempted to explain the number of fatalities by both time development and traffic volume as a measure of exposure. No predictions for the number of fatalities are made. Also no effort is made to explain the theoretical basis for the respective models.

One of the goals of this research was to establish a framework in which it is relatively easy to estimate models and to compare those models with alternatives. Once models are estimated, a number of statistical measures are derived that can be compared. A detailed description is given of the selected statistical measures.

Five countries, the Netherlands, Great-Britain, the former West-Germany, Japan and the United States are considered. Based on those countries, six models are evaluated and compared. None of the models seems globally acceptable. Probably, the set of models should be extended.

Restricting to the situation in the Netherlands, the logistic model seems to fit best of the six considered. It fits substantially better than the exponential model used in BIS-V so far. It is suggested to replace the latter model with the logistic model.

Inhoud

1.	<i>Inleiding</i>	7
1.1.	Doelstelling en opzet nieuw rapport	7
1.2.	Beknopte geschiedenis van de ontwikkeling van macromodellen	7
1.3.	Korte beschrijving van de procedure zoals toegepast in BIS-V	9
2.	<i>Gebruikte Modellen</i>	12
2.1.	Risico-modellen	12
2.2.	Verkeersprestatie-modellen	17
2.3.	Samenvatting gebruikte modellen	17
3.	<i>Statistische aspecten</i>	18
3.1.	Niet-lineaire regressie	18
3.2.	Diagnostische analyses	22
3.3.	Vergelijking van modellen	27
4.	<i>Resultaten</i>	30
4.1.	Overzicht	30
4.2.	Primaire resultaten	31
4.3.	Van gewogen analyse afgeleide resultaten	38
4.4.	Vergelijking landen	40
4.5.	Golf-analyse	41
5.	<i>Conclusies</i>	45
	<i>Lijst van tabellen</i>	49
	<i>Lijst van figuren</i>	51
	<i>Bijlage</i>	
A.	<i>Data</i>	53
A.1.	West-Duitsland	53
A.2.	Engeland	55
A.3.	Nederland	57
A.4.	Verenigde Staten	59
A.5.	Frankrijk	61
A.6.	Israël	63
A.7.	Japan	65
A.8.	Michigan	67
B.	<i>Geschatte parameterwaarden</i>	69
B.1.	Gewone kleinste kwadraten	69
B.2.	Gewogen kleinste kwadraten	71
B.3.	Maximum likelihood	73
C.	<i>Afbeeldingen modellen voor doden</i>	75
C.1.	Doden West-Duitsland	75
C.2.	Doden Engeland	78
C.3.	Doden Japan	81
C.4.	Doden Nederland	84

C.5.	Doden Verenigde Staten	87
D.	<i>Afbeeldingen modellen voor Risico's</i>	90
D.1.	Risico's West-Duitsland	90
D.2.	Risico's Engeland	93
D.3.	Risico's Japan	96
D.4.	Risico's Nederland	99
D.5.	Risico's Verenigde Staten	102
E.	<i>Voorlopige vergelijking van de Amerikaanse staat Michigan met de Verenigde Staten als geheel</i>	105
E.1.	Inleiding	105
E.2.	Resultaten	105

1. Inleiding

1.1. Doelstelling en opzet nieuw rapport

Dit rapport behelst een gedetailleerd vergelijkend onderzoek naar verschillen en overeenkomsten tussen de alternatieve risico-modellen voor een aantal landen. Er wordt geen poging ondernomen een verklaring te geven voor het waarom van verloop van de reeksen; er wordt slechts onderzoek gedaan naar een aantal verschillende mathematische beschrijvingen voor de reeksen.

De praktische doelstelling is het totstandbrengen van een systeem waarmee de verschillende modellen met elkaar vergeleken kunnen worden. Daarbij zal ten eerste een maat voor de 'fit' van een model opgesteld moeten worden waarmee, als bepaalde veronderstellingen juist zijn, een maat voor de overeenkomst tussen de data en het model verkregen kan worden. Aan de hand van een dergelijke maat kan een 'optimaal' model geselecteerd worden. Daarnaast zullen criteria opgesteld moeten worden aan de hand waarvan gecontroleerd kan worden of aan de bovengenoemde veronderstellingen wordt voldaan. Een toepassing van de resultaten is verder om te komen tot een keuze van een model voor toepassing op Nederlandse gegevens, met name in verband met het Beleids Informatie Systeem Verkeersveiligheid (BIS-V). In deze zin is dit rapport als een vervolg op Bijleveld & Oppe (1992) te zien en als ondersteuning van een toekomstig rapport over de aanpassingen in BIS-V.

1.2. Beknopte geschiedenis van de ontwikkeling van macromodellen

Sinds mobiliteits- en verkeersonveiligheidsgegevens op systematische wijze worden geregistreerd, is belangstelling ontstaan voor de mogelijkheid om aan de hand van deze gegevens prognoses te maken voor de mobiliteit en de onveiligheid in het verkeer. Een vroeg voorbeeld hiervan is Smeed (1949). In dat artikel wordt aan de hand van gegevens over een periode van 40 jaar de verkeers- onveiligheid in verschillende landen met elkaar vergeleken. De resultaten van dit onderzoek hebben in zekere mate geleid tot de acceptatie van 'de wet van Smeed'. Een belangrijk verschijnsel bij het analyseren van ontwikkelingen in de tijd met behulp van modellen is dat na verloop van tijd telkens weer blijkt dat de gebruikte modellen aangepast dienen te worden. Zo is het ook met de wet van Smeed gegaan. Dit verschijnsel leidde tot een voor dit rapport belangrijke mijlpaal in het artikel van Appel (1982) en het onafhankelijk daarvan opgestelde rapport van Blokpoel (1982). In deze stukken is aannemelijk gemaakt dat de ontwikkeling van het aantal doden (f_t) op een eenvoudige manier ontbonden kon worden in een verkeersprestatie-component (v_t) en een risico-component (r_t), zodat:

$$f_t = v_t \times r_t. \quad (1.1)$$

In Appel (1982) en Blokpoel (1982) werd aangenomen dat voor v_t en r_t van lineaire trends uitgegaan kan worden. Dit komt neer op een *kwadratisch* model voor het aantal doden. Blokpoel (1982) voegt hier wel aan toe dat een lineaire trend in zou houden dat aan het einde van de jaren tachtig het aantal doden het nulpunt zou moeten bereiken. Dit werd destijds al niet aannemelijk geacht.

Het fundamentele en belangrijkste resultaat dat uit Appel (1982) en Blokpoel

(1982) gedistilleerd kan worden is de ontbinding van het aantal doden in twee tijdreeksen. Deze tijdreeksen zijn ieder voor zich eenvoudiger van structuur. Prognoses voor het aantal doden kunnen vervolgens berekend worden door prognoses voor beide tijdreeksen te berekenen en deze te vermenigvuldigen. Dit is bijvoorbeeld de werkwijze in Brüning et al. (1986) geweest en is tot op heden nog steeds in gebruik (bijvoorbeeld Bijleveld & Oppe, 1992). Wel zijn een aantal aanpassingen en daarop weer nieuwe aanpassingen voorgesteld op de bovenstaande werkwijze. Het merendeel van deze aanpassingen houden variaties in op de modellen voor zowel de mobiliteits- als de risico-ontwikkeling. Opvallend is dat ze bijna alle (1.1) intact blijven. Hierin zal zeer waarschijnlijk verandering komen in de toekomst.

De belangrijkste aanpassing die tot nog toe is aangebracht, is het toepassen van een *exponentiële* trend voor het risico in plaats van de in Appel (1982) en Blokpoel (1982) gebruikte lineaire trend. Ons is niet bekend wie deze aanpassing het eerste heeft toegepast. Deze aanpassing is gebruikt in drie varianten:

1. Risico naderend naar nul.
2. Risico naderend naar een getal groter dan nul.
3. Risico naderend naar een getal dat mogelijk *kleiner* is dan nul.

Het laatste model kan leiden tot dezelfde problemen als Blokpoel (1982) aangeeft. De eerste variant, waarbij het risico naar nul nadert, lijkt het meest gebruikt te worden.

Behalve vele aanpassingen op het risico zijn ook verschillende varianten voor de mobiliteitsontwikkeling voorgesteld, vaak in combinatie met risico-modellen. Een belangrijke overweging blijkt telkens weer de eenvoud van de gebruikte modellen te zijn. Een belangrijk gebleken aanname voor de mobiliteitsontwikkeling is de aanname dat de mobiliteit tot een soort van verzadigingspunt zal stijgen. Praktisch betekende dit dat de mobiliteitsontwikkeling door een S-vormige kromme beschreven kon worden. Een veel gebruikte S-kromme is de *logistische* kromme geweest. Tanner (1958) gebruikte deze kromme reeds voor *aantallen* auto's op de weg. Deze kromme is in combinatie met een exponentieel model voor het risico te vinden in Oppe et al. (1988). In reactie hierop is dit model verder gegeneraliseerd door Koomstra (1988). In dit rapport dat enigszins voortborduurde op het gedeelte van Koomstra in Oppe et al. (1988) worden ook andere dan de logistische aanname bestudeerd. Ook wordt een poging ondernomen om tot een inhoudelijke onderbouwing van de theorie te komen. Daarbij wordt ook van (1.1) afgeweken door de mogelijkheid te opperen dat er sprake kan zijn van een machtsverheffing van de mobiliteit.

$$f_t = v_t^q \times r_t, \quad \text{voor } q > 0$$

In Oppe (1989) worden modellen voor een aantal landen met elkaar vergeleken. Deze landen zijn Nederland, West-Duitsland, Groot-Britannië en de Verenigde Staten. Er lijkt een lineair verband te bestaan tussen gefitte coëfficiënten van mobiliteitsmodellen enerzijds en risico-modellen anderzijds. Dit resultaat suggereert een algemeen model voor de ontwikkeling van de verkeersveiligheid. Ook lijkt het een ordening van de individuele landen te impliceren. Een dergelijk resultaat heeft nieuw onderzoek gestimuleerd, onder meer naar de relaties tussen beide ontwikkelingen. Het model is uitgebreid

met twee landen (Japan en Israël) en enkele analyses in Oppe & Koornstra (1990), Oppe (1991b). Ook is een theoretische onderbouwing gegeven.

In Oppe (1991a) wordt melding gemaakt van systematische fluctuaties van afwijkingen van de diverse modellen. Er blijkt een samenhang te zijn tussen de afwijkingen in het mobiliteitsmodel en de afwijkingen in het risico-model. Dit is onderzocht met behulp van een zogenaamde *polynoom*-analyse. Op zich is het fitten van hogere-orde- polynomen ((Oppe, 1991a) gaat tot de 11^e graad) een kwetsbare methode, maar als geheel blijkt het als descriptieve methode toch een exploratieve waarde te hebben. Een belangrijk resultaat is de constatering van een min of meer structurele periodieke afwijking van de modellen. Koornstra (1992) gebruikt onder meer de risico-adaptatie- theorie om dit aspect te verklaren.

In tegenstelling tot de bovenstaande literatuur beoogd het onderhavige rapport niet tot een verklaring van de ontwikkeling van de verkeersveiligheid te komen en al helemaal niet een relatie te leggen tussen deze ontwikkeling en de mobiliteitsontwikkeling. Dit onderzoek probeert met behulp van een redelijk geavanceerde theorie van niet-lineaire regressie-modellen een analyse te geven van de bovengenoemde modellen. Daarbij is een belangrijk oogmerk de toepassing binnen prognosemodellen, zowel voor de korte als langere termijn. Dit gegeven beïnvloedt de keuze en uitwerking van de gebruikte technieken.

1.3. Korte beschrijving van de procedure zoals toegepast in BIS-V

In Bijleveld & Oppe (1992) is ten behoeve van het Beleids Informatie Systeem Verkeersveiligheid (BIS-V) een prognose-model voor de verkeersonveiligheid ontwikkeld.

Het model in Bijleveld & Oppe (1992) is ontwikkeld om prognoses voor verkeersonveiligheid te maken, zowel in termen van het aantal doden als het aantal ziekenhuisgewonden per jaar. Hiervoor zijn zowel modellen voor de risico-ontwikkeling opgesteld als modellen voor de mobiliteitsontwikkeling. Deze modellen zijn zelfs gedisaggregeerd naar leeftijd en wijze van vervoer.

Ten opzichte van de uitgangspunten in Bijleveld & Oppe (1992) zijn er in dit stuk een aantal veranderingen aangebracht. Zo is de doelstelling enigszins veranderd. Er wordt nu ook een vergelijking van analyse-resultaten over de verkeers(onveiligheids)gegevens tussen verschillende ontwikkelde landen als doel gesteld.

Het door Bijleveld & Oppe (1992) ontwikkelde model bleek een aantal praktische bezwaren te hebben. Ook bleken er mogelijkheden te bestaan tot verbetering van zowel de gebruikte modellen als de toegepaste technieken. Bovendien bestaat de reeds vermelde behoefte aan een vergelijking van dergelijke modellen met gelijksoortige modellen voor andere landen. Het onderhavige stuk bevat een verslag van de activiteiten die zijn uitgevoerd om bovenstaande aanpassingen en toevoegingen aan te brengen aan het in Bijleveld & Oppe (1992) beschreven model.

Aan de activiteiten zijn een aantal beperkingen opgelegd:

- In eerste instantie is gekozen voor het analyseren van gegevens uit
 - Nederland;

- West-Duitsland;
- Engeland;
- Verenigde Staten;
- Japan.

Behalve deze landen zijn ook (beperkte) gegevens beschikbaar over:

- Frankrijk;
- Israël.

Mede om de hoeveelheid gegevens te beperken zijn deze laatste landen niet in deze rapportage verwerkt.

- Van bovenstaande landen zijn telkens voor een zo lang mogelijke periode het aantal overleden slachtoffers en een maat voor de mobiliteit verzameld. De precieze aard van deze gegevens zal niet voor ieder land hetzelfde zijn. Er moet bijvoorbeeld rekening worden gehouden met mogelijke verschillen in de definitie van verkeersdoden. Ook kan het mobiliteitscijfer verschillen. Dit kan zowel qua registratie-methode (en daaruit volgende consequenties) als qua geregistreerde eenheid (voertuigkilometers of -mijlen dan wel personenkilometers). De uiteindelijk gebruikte gegevens staan in *Bijlage A*.
- Alleen het totaal aantal doden is geanalyseerd. Er zijn geen disaggregaties gemaakt. In Bijleveld & Oppe (1992) is dat wel gedaan. Daar zijn ook gegevens over gewonden geanalyseerd.
- Er zijn geen assumpties over de ontwikkeling van de mobiliteit in de diverse landen gemaakt. Praktisch houdt dit in:
 - Dat er geen, zoals in Bijleveld & Oppe (1992) gebruikte, gesmoothde mobiliteitsgegevens gebruikt kunnen worden. Deze beperking heeft overigens voor totalen geen grote gevolgen. Wel wordt het onderling vergelijken helderder door deze vereenvoudiging. In de plaats van gesmoothde mobiliteitscijfers zijn hier de geobserveerde cijfers gebruikt, zoals dat meestal gebeurt. De relevante afbeeldingen zijn in *Bijlage A* weergegeven.
 - Dat er geen prognoses voor de mobiliteit in de diverse landen beschikbaar is. Als gevolg hiervan kunnen dus ook geen prognoses voor het aantal slachtoffers gemaakt worden.

Behalve de boven opgesomde beperkingen zijn er ook uitbreidingen aangebracht ten opzichte van Bijleveld & Oppe (1992):

- Reeds genoemd: er zijn meer landen geanalyseerd.
- Er is een groep alternatieve modellen met elkaar vergeleken. Het in Bijleveld & Oppe (1992) gebruikte model voor de risico-ontwikkeling was een exponentieel model. (Voor de verkeersprestatie werd overigens een *logistisch* model gebruikt). In Bijleveld & Oppe (1992) werd aangenomen dat het risico exponentieel zou *dalen* naar het nul- niveau.
- Behalve deze aanname in Bijleveld & Oppe (1992) is bovendien verondersteld dat het te verwachten aantal slachtoffers (doden) het produkt is van de verkeersprestatie en het risico in een bepaald jaar. Zowel de precieze keuze van het risico-model (exponentieel) als de aanname van de lineaire afhankelijkheid van het aantal doden van de verkeersprestatie (mobiliteit) zijn in het onderhavige stuk losgelaten. De oplossing is vergeleken met een aantal alternatieven. Deze alternatieven worden in *Hoofdstuk 2* behandeld.
- De statistische methodiek in Bijleveld & Oppe (1992) is gebaseerd op de mogelijkheden die de procedure PROC NLIN in de module STAT van het statistisch pakket SAS biedt. Voor dit project is verder gebruik

gemaakt van het pakket Mathematica, waardoor een ruime hoeveelheid specialistische toetsingsmogelijkheden ter beschikking is gekomen. Dit heeft onder meer geleid tot de mogelijkheid verschillende wijzen 'slinger'- effecten te onderzoeken. Ook bleek een onderzoek naar de betrouwbaarheid van de asymptotische parameterinformatie mogelijk. Deze onderwerpen komen in *Hoofdstuk 3* aan de orde.

2. Gebruikte Modellen

2.1. Risico-modellen

2.1.1. Inleiding

De meeste modellen voor de verkeersonveiligheid zijn toegepast op het jaarlijks aantal doden, als maat voor de onveiligheid en van de hoeveelheid verkeer en de omstandigheden waarin dat verkeer zich beweegt. De hoeveelheid verkeer noemt men ook wel de expositie, waarmee blootstelling aan gevaar bedoeld wordt. Deze expositie kan worden uitgedrukt in het totaal van verplaatsingen, verplaatsingskilometers of gebeurtenissen in het verkeer waarbij een kans op ongevallen bestaat. De kans dat hieruit ook werkelijk een ongeval komt en erger, wellicht doden vallen zou men als wordt verkeersrisico genoemd. Dit risico hangt van vele omstandigheden af. Een aantal van deze omstandigheden zullen met de tijd veranderen. In principe lijkt het redelijk aan te nemen dat het risico afneemt als gevolg van betere kwaliteit van de verkeersinfrastructuur, vervoermiddelen en meer aangepast verkeersgedrag, dus vooral ook doordat de samenleving als geheel zich beter instelt op de gevaren van het deelnemen aan het verkeer.

Het te verwachten aantal doden hangt dus zowel van de hoeveelheid verkeer als van de omstandigheden af. Een natuurkundig gas-model voor het botsen van moleculen is gelukkig te simplistisch voor het verkeersonveiligheidsmodel, toch zal het aantal ongevallen of slachtoffers op de één of andere manier toenemen bij een toename van de hoeveelheid verkeer. Meestal wordt het risico uitgedrukt in het aantal slachtoffers (bijvoorbeeld doden) per expositie-eenheid (bijvoorbeeld personenkilometers). Daarbij wordt bijvoorbeeld verondersteld dat het te verwachten aantal doden gelijk is aan de hoeveelheid personenkilometers vermenigvuldigd met het risico per personenkilometer. Dit is ook de gedachte in Bijleveld & Oppe (1992). Terugkomende op het gas-model; een complicatie van deze benadering is dat het aantal botsingen van bijvoorbeeld zuurstof-moleculen met bijvoorbeeld waterstof-moleculen van zowel de concentratie zuurstof als de concentratie waterstof af zal hangen. Het aantal waterstof-moleculen dat botst hangt dus af van zowel de hoeveelheid waterstof (essentieel) als van de hoeveelheid zuurstof.

In termen van verkeersveiligheid betekent dit (bijvoorbeeld) dat het aantal fietsers dat gedood wordt in het verkeer afhangt van de hoeveelheid fietsers waarmee gebotst kan worden, maar ook op een speciale manier, van de hoeveelheid andersoortige verkeersdeelnemers. In het geval van de personenauto als botspartner van de fietser geldt natuurlijk dat een groot deel van de mobiliteit van personenautos zich ver van het fietsverkeer afspeelt, zodat een complexe samenhang zal gelden.

Behalve dit meer gedetailleerde expositie-model, gedefinieerd in termen van ontmoetingen tussen verkeersdeelnemers als eenheid voor de expositie, is de expositie op een meer geaggregeerd niveau te beschrijven als functie van het gemotoriseerd verkeer: wat is de kans op een slachtoffer per kilometer, afgelegd met een gemotoriseerd voertuig? Eigenlijk zouden dan de slachtoffers waarbij geen gemotoriseerd verkeer is betrokken daarvan moeten worden afgetrokken. Om praktische redenen is hiervan afgezien.

2.1.2. De hier gebruikte definitie van het risico

In Bijleveld & Oppe (1992) is verondersteld dat het risico in de tijd een exponentieel dalende functie is, welke in het 'oneindige' het nul-niveau zou bereiken. In de huidige analyse is hiervan op een aantal punten afgeweken:

- Er wordt niet meer expliciet geëist dat een risico-functie *dalend* is. Dit is een minder vergaande verandering dan het lijkt want, het effect van deze aanname in Bijleveld & Oppe (1992) beperkte zich tot een aantal kleine sub-groepen. In Bijleveld & Oppe (1992) is voor deze groepen een gelijkblijvend risico verondersteld. De ontwikkelingen van het totaal aantal doden bleken vanzelf al een dalend risico te hebben.
- Er wordt niet meer geëist dat het risico het nul-niveau nadert. De voorspellingswaarde van de huidige modellen is niet van dien aard dat een zeer lange periode voorspeld kan worden. Wel kan geconcludeerd worden dat modellen waarvan het risico op een gegeven moment het nul-niveau passeert, waarschijnlijk enige tijd daarvoor al hun voorspellende waarde zullen hebben verloren. In *Tabel 16* zijn de jaartallen waarop de diverse modellen het nulniveau doorsnijden aangegeven.
- Er wordt niet meer geëist dat het risico een exponentieel verloop heeft. Voor sommige landen blijkt een logistische ontwikkeling beter te passen. Wellicht zijn er nog andere vormen van een risico-ontwikkeling mogelijk. In dit stuk zijn de exponentiële ontwikkeling (zie *Afbeelding 1* voor een voorbeeld) en de logistische ontwikkeling onderzocht (zie *Afbeelding 2* voor een voorbeeld).
- In een aantal gevallen wordt zelfs onderzocht of het 'risico' nog wel op de klassieke wijze gedefinieerd moet worden. Zie ook § 2.1.1. De klassieke benadering is $d_t = \frac{d_t}{v_t} \times v_t = r_t \times v_t$, waarbij v_t de verkeersprestatie is op het tijdstip t , r_t het risico op dat tijdstip en d_t het aantal doden. Er wordt dus verondersteld dat het risico uit te drukken is in een aantal doden (in dit geval) per mobiliteitseenheid. Voor deze mobiliteitseenheid kiest men meestal de gemotoriseerde mobiliteit, bijvoorbeeld personenautokilometers of personenauto-inzittendenkilometers.

Hoewel het theoretisch de voorkeur verdient om de gekozen exponentieel-grootheid (welke dit ook mag zijn) als basis te gebruiken in de analyse, is het om empirische redenen nuttig om na te gaan of een bij v_t gekozen exponent ongelijk aan één, niet beter kan worden voorspeld. Uiteindelijk kan de uitkomst bij een exponent $\neq 1$ weer herschreven worden als een (meer gecompliceerde) formule met betrekking tot v_t in plaats van v_t^ϵ .

Om deze redenen en omdat uit eerder onderzoek van Oppe & Koomstra (1990) is gebleken dat de verklaring van de modellen soms beter is als $(v_t)^\epsilon$ in plaats van v_t wordt gebruikt, is een dergelijke uitbreiding in aanmerking genomen. In het standaard geval wordt dus $\epsilon \equiv 1$ verondersteld. Voor alle risico-typen is onderzocht of $\epsilon \neq 1$ een verbetering oplevert.

2.1.3. De exponentiële ontwikkeling

De exponentiële ontwikkeling is nog steeds de meest veronderstelde ontwikkeling van het risico. In het huidige onderzoek zijn een aantal varianten

van deze ontwikkeling in aanmerking genomen. De algemene vorm van de exponentiële ontwikkeling is:

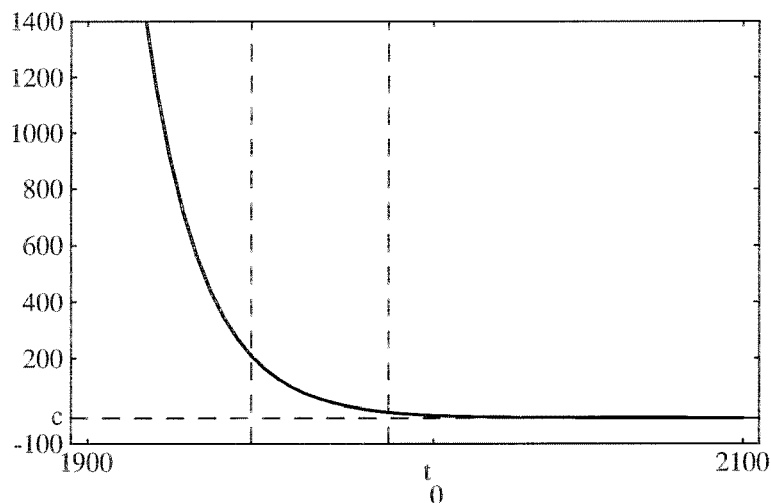
$$r_e(t) = c + e^{at+b} \quad (2.1)$$

Met parameters:

- a De grootte waarmee de risico verandering in de tijd wordt geregeld. In Bijleveld & Oppe (1992) is verondersteld dat a kleiner is dan nul, hetgeen dus een risicodaling impliceert.
- b Vervult de rol van het niveau op tijdstip $t = 0$, in samenhang met c bespeelt deze parameter de verschuiving in de tijd.
- c Vervult de rol van het laagste-niveau, het risico niveau dat in het 'oneindige' 'bereikt' wordt. In Bijleveld & Oppe (1992) is verondersteld dat $c = 0$. Uiteraard wordt geëist dat het risico groter of gelijk aan nul moet zijn. Dit betekent dat óf $c \geq 0$, óf de bruikbaarheid van het model is beperkt tot een tijdstip ruim voordat het nul-niveau gepasseerd wordt. In *Tabel 16* zijn de jaartallen waarop sommige modellen het nulniveau doorsnijden aangegeven.

Het grote voordeel van het exponentiële model is het beperkte aantal parameters dat geschat moet worden, tezamen met de toch redelijke voorspellingsresultaten. Zeker bij geaggregeerde modellen, als er slechts voor een klein aantal jaren observaties ter beschikking staan, bijvoorbeeld voor bijzondere categorieën waarvoor alleen OVG-gegevens beschikbaar zijn, zal het waarschijnlijk het beste zijn om dit model voor het risico te blijven kiezen. Dit is dan ook één van de redenen waarom dit model zo vaak gebruikt is, bijvoorbeeld in Bijleveld & Oppe (1992) en Bijleveld (1994).

In *Afbeelding 1* is het risico voor Nederland opgetekend. c is in dit geval kleiner dan nul. t_0 is het tijdstip waarop het risico gelijk is aan nul. Dit blijkt in de tweede helft van het jaar 2005 te gaan gebeuren. Dit jaar moet dus achter de voorspellings-'horizon' van dit model liggen.



Afbeelding 1. Voorbeeld exponentiële risico kromme

Gewogen regressie van Nederland-gegevens. Op tijdstip t_0 wordt het risico negatief. Dit gebeurt in het jaar 2005.

2.1.4. De logistische ontwikkeling

De klassieke vorm van de logistische kromme is:

$$f(t) = \frac{1}{1 + e^{at+b}} \quad (2.2)$$

Voor deze functie geldt:

- $\lim_{t \rightarrow \infty} f(t) = 1$.
- $\lim_{t \rightarrow -\infty} f(t) = 0$.
- $0 < f(t) < 1$ voor reële t .

Een belangrijk verschil met de exponentiële risico-kromme is dat de logistische kromme twee horizontale asymptoten heeft. De exponentiële kromme heeft er slechts één. De logistische kromme heeft zowel een boven als een ondergrens. Ook heeft de logistische kromme een zogenaamd buigpunt. Dit betekent dat, als het risico werkelijk volgens een logistische ontwikkeling daalt, het risico eerst steeds sterker gaat dalen om vanaf een zeker moment weer steeds minder snel te dalen. Bij de exponentiële risicodaling is de dalingsfactor constant over de tijd. De absolute daling wordt natuurlijk wel steeds kleiner.

Deze functie, als (2.2) geformuleerd, is op zich niet bruikbaar als risico-kromme. Daarvoor moet de functie eerst worden 'geschaald':

$$r_l(t) = (p - c)f(t) + c$$

waarbij $p > c$, zodat het resultaat $r_l(t)$ van boven begrensd is door de waarde p en aan de onderkant begrensd wordt door de waarde c , welke dezelfde rol speelt als de c in vergelijking 2.1. Zie ook *Afbeelding 2*. Het resultaat levert (2.3).

$$r_l(t) = p + \frac{c}{1 + e^{at+b}} - \frac{p}{1 + e^{at+b}} \quad (2.3)$$

Volgens de bovenstaande definitie gaat $r_l(t)$ over in een exponentieel model als $p \rightarrow \infty$. Dit is een praktisch voordeel dat numeriek echter niet praktisch is. Daarom is gekozen voor een iets andere formulering:

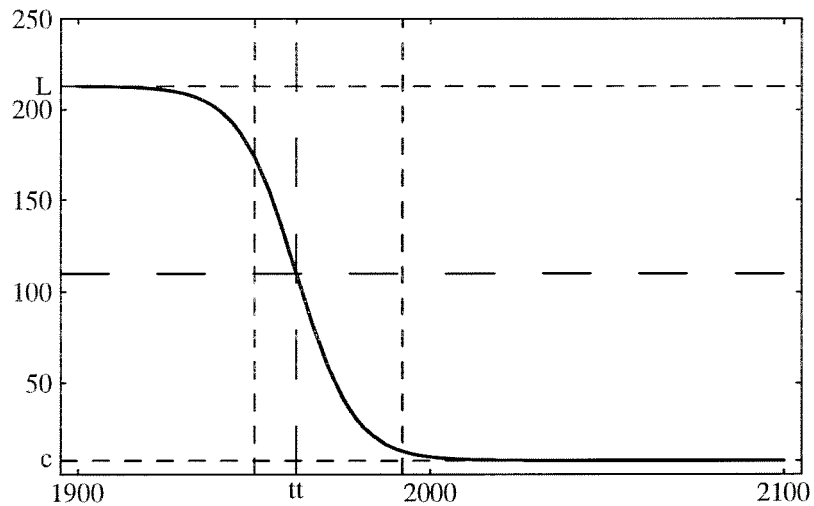
$$r_l(t) = \frac{e^{at+b}}{1 + qe^{at+b}} + c$$

Waarbij $q \geq 0$. Als $q = 0$ dan komt dit overeen met een exponentieel model. De bovenlimiet is nu gelijk aan $c + \frac{1}{q}$. Om praktische redenen is in plaats van q het kwadraat d^2 gezet, zodat $d^2 \geq 0$. Deze functie is verder in de analyses gebruikt.

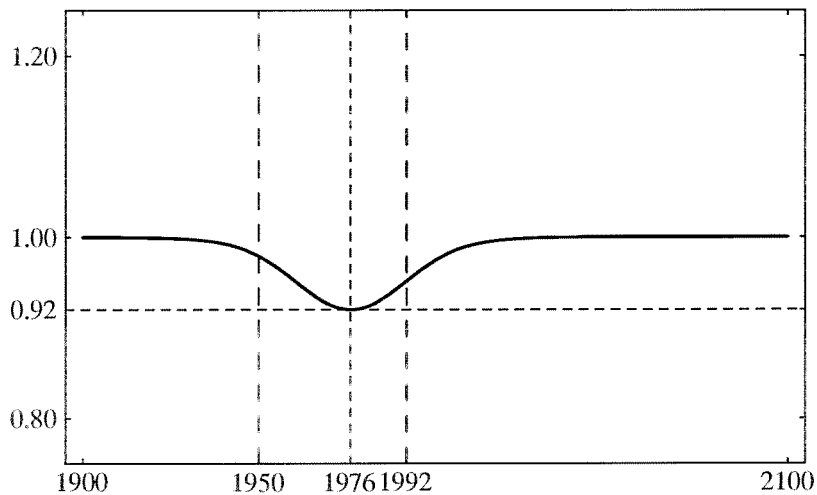
$$r_l(t) = \frac{e^{at+b}}{1 + d^2 e^{at+b}} + c \quad (2.4)$$

In *Afbeelding 2* staat een voorbeeld van een geschatte logistische risico-kromme gebaseerd op gewogen regressie van Nederlandse gegevens. Bij de waarde tt buigt de kromme. Volgens deze analyse zou het keerpunt in de afname van het risico zich reeds in 1960 hebben voorgedaan.

In *Afbeelding 3* staat de verhouding tussen het risico in een jaar gedeeld door het risico het jaar daarvoor.



Afbeelding 2. Voorbeeld logistische risico-kromme
Gewogen regressie van Nederland-gegevens. Tijdstip tt is het buigpunt in de ontwikkeling.



Afbeelding 3. Voorbeeld verhoudingen logistische risico-kromme

Deze verhouding kan worden gezien als een maat voor het afnemen van het risico. In het jaar 1976 blijkt deze afname het grootst te zijn met als verhouding (0.921). Volgens deze ontwikkeling zal de verhouding van de risico's in het jaar 2011 ongeveer gelijk zijn aan 0.99, of wel in het jaar 2011 zal het risico nog maar één procent zijn afgenomen ten opzichte van het jaar daarvoor (2010). Het is onvermijdbaar dat de risico-daling in absolute zin kleiner wordt. Sommigen zullen zich afvragen of de afname in de risico-daling zich al zo snel in deze mate zal voordoen. Als daarbij rekening wordt gehouden met een voldoende sterke groei in de mobiliteit, zal blijken dat de verkeersonveiligheid in termen van het totaal aantal slachtoffers of ongevallen zelfs zal gaan stijgen. In *Tabel 18* staan vergelijkbare jaartallen van risicodaling voor alle geschatte modellen. Hierbij moet echter wel rekening worden gehouden met een beperkte nauwkeurigheid van de voorspelling van de diverse modellen.

2.2. Verkeersprestatie-modellen

Behalve aannames over de risico-ontwikkeling is in beperkte mate rekening gehouden met verschillende verkeersprestatie-modellen. Er is uitsluitend gebruik gemaakt van een mobiliteitscijfer gebaseerd op het gemotoriseerd verkeer. Daarbij is een algemene transformatie toegepast:

$$v(t) = v_t^e \quad (2.5)$$

waarbij v_t het geobserveerde mobiliteitscijfer is. Telkens is getoetst of $\equiv 1$. In een aantal gevallen blijkt toevoeging van de parameter e tot significant betere resultaten te leiden.

2.3. Samenvatting gebruikte modellen

Hieraan voorafgaand zijn een aantal modellen gedefinieerd. Uitgegaan is van het model zoals dat in Bijleveld & Oppe (1992) is gebruikt: het exponentiële model (2.1) met $c = 0$. Dit model is Exp0 genoemd. Telkens is een risico-model met een verkeersprestatie-model gecombineerd. In *Tabel 1* zijn de combinaties weergegeven.

Model		mobiliteit	
Risico	Bijzonderheid	$e = 1$	$e \neq 1$
Exponentieel	$c = 0, d = 0$	Exp0	ExpV0
	$c \neq 0, d = 0$	Exp	ExpV
Logistisch	$c \neq 0, d \neq 0$	Logis	LogisV

Tabel 1. *Schema predictie-modellen*

Het model Exp0 het model van Bijleveld & Oppe (1992), is op twee manieren uitgebreid. Ten eerste naar Exp door $c \neq 0$ toe te laten en ten tweede naar ExpV0 door het andere verkeersprestatie model toe te passen. Beide modellen kunnen uitgebreid worden tot ExpV. Verder zijn de modellen Logis en LogisV geschat.

De formules staan in *Tabel 2*.

Model	Risico	Predictie
Exp0	e^{b+at}	$e^{b+at} v$
Exp	$c + e^{b+at}$	$(c + e^{b+at}) v$
ExpV0	e^{b+at}	$e^{b+at} v^e$
ExpV	$c + e^{b+at}$	$(c + e^{b+at}) v^e$
Logis	$c + \frac{e^{b+at}}{1+d^2 e^{b+at}}$	$\left(c + \frac{e^{b+at}}{1+d^2 e^{b+at}} \right) v$
LogisV	$c + \frac{e^{b+at}}{1+d^2 e^{b+at}}$	$\left(c + \frac{e^{b+at}}{1+d^2 e^{b+at}} \right) v^e$

Tabel 2. *Gebruikte modellen*

3. Statistische aspecten

3.1. Niet-lineaire regressie

3.1.1. Keuze afhankelijke variabele

Zoals al in Bijleveld & Oppe (1992) beschreven staat, is ook in dit geval gekozen voor een aanpak waarin de risicokromme geschat wordt door het aantal doden te voorspellen. Dit in tegenstelling tot het berekenen van een empirisch risico, door het aantal doden te delen door de verkeersprestatie en vervolgens dit quotiënt te gebruiken als te verklaren grootte.

3.1.2. Over Poisson-variantie

De meest gebruikte veronderstelling in de analyses van aantallen ongevallen per tijdseenheid is waarschijnlijk dat de aantallen ongevallen per tijdseenheid bij benadering Poisson verdeeld zijn. Daar het aantal doden per ongeval gelukkig klein is, wordt meestal (en ook in dit geval) verondersteld dat het aantal doden per tijdseenheid wederom bij benadering Poisson verdeeld is. Deze veronderstelling houdt in dat het uiteindelijk resulterende aantal slachtoffers volgens een bepaald patroon toevallig afwijkt van het meest waarschijnlijke aantal. Grotere afwijkingen kunnen zich voordoen, doch de kansen daarop zijn kleiner. De verdeling van deze kansen zijn (karakteristiek) bepaald door de (Poisson)verdeling. In Bijleveld & Oppe (1992) is deze aanname ook gedaan.

Zoals de meeste modellen, is ook het Poisson-model slechts een benadering van de werkelijkheid. Bovendien blijkt dat, gezien het feit dat in dit stuk slechts de totale aantallen slachtoffers geanalyseerd zijn, de normale verdeling onder omstandigheden weer een redelijke benadering van de Poisson-verdeling te kunnen leveren. Hierbij kunnen op deze verdeling gebaseerde toetsen, zoals de χ^2 -toets en de F-toets worden toegepast. Een aantal zaken dienen hierbij in ogenschouw genomen te worden.

Twee fundamentele eigenschappen van het Poisson-model zijn dat:

- a. De uitkomsten slechts gehele getallen zijn, die groter of gelijk aan nul zijn.
- b. De variantie van de uitkomsten even groot is in magnitude als de verwachte waarde van de uitkomsten.

Aan eigenschap (a.) is natuurlijk bij het analyseren van slachtoffergegevens automatisch voldaan. Eigenschap (b.) daarentegen heeft verstrekkender gevolgen. Vaak zal in de praktijk blijken dat deze variantie groter is. Een dergelijk uit de literatuur (Bijvoorbeeld Cox, 1983) bekend verschijnsel heet 'bovenmatige spreiding' ofwel *overdispersion*. Vaak wordt verondersteld dat bovenmatige spreiding wordt veroorzaakt door het tellen van aantallen (slachtoffers) uit inhomogene groepen.

In het onderhavige geval zal eventuele 'bovenmatige spreiding' veroorzaakt kunnen worden doordat in plaats van ongevallen doden (of zelfs slachtoffers) geanalyseerd worden. Dit betekent dat in feite 'gebeurtenissen' niet alleen komen, maar juist in groepjes (clusters). Dit kan een grotere variantie veroorzaken.

De oorzaak voor ‘bovenmatige spreiding’ zal in dit onderzoek eerder liggen in het feit dat de gebruikte risico- en mobiliteitsmodellen de reeksen observaties matig tot redelijk beschrijven. Daardoor zullen de ‘verwachte aantallen’ slachtoffers extra afwijken van de geobserveerde aantallen slachtoffers. Een andere reden tot afwijking van het model van de werkelijkheid is het feit dat de mobiliteitscijfers welke gebruikt worden om de aantallen slachtoffers te voorspellen zelf ook aan een zekere onnauwkeurigheid lijden. Deze onnauwkeurigheid is in het algemeen onbekend.

Er is besloten rekening te houden met deze effecten door óf uit te gaan van een model waar deze effecten verwaarloosbaar worden geacht óf door aan te nemen dat dit effect zo groot is dat deze alle variatie overheerst.

Zuiver Poisson-model Modellen zullen met behulp van methoden worden geschat, die gebaseerd zijn op meest aannemelijke schatters (maximum-likelihood). Er wordt in dit geval aangenomen dat de extra spreiding geheel afwezig is of proportioneel is met de Poisson-spreiding.

Gewone kleinste kwadraten Deze aanname levert schattingen voor de parameters, uitgaande van de veronderstelling dat de extra spreiding de Poisson-spreiding zodanig overschrijdt dat de laatste verwaarloosd kan worden.

Tussen beide bovenstaande uitersten zou men de gewogen kleinste kwadraten kunnen plaatsen. Dit model is niet gebaseerd op de Poisson-aanname, maar houdt wel rekening met de verschillen in variantie. Omdat dit model geen expliciete verdelingsaanname heeft in de zin dat het alle observaties als gelijkwaardig beschouwt, is het robuuster tegen grote afwijkingen van de Poisson-verdeling. Vele diagnostische technieken zijn beschikbaar voor deze modellen. Om die reden is het merendeel van de controlerende analyses gebaseerd op dit model.

3.1.3. *Niet-lineaire-regressie implementatie*

Niet-lineaire regressie-analyse is geen eenvoudig probleem. De meeste bestaande technieken zijn gebaseerd op klassieke lineaire regressie-technieken, waarbij het niet-lineaire model lokaal ‘gelineariseerd’ wordt. Het proces van het uitvoeren van een dergelijke analyse valt uiteen in twee stappen:

- De numerieke stap: het vinden van parameterwaarden waarvoor een criterium optimaal is. In het (gewogen) kleinste kwadraten geval betekent dit het minimaliseren van een zogenaamde doelfunctie, in het maximum-likelihood geval betekent dit het maximaliseren van een functie. In de praktijk wordt meestal minus de likelihoodfunctie geminimaliseerd. Dit probleem is op zich al geen sinecure, maar meestal niet het grootste probleem, zeker als redelijke beginafschattingen beschikbaar zijn. Het algoritme dat zowel SAS als Mathematica hiervoor kunnen gebruiken, heet het Marquardt-Levenberg algoritme (Fletcher, 1981p. 82). In § 3.1.4, 3.1.5 en 3.1.6 worden de doelfuncties van de verschillende criteria omschreven.
- Nadat op basis van numerieke criteria een passende oplossing gevonden is, moet deze oplossing aan statistische analyse worden onderworpen. Dit gedeelte is technisch het meest ingewikkeld. SAS levert zelf uitsluitend een asymptotische analyse (zie § 3.2.1). Hoewel een dergelijke analyse

over het algemeen behoorlijke resultaten oplevert, is er geen aanleiding niet wat specifiek naar een oplossing te kijken. In § 3.2.2 staat kort beschreven waarom een toets voor het afwijken van het lineair model nuttig kan zijn.

In de nu volgende sub-paragrafen wordt een beknopte uiteenzetting gegeven van de doelfuncties behorende bij de diverse optimaliteitscriteria. Daarbij worden de volgende symbolen gedefinieerd:

- t is het symbool voor de tijd. Het symbool wordt ten eerste zelf als variabele gebruikt. Ten tweede wordt het ook als index gebruikt voor waargenomen gegevens zoals:
- d_t is het aantal doden in het jaar t .
- v_t is het cijfer voor de mobiliteit in het jaar t .

In Bijleveld & Oppe (1992) is het model Exp0 geschat; verondersteld is dat het risico op het tijdstip t r_t gelijk is aan:

$$r_t = e^{at+b}$$

Daarbij veronderstellend dat de schatting van d_t :

$$\hat{d}_t = v_t \times r_t = v_t e^{at+b} = p_t.$$

Deze p_t is in dit geval feitelijk een functie van v_t , t en de parameters a en b . In andere gevallen zullen er nog enkele parameters bijkomen. Essentieel in dit onderdeel is dat er een vergelijking tussen de geobserveerde d_t enerzijds en de geschatte p_t anderzijds optreedt.

3.1.4. Gewone regressie

In het geval van gewone regressie (OLS) worden de kwadraten van de verschillen tussen observaties (d_t) en hun voorspellingen (p_t) opgeteld. Meestal wordt dit getal gedeeld door het aantal getelde termen minus één. Dit is dus een soort gemiddelde van het kwadraat van de afwijkingen:

$$ssq_{ols} = (d_{1950} - p_{1950})^2 + (d_{1951} - p_{1951})^2 + \dots + (d_{1992} - p_{1992})^2$$

Belangrijk is dat verondersteld wordt dat van de verschillen $d_t - p_t$ wordt aangenomen dat ze enigermate gelijksoortig verdeeld zijn. Dit houdt praktisch in dat men verwacht dat de getallen $(d_t - p_t)^2$ ongeveer even groot zijn. Een andere manier om het bovenstaande te zeggen is dat je aanneemt dat alle p_t gelijkwaardige voorspellers zijn van hun respectievelijke d_t 's. Men veronderstelt dat de varianties van de voorspellingen ongeveer gelijk aan elkaar zijn. Als dat niet het geval is, dan is de volgende aanpassing meestal bruikbaar.

3.1.5. Gewogen regressie

In het geval van gewogen regressie (WLS) wordt de kwadraten som van de verschillen *gewogen* opgeteld. De weging door getallen w_t wordt zo gedaan dat de getallen $w_t(d_t - p_t)^2$ ongeveer even groot zijn. Voor dit getal w_t wordt hier $1/d_t$ gebruikt.

$$ssq_{wls} = (d_{1950} - p_{1950})^2/d_{1950} + \dots + (d_{1992} - p_{1992})^2/d_{1992}$$

De keuze van $w_t = 1/d_t$ in dit geval is gemaakt omdat aangenomen wordt dat de aantallen doden ongeveer Poisson verdeeld zijn. Er wordt gedacht dat er achter het geobserveerde aantal doden in feite een Poisson verdeelde stochastische grootheid zit, die toevallig d_t doden heeft opgeleverd. We kunnen stellen dat het te verwachten aantal doden in een bepaald jaar gelijk is aan δ_t . Dit normaal te verwachten aantal doden kan in eerste instantie alleen met het geobserveerde aantal worden geschat. Eigenschap (b) aan het begin van dit hoofdstuk stelde dat de variantie in grootte ongeveer gelijk is aan de verwachting, δ_t , dus de variantie schatten we ook door middel van het geobserveerde aantal. Door dus telkens ieder kwadraat door d_t te delen zouden we dus weer in de situatie terecht komen van de vorige paragraaf. Deze methode wordt ook wel de gemodificeerde chikwadraat-analyse genoemd.

Nu kleven hier problemen aan: het aantal doden zou gelijk aan nul kunnen zijn. Als het aantal doden klein is, dan zou de weging te grof kunnen zijn. Een jaar waarin één dode valt telt twee keer zo zwaar mee in de analyse als een jaar waarin twee doden vallen. Dit verschil kan puur toevallig zijn. In het onderhavige geval is er telkens sprake van veel grotere aantallen, zodat dit effect verwaarloosbaar zal zijn.

Een tweede punt is dat de gewichten zelf schatters zijn van iets dat we met het model proberen te schatten. Sterker, als verondersteld wordt dat het gebruikte model voldoende is, dan zullen de getallen p_t betere schattingen voor de δ_t zijn dan de getallen d_t . Het model met de keuze $w_t = 1/p_t$ is moeilijker te schatten. Een bekende methode is door gebruik te maken van de zogenaamde Iteratively Reweighted Least Squares of wel (IRLS):

$$ssq_{irls} = (d_{1950} - p_{1950})^2/w_{1950,i} + \dots + (d_{1992} - p_{1992})^2/w_{1992,i} \quad (3.1)$$

Deze methode schat een oplossing door herhaald een gewogen regressie-model te schatten, net zolang totdat aan een bepaald convergentie-criterium is voldaan.

Stap(1) $w_{t,1} \equiv 1/d_t$. Dit levert de gewogen kleinste kwadraten schatting $\hat{p}_{t,1}$ voor δ_t op.

Stap(i) $w_{t,i} \equiv 1/\hat{p}_{t,i-1}$. Dit levert $\hat{p}_{t,i}$ op.

Deze methode is gebruikt in Bijleveld & Oppe (1992). De resultaten zijn uitwisselbaar met het volgende Poisson-model. Deze methode wordt soms de minimum-chikwadraat methode genoemd.

3.1.6. Poisson-regressie

Bij Poisson-regressie wordt van het volgende uitgegaan. Onder de Poisson-assumptie wordt gesteld dat de kans op een aantal doden in een bepaald jaar gelijk is aan

$$P(d_t) = \frac{e^{-\delta_t} \delta_t^{d_t}}{d_t!}$$

Aangezien wordt aangenomen dat het aantal doden in het ene jaar niet het aantal doden in het andere jaar beïnvloedt, is de totale kans op het feit dat $d_{1950}, \dots, d_{1992}$ wordt waargenomen gegeven de functie δ_t :

$$P(d_{1950}) \times \dots \times P(d_{1992})$$

δ_t wordt zo gekozen dat het bovenstaande maximaal is. In de praktijk minimaliseert men meestal minus de logaritme van bovenstaande functie. Dit gebeurt door het sommeren van termen van het type:

$$\delta_t - d_t \times \log \delta_t + \log d_t! \quad (3.2)$$

De laatste term van (3.2) is constant en dus niet interessant bij het optimaliseren, maar de tweede term ($\log \delta_t$) kan in de praktijk tot problemen leiden, als δ_t volgens een model waarden kan aannemen die kleiner dan of gelijk aan nul zijn. Zowel theoretisch als praktisch is dit echter geen reëel probleem. Uit voorzorg is echter toch eerst een OLS-oplossing gezocht, vervolgens is daarvan uitgaande een WLS-oplossing gezocht om daarna de Poisson (ML) oplossing te zoeken. Langs deze weg heeft het groter of gelijk aan nul probleem in de praktijk geen problemen gegeven.

3.2. Diagnostische analyses

Het analyseren van resultaten van niet-lineaire regressie modellen kan uiteenvallen in de volgende stappen:

- Controle of aan numerieke condities is voldaan. Bijvoorbeeld of er feitelijk een optimale waarde is aangenomen. De meeste software zorgt hier zelf voor.
- Visuele inspectie van het resultaat, in dit geval meestal een ‘plotje’.
- Indien gebruik wordt gemaakt van asymptotische betrouwbaarheidsintervallen moet gecontroleerd worden of deze asymptotische benadering redelijk geacht mag worden. Dit wordt gedaan via ‘curvature-tests’. Zie hiervoor § 3.2.2.
- Indien besloten wordt om gebruik te maken van asymptotische betrouwbaarheidsintervallen, dan kan gecontroleerd worden of bepaalde in het model aanwezige parameters wel terecht in het model aanwezig zijn.
- Als alternatief voor asymptotische betrouwbaarheidsintervallen kan de relevantie van een parameter worden getest door twee versies van een model te schatten, één met en één zonder de betreffende parameter. Dan kan via een F-ratio toets in het (gewogen) kleinste-kwadraten-geval of een likelihood-ratio toets in het maximum-likelihood geval de relatieve bijdrage vastgesteld worden.
- Indien daar behoefte aan is kan via een Lagrange-Multipliertoets onderzocht worden of er voordeel te behalen valt als een parameter aan het model wordt toegevoegd. Een omschrijving van de Lagrange-Multipliertoets staat in § 3.2.4.

3.2.1. Asymptotische analyse

In het klassieke geval van multiple-regressie wordt uitgegaan van een model waarbij de te verklaren variabele y_t verklaard wordt uit een lineaire combinatie van verklarende variabelen X_{tj} :

$$y_t = \beta_1 X_{t1} + \dots + \beta_n X_{tn} + \varepsilon_t$$

terwijl voor de voorspellingen geldt:

$$\hat{y}_t = \hat{\beta}_1 X_{t1} + \dots + \hat{\beta}_n X_{tn}$$

$$\hat{y}_t = y_t + \varepsilon_t$$

waarbij de ε_t worden verondersteld onderling onafhankelijk te zijn en normaal verdeeld met verwachting nul en variantie σ^2 .

In de huidige situatie is er sprake van een model

$$y_t = f_t(\theta) \quad (3.3)$$

De waarden y_t worden dus verondersteld een functie van de tijd en een (onbekende) parameter θ te zijn. Ook kan het zijn dat bijvoorbeeld andere gegevens zoals de verkeersprestatie een rol speelt.

Indien θ^* de werkelijke waarde van θ is en $\hat{\theta}$ een schatting hiervoor, dan kan voor $\hat{\theta} - \theta^* \approx 0$ gesteld worden dat (Seber & Wild, 1988), in het lineaire geval:

$$\hat{y}_t - y_t = \sum_{j=1}^n X_{tj}(\hat{\beta}_j - \beta_j)$$

In het niet-lineaire geval:

$$f_t(\hat{\theta}) - f_t(\theta^*) \approx \sum_{j=1}^n \frac{\partial f_t}{\partial \theta_j}(\hat{\theta}_j - \theta_j^*) \quad (3.4)$$

Dus, in het niet-lineaire geval gedragen de termen

$$\frac{\partial f_t}{\partial \theta_j}$$

zich als de X_{tj} in het lineaire model. De covariantie matrix van $\hat{\theta}$ wordt in dit geval op dezelfde wijze berekend als in het lineaire geval. Er wordt echter sterk 'geleund' op de lineaire benadering van f_t rondom θ^* . Het probleem kan zijn dat de afwijking $\hat{\theta} - \theta^*$ groter is dan volgens deze lineaire benadering van f_t toelaatbaar is. § 3.2.2 behandelt dit onderwerp.

3.2.2. Krommings analyse

In § 3.3.2 en *Hoofdstuk 4* van Seber & Wild (1988) wordt een uitgebreide uiteenzetting gegeven van de gevolgen van de niet-lineariteit van f_t . Ook Ross (1990) besteedt aandacht aan dit onderwerp, zij het op een beperkt gebied. Refererend aan ouder werk, stellen beide auteurs dat de niet-lineariteit van een probleem kan worden gesplitst in een component 'intrinsieke' niet-lineariteit en een component 'parameter' niet-lineariteit. De eerste term wordt in de literatuur *intrinsic curvature* genoemd, de tweede *parameter-effects curvature*.

Voor een beschrijving van de methode wordt verwezen naar Seber & Wild (1988). Zie ook het eenvoudiger voorbeeld in hun § 3.3.2.

Praktisch is er voor ieder model de benadering (Seber & Wild, 1988p. 129):

$$\begin{aligned} f(\theta) - f(\hat{\theta}) &\approx \hat{F}'(\theta - \hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})' \hat{F}''(\theta - \hat{\theta}) \\ &= \hat{F}'\delta + \frac{1}{2}\delta' \hat{F}''\delta \end{aligned} \quad (3.5)$$

In het geval dat er sprake is van n observaties of tijdstippen zullen $f(\theta)$ en $f(\hat{\theta})$ n -dimensionale vectoren zijn. Indien p parameters geschat worden

is $\hat{F}_{..}$ een $n \times p \times p$ dimensionale matrix, bestaande uit $n \times p$ matrices van tweede-orde-afgeleiden van het model f in $\hat{\theta}$. Daardoor is $\delta' \hat{F}_{..} \delta$ een n -dimensionale vector.

De relatieve kromming is nu de verhouding tussen deze laatste vector (de tweede-orde-informatie) en het kwadraat van de lengte van de vector $\hat{F}_{..}(\theta - \hat{\theta})$, (de eerste-orde-informatie).

$$\frac{\|\delta' \hat{F}_{..} \delta\|}{\|\hat{F}_{..} \delta\|^2}$$

De vector $\delta' \hat{F}_{..} \delta$ wordt verder ontbonden in:

T_{δ} een vector liggend in het raakvlak van de oplossing en

N_{δ} een component daar loodrecht op.

De eerste component levert de intrinsic curvature op:

$$K_{\delta}^T = \frac{\|T_{\delta}\|}{\|\hat{F}_{..} \delta\|^2}$$

en de tweede de parameter-effects curvature:

$$K_{\delta}^N = \frac{\|N_{\delta}\|}{\|\hat{F}_{..} \delta\|^2}.$$

Uit (3.5) volgt dat beide maten voldoende klein moeten zijn zodat $\frac{1}{2} \delta' \hat{F}_{..} \delta$ relatief klein is, wil een lineaire benadering (3.4) voldoende zijn.

De verdeling van de curvature maten Beide curvature maten bevatten het quotiënt van de standaardfout van de observaties (het kwadraat van de lengte van de vector $\hat{F}_{..}(\theta - \hat{\theta})$) en een kwadratische vorm in $p = \dim(\theta)$ dimensies. De maximale waarden blijken $\sqrt{F_{\alpha}}$ niet te mogen overschreiden, met α de onbetrouwbaarheidsmarge en F de geschikte F-verdeling.

Gevolgen van onvoldoende lineariteit In § 5.8 van Seber & Wild (1988) staat omschreven welk effect curvature heeft op 'Linearized regions'. Indien de intrinsieke kromming te groot wordt, blijkt het mogelijk een schatting te geven voor de aanpassing van de F-statistics zodat daarvan afgeleide betrouwbaarheidsintervallen redelijk blijken. In de praktijk komt dit zelden voor, in slechts twee gevallen in dit onderzoek. Zie hiervoor *Tabel 3*.

	BRD	GBR	JAP	NLD	USA
Exp0	0.02	0.01	0.10	0.03	0.02
Exp	0.04	0.03	0.08	0.05	0.05
ExpV0	0.03	0.01	0.06	0.02	0.04
ExpV	0.06	0.03	0.12	0.05	0.09
Logis	0.12	0.07	0.14	0.09	*38.57
LogisV	0.31	0.40	0.53	0.59	*658220.91

Tabel 3. *Intrinsieke krommings-waarden en significantie tegen 5% * voor gewogen modellen*

Indien niet aan de ‘parameter effects curvature-toets’ is voldaan, is het waarschijnlijk verstandiger voor het berekenen van de betrouwbaarheidsintervallen gebruik te maken van zogenaamde exacte betrouwbaarheidsintervallen (zie Ross, 1990, § 2.4.4). Deze intervallen worden berekend zonder dat er wordt uitgegaan van dat zij symmetrisch om de verwachte waarde liggen. Er wordt hier echter ook gebruik gemaakt van asymptotische grootheden. Deze zijn echter wel betrouwbaar als aan de intrinsieke kromming wordt voldaan. Bij onbetrouwbaarheid van de betrouwbaarheidsintervallen van de parameters is het raadzaam ook voorzichtig te zijn met het berekenen van betrouwbaarheidsintervallen voor de prognoses.

Een alternatief bij onbetrouwbaarheid van de parameterbetrouwbaarheidsintervallen kan het her-parametriseren van het model zijn. Bijvoorbeeld door in plaats van b de vorm $\log b$ te gebruiken, of omgekeerd. Dit heeft op zich dus geen invloed op de mathematische vorm van een model zelf, maar alleen op hoe de parameters deze vorm kunnen beïnvloeden. Daardoor zullen ze ook op een iets andere, hopelijk minder problematische wijze geschat worden. Dit soort aanpassingen kan het probleem vergaand verminderen.

	BRD	GBR	JAP	NLD	USA
Exp0	0.03	0.02	0.09	0.03	0.03
Exp	*1.29	*1.82	0.22	*1.07	*1.83
ExpV0	0.05	0.03	0.13	0.04	0.06
ExpV	*5.75	*8.20	0.39	*4.80	*10.68
Logis	*447553.99	*1154.58	0.48	*1.01	*3.91 × 10 ¹⁶
LogisV	*4.02	*1.73 × 10 ⁹	*8.52	*17.38	*8.47 × 10 ²⁰

Tabel 4. *Parameter krommings-waarden en significantie tegen 5% * voor gewogen modellen*

Aangezien uit *Tabel 4* blijkt dat ongeveer alle hier gebruikte modellen een te grote ‘parameter-effects curvature’ hebben, lijkt het verstandig een procedure voor exacte betrouwbaarheidsintervallen te ontwikkelen, zodat dit probleem opgelost kan worden. De betrouwbaarheidsintervallen van de parameters gebaseerd op de gebruikelijke asymptotische normaliteitsaannames zullen met een behoorlijke terughoudendheid moeten worden geïnterpreteerd. Dit geldt dus bijvoorbeeld voor *Tabel 15*.

3.2.3. Residuele analyse

Een belangrijk onderdeel van de diagnose van analyse-resultaten is het controleren of er geen belangrijke aspecten vergeten zijn.

Het analyseren van residuen kan hiervoor gebruikt worden. De residuen zijn de verschillen tussen de geobserveerde waarden en de voorspellingen of in dit geval het aantal doden en de voorspelling daarvan:

$$r_t = y_t - \hat{p}_t \equiv d_t - \hat{p}_t$$

Bij de analyse van tijdsafhankelijke gegevens is het verstandig te onderzoeken of er een zeker patroon in de residuen bestaat: bijvoorbeeld veel waarden achter elkaar positief of juist veel waarden achter elkaar negatief.

Voor dit doel is de Box-Ljung-Q toets gebruikt om correlaties tussen de residuen op te sporen. In formulevorm: (Harvey, 1989)

$$Q = n(n + 2) \sum_{k=1}^K \frac{\rho_k^2}{n - k}$$

- n is het aantal observaties.
- $\rho_k = \sum_{i=k+1}^n r_i r_{i-k} / \sum_{i=1}^n r_i^2$. De ρ_k zijn de auto-correlaties met lag k .
- K het aantal auto-correlatie termen dat wordt gebruikt, 30 in dit geval.

Onder de nul-hypothese van géén tijdsafhankelijke residuen heeft Q een χ^2 verdeling met K vrijheidsgraden. We moeten opmerken dat deze toets een omnibustoets is die een groot aantal alternatieven moet onderscheiden. Dit betekent dat het mogelijk is dat een toets die een specifiek alternatief moet onderscheiden wel de noodzaak van een alternatief vindt, waar de Box-Ljung-Q toets dit niet mocht aangeven.

	BRD	GBR	JAP	NLD	USA
Exp0	*158.19	*87.73	*266.02	*159.05	*219.55
Exp	*70.27	*44.85	*202.00	*104.30	*214.33
ExpV0	*130.16	*47.40	*264.60	*75.91	*216.13
ExpV	*65.60	*45.14	*119.96	*63.01	*224.15
Logis	*70.27	*44.85	*132.23	*46.06	*214.33
LogisV	40.09	*45.14	*124.43	*46.41	*224.15

Tabel 5. *Box-Ljung Q statistic voor gewogen modellen. * betekent significant onder aanname van ongecorrleerde residuen.*

In Tabel 5 zijn voor de gebruikte modellen de diverse Box-Ljung-Q toetsen weergegeven. Met een ster (*) is aangegeven welke significant afwijken (tegen 5% onbetrouwbaarheid). Waarschijnlijk is behalve het Duitse model (en eventueel het Engelse model) geen van de gebruikte modellen statistisch gezien voldoende. Het is zeer aannemelijk dat er nog componenten aan de diverse modellen toegevoegd moeten worden.

Een bekend verschijnsel is periodiek te hoog of te laag voorspellen van het aantal doden. Een bekende uitbreiding om hiervoor te compenseren is het toevoegen van een golfbeweging (hier en daar ook 'slingerbeweging' genoemd). Voor dit specifieke soort modellen is dit reeds eerder gesuggereerd, door bijvoorbeeld Oppe (1991a), Koornstra (1992). Een methode voor onderzoek naar de eventuele noodzaak en de optimale frequentie die daarbij ongeveer gebruikt moet worden, staat uitgewerkt in § 3.2.4. De resultaten staan in § 4.5.

Wel moet hierbij bedacht worden dat de analyses betrekking hebben op grote aantallen doden. Daarbij kunnen ook relatief kleine modelafwijkingen die niet echt relevant behoeven te zijn, tot een significant effect leiden.

3.2.4. Lagrange-multiplier analyse

Een methode voor het zoeken naar specifieke termen is gebaseerd op de methode van Lagrange-multipliers. Het idee hierachter is het volgende. Als een specifieke term die wordt toegevoegd aan een model een bijdrage aan

het model levert, moet de kwadratensom van een model *met* deze term meer dan triviaal kleiner zijn dan het model dat anderszins gelijk is *zonder* deze term. Dit kan onderzocht worden door twee modellen apart te schatten en de resultaten te vergelijken. Deze methode heeft echter het belangrijke nadeel dat dit vaak (en ook in dit geval) een enorme hoeveelheid rekenwerk oplevert.

De truc die hier gebruikt wordt, is gebaseerd op het volgende voorbeeld: stel algemeen dat het model $y_t = f(t)$ is geschat. Als een additionele bijdrage van een component $p \times g(t)$ niets oplevert, moet de afgeleide van

$$ssq(p) = \sum_{t=1}^n (y_t - f(t) - p g(t))^2 \quad (3.6)$$

naar p ongeveer gelijk aan nul zijn.

De Lagrange-multipliertoets is op dit fenomeen gebaseerd. Met behulp van deze toets wordt de hypothese $p = 0$ getoetst.

In dit onderzoek is op subtiele plaatsen in een aantal modellen, bijvoorbeeld in de exponent, de term $p \cos(2t\pi/q + r)$ toegevoegd. Gegeven het vooraangenomen feit dat $p \equiv 0$, fit het model met deze term even goed als zonder. Vervolgens kan met de Lagrange-multipliertoets de lineaire hypothese $p = 0, q = q_0, r = r_0$ getest worden. De toetsingsgrootheid is χ_3^2 verdeeld, zie *Hoofdstuk 5.3* van Seber & Wild (1988). In § 4.5 staat voor een selecte groep modellen een analyse-resultaat.

3.3. Vergelijking van modellen

Een belangrijk punt in de analyse van een reeks gegevens met behulp van een aantal modellen is de onderlinge vergelijking van deze eenheden. Dit geldt dus zowel voor de vergelijking van de modellen, als de vergelijking van de diverse reeksen (landen).

In dit stuk is als eerste de vergelijking van de diverse modellen *binnen* de diverse landen benadrukt. Vervolgens is tussen de diverse landen gekeken welke modellen de gegevens blijkbaar het beste benaderden. Dit is op deze wijze gedaan, omdat het nauwelijks zinnig lijkt sterk verschillende modellen, op verschillende reeksen gefit, te vergelijken.

Ten behoeve van het vergelijken van modellen staan in principe drie methoden ter beschikking:

1. Gebruik maken van de R^2 -waarde. Deze is gedefinieerd als:

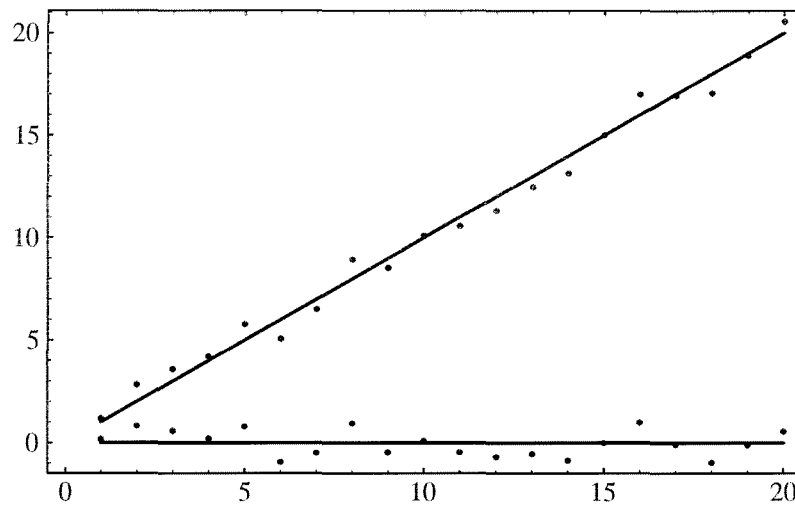
$$R^2 = 1 - \frac{\sum (y_t - \hat{p}_t)^2}{\sum (y_t - \bar{y})^2}$$

Deze grootheid wordt bijna als standaard in de regressie-analyse gebruikt. In de appendix van Harvey (1984) wordt, enigszins summier, uiteengezet waarom deze maat hier beter niet gebruikt kan worden, onder meer omdat deze maat soms een te optimistisch beeld van een fit kan geven. Dit zou vooral het geval zijn als een model enigszins in staat blijkt een trend te volgen. In Kendall & Ord (1990) wordt dit standpunt bevestigd. Het

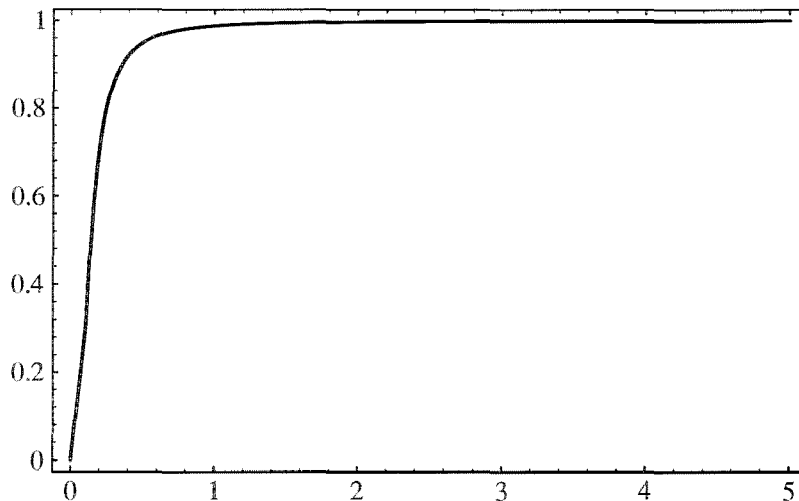
alternatief dat beide auteurs leveren is:

$$R_D^2 = 1 - \frac{\sum (y_t - \hat{p}_t)^2}{\sum (\Delta y_t - \overline{\Delta y})^2}$$

Hierbij is $\Delta y_t = y_t - y_{t-1}$. $\overline{\Delta y}$ is dus de gemiddelde waarde van Δy , ofwel $(y_n - y_1)/(n - 1)$. De rationale achter deze aanpassing lijkt dat een tijdreeksmodel (met trend) vergeleken moet worden met een simpele 'random walk with drift'. Een andere kijk is om aan te nemen dat in vergelijking met gewone regressie, de (eerste orde) bijdrage de tijd als vanzelfsprekend moet worden beschouwd.



Afbeelding 4. Twee reeksen data met precies dezelfde modelfouten.



Afbeelding 5. Ontwikkeling R^2 (verticaal) als functie van de steilheid (horizontaal) van de reeks in Afbeelding 4.

In de Afbeeldingen 4 en 5 wordt duidelijk dat de waarde van de R^2 -grootte afhankelijk is van de variatie in de reeks gegevens. In Afbeelding 4 staan twee reeksen waarden en voorspellingen daarvoor. De voorspellingsfouten voor beide reeksen zijn exact gelijk. Toch heeft de schuine lijn een veel hogere R^2 -waarde dan de horizontale lijn. In het algemeen neemt de R^2 -waarde met de richtingscoëfficiënt (tangens van de hoek met

de horizontale as) toe. Dit staat in *Afbeelding 5* aangegeven. Hoe steiler de lijn in *Afbeelding 4*, hoe groter de R^2 -waarde, terwijl de modelafwijkingen precies hetzelfde blijven. De R_D^2 -groottheid blijft in dit geval constant. In het geval dat er in plaats van een rechte lijn een kwadratische functie gebuikt wordt, is dit niet meer het geval en zal de kromme van de R_D^2 -groottheid een gelijksoortige vorm hebben als de kromme van de R^2 -groottheid in *Afbeelding 5*.

Een opvallend fenomeen van de maat R_D^2 is dat het kleiner dan nul kan uitvallen. In dit geval fit een model dus minder goed dan een op 'random walk with drift' lijkend model. Volgens Harvey (1984) moet zo'n model niet in aanmerking genomen worden.

2. Gebruikmakend van de Poisson-oplossing kan een Likelihood-Ratio toets worden toegepast.
3. Twee modellen kunnen door middel van een F-ratio toets worden vergeleken. Daarbij wordt het quotiënt van de kwadratensommen van beide modellen berekend en vergeleken met wat statistisch te verwachten is.

Een eerste stap voor de analyses is het berekenen van de R^2 , R_D^2 , likelihoodwaarden en kwadratensommen van de respectievelijke modellen. De R^2 -groottheid is voor ongewogen-, gewogen-en likelihoodmodellen in de *Tabellen 6, 8 en 13* aangegeven. De R_D^2 -resultaten staan, uitsluitend voor ongewogen modellen, in *Tabel 7*.

De likelihoodgegevens staan in *Tabel 14* en in *Tabel 9* staan de kwadratensommen voor de gewogen analysemodellen. Bij deze laatste tabel is verondersteld dat de kwadratensom als volgt is gedefinieerd:

$$ssq = \sum \frac{(O_i - E_i)^2}{O_i}$$

Deze kwadratensom wordt soms de gemodificeerde kwadratensom genoemd. Een zuiverder variant is gegeven in *Tabel 11*, waarbij de kwadratensom is berekend als:

$$ssq = \sum \frac{(O_i - E_i)^2}{E_i} \quad (3.7)$$

Dit is de kwadratensom behorende bij de minimum-chi-kwadraat methode. De laatste som is bij benadering χ_{n-p-1}^2 verdeeld, indien het model correct is gespecificeerd.

De procedure om modellen te selecteren is nu te zoeken naar dat model waarvoor geldt dat de kwadratensom (3.7) van alle andere modellen gedeeld door zijn eigen kwadratensom significant groter is dan één. Dit wordt telkens per land uitgevoerd. De resultaten staan in de *Tabellen 10a-e* voor de gemodificeerde chi-kwadraat en *Tabellen 12a-e* voor de zuivere chi-kwadraat op basis van (3.7).

4. Resultaten

4.1. Overzicht

In het nu volgende hoofdstuk wordt een opsomming gegeven van de resultaten van de diverse analyses.

De volgende analyses zullen daarbij aan de orde komen:

1. Gewone kleinste kwadraten. Van alle gefitte modellen zijn de ondergewone kleinste kwadraten optimale parameterwaarden weergegeven. Zie hiervoor *Tabellen 28a–e* (bijlage). Verder zijn de zogenaamde R^2 -grootheden in *Tabel 6* gegeven. Voor het gewone kleinste kwadratengeval is ook een R_D^2 -tabel gegeven in *Tabel 7*. Zie voor achtergronden van deze tabellen § 3.3.
2. Gewogen kleinste kwadraten (gemodificeerde chi-kwadraat). Bij deze techniek (en de volgende) worden de volgende grootheden berekend:
 - Onder gewogen kleinste kwadraten optimale parameterwaarden. Zie *Tabellen 29a–e* (bijlage).
 - R^2 -grootheden in *Tabellen 8*. Géén R_D^2 -grootheid.
 - Bij de modellen behorende chi-kwadraat waarden. Zie hiervoor *Tabel 9*.
 - F-toetsen voor de chi-kwadraat ratio's van de modellen, voor ieder land apart. Zie hiervoor *Tabellen 10a–e*.Zie voor verdere toelichting ook de tekst.
3. Maximum likelihoodmodellen. Op basis van de maximum likelihood-schattingen zijn twee soorten analyses uitgevoerd:
 - (a) chi-kwadraat analyse. In feite is hetzelfde uitgevoerd als voor de chi-kwadraat analyse, behalve dat nu gewogen is met de geschatte aantallen doden in plaats van het geobserveerde aantal. Deze variant komt in § 4.2.3 aan de orde. Onder deze aanname zijn in deze paragraaf de volgende tabellen weergegeven:
 - Bij de modellen behorende chi-kwadraat waarden. Zie hiervoor *Tabel 11*.
 - F-toetsen voor de chi-kwadraat ratio's van de modellen, voor ieder land apart. Zie hiervoor *Tabellen 12a–e*.
 - (b) Maximum likelihoodanalyse. Uiteindelijk zijn uitsluitend de log-likelihood waarden berekend (*Tabel 14*). Behalve deze tabel zijn analoog aan de gewogen analyse de volgende tabellen aangemaakt:
 - Onder Poisson-maximum likelihood aanname optimale parameter waarden. Zie hiervoor *Tabellen 30a–e* in de bijlage.
 - R^2 -grootheden in *Tabel 13*. Er is géén R_D^2 -grootheid berekend.

4.2. Primaire resultaten

4.2.1. Gewone kleinste kwadraten

Zie voor een opmerking over het gebruik van de R^2 -grootheid § 3.3. Ter vergelijking zijn hieronder de R^2 -waarden voor de gewone kleinste kwadraten modellen gegeven (Tabel 6). Ter vergelijking staan in Tabel 7 de R_D^2 -waarden van de diverse modellen.

	BRD	GBR	JAP	NLD	USA
Exp0	0.89	0.88	0.01	0.91	0.85
Exp	0.91	0.94	0.75	0.93	0.86
ExpV0	0.89	0.91	0.56	0.94	0.85
ExpV	0.92	0.94	0.79	0.95	0.86
Logis	0.91	0.94	0.90	0.97	0.86
LogisV	0.95	0.94	0.90	0.97	0.86

Tabel 6. R^2 statistic voor niet-gewogen modellen.

	BRD	GBR	JAP	NLD	USA
Exp0	-0.31	0.10	-9.54	-0.30	-0.58
Exp	0.02	0.55	-1.67	0.06	-0.52
ExpV0	-0.22	0.34	-3.65	0.23	-0.56
ExpV	0.11	0.55	-1.27	0.25	-0.52
Logis	0.02	0.55	-0.07	0.61	-0.52
LogisV	0.40	0.55	-0.07	0.61	-0.52

Tabel 7. R_D^2 statistic voor niet-gewogen modellen.

Als men de theorie van Harvey (1984) aanneemt, dan zou men geen van de modellen voor Japan of Verenigde Staten aannemen. Bovendien blijkt dat uit deze reeks modellen eigenlijk alleen voor West-Duitsland modellen met een exponent voor de verkeersprestatie in aanmerking komen.

Op grond van deze resultaten zou men voor West-Duitsland voor LogisV kiezen, voor Engeland voor Exp en voor Nederland voor Logis of misschien LogisV. Dan wordt wel uitgegaan van het ideaal zijn van zowel de R_D^2 -grootheid als het gewone kleinste kwadraten model. Hoewel dat laatste geen groot effect lijkt te hebben gehad.

4.2.2. Gewogen analyse

Om te beginnen is voor iedere gewogen analyse de R^2 -grootheid weer uitgerekend. De resultaten hiervan staan in Tabel 8. Aangezien er eigenlijk geen voor de handliggend alternatief is voor de R_D^2 -grootheid onder gewogen analyse, is deze hier weggelaten. Opvallend is dat de resultaten in Tabel 8 nauwelijks afwijken van vergelijkbare resultaten voor niet-gewogen modellen in Tabel 6.

	BRD	GBR	JAP	NLD	USA
Exp0	0.88	0.88	0.00	0.90	0.85
Exp	0.91	0.94	0.73	0.93	0.86
ExpV0	0.89	0.91	0.54	0.94	0.85
ExpV	0.92	0.94	0.78	0.94	0.86
Logis	0.91	0.94	0.90	0.97	0.86
LogisV	0.95	0.94	0.90	0.97	0.86

Tabel 8. R^2 statistie voor gewogen modellen.

Vergelijking kwadratensommen met behulp van F -waarden

In Tabel 9 staan de kwadratensommen

$$ssq_{wts} = \sum_t \frac{(d_t - p_t)^2}{d_t}$$

weergegeven. Hierbij stelt d_t het aantal doden op tijdstip t voor en p_t de voorspelling daarvan volgens een bepaald model.

	BRD	GBR	JAP	NLD	USA
Exp0	3607.11 (37)	902.73 (40)	36540.32 (39)	918.85 (40)	23202.61 (67)
Exp	2674.87 (36)	409.46 (39)	10644.85 (38)	675.68 (39)	22697.62 (66)
ExpV0	3353.56 (36)	647.04 (39)	17224.54 (38)	562.24 (39)	23181.43 (66)
ExpV	2445.01 (35)	404.47 (38)	8433.29 (37)	552.22 (38)	22577.76 (65)
Logis	2674.87 (35)	409.46 (38)	3436.39 (37)	240.21 (38)	22697.62 (65)
LogisV	1523.27 (34)	404.47 (37)	3424.38 (36)	240.20 (37)	22577.76 (64)

Tabel 9. χ^2 -waarden en vrijheidsgraden (df) voor gewogen modellen

In principe is men geïnteresseerd in die oplossing waarvoor ssq_{wts} minimaal is. Een model met een kleinere kwadratensom past beter op de data en zal daardoor waarschijnlijk de werkelijkheid beter benaderen dan een model dat een grotere kwadratensom heeft. Daarbij dient echter te worden opgemerkt dat ook rekening moet worden gehouden met het feit dat bepaalde modellen deze kwadratensom bereiken door gebruik te maken van meer parameters. Dergelijke modellen hebben daardoor meer

mogelijkheden zich aan de gegevens aan te passen. Er moet dus zowel op de kwadratensom gelet worden als op het aantal 'vrijheidsgraden'. Hoe kleiner het aantal gebruikte vrijheidsgraden van de oplossing, hoe meer parameters zijn gebruikt.

Opvallend in Tabel 9 is dat de relatieve verschillen in kwadraten-sommen voor de Verenigde Staten onderling nauwelijks verschillen vergeleken met de andere landen. Exp doet het relatief goed. Men kan in één oogopslag zien dat het model LogisV voor West-Duitsland het beste zal fitten, terwijl het model Logis dat doet voor Nederland (LogisV is slechts marginaal beter, zoals uit Tabel 10d zal blijken). Voor Engeland lijkt de keuze op het model Exp te vallen. Het model Logis lijkt voor Japan de beste keuze.

West-Duitsland						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.76 (0.79)	0.96 (0.55)	0.72 (0.84)	0.78 (0.76)	0.46 (0.99)
Exp	1.31 (0.21)	1.00 (0.50)	1.25 (0.25)	0.94 (0.57)	1.03 (0.47)	0.60 (0.93)
ExpV0	1.05 (0.45)	0.80 (0.75)	1.00 (0.50)	0.75 (0.80)	0.82 (0.72)	0.48 (0.98)
ExpV	1.40 (0.16)	1.06 (0.43)	1.33 (0.20)	1.00 (0.50)	1.09 (0.40)	0.64 (0.90)
Logis	1.28 (0.24)	0.97 (0.53)	1.22 (0.28)	0.91 (0.60)	1.00 (0.50)	0.59 (0.94)
LogisV	2.18 (0.01)	1.66 (0.07)	2.08 (0.02)	1.56 (0.10)	1.71 (0.06)	1.00 (0.50)

Tabel 10a. *F*-toetsen voor gewogen modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Engeland						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.47 (0.99)	0.74 (0.83)	0.47 (0.99)	0.48 (0.99)	0.48 (0.99)
Exp	2.15 (0.01)	1.00 (0.50)	1.58 (0.08)	1.01 (0.48)	1.03 (0.47)	1.04 (0.45)
ExpV0	1.36 (0.17)	0.63 (0.92)	1.00 (0.50)	0.64 (0.91)	0.65 (0.91)	0.66 (0.90)
ExpV	2.12 (0.01)	0.99 (0.52)	1.56 (0.09)	1.00 (0.50)	1.01 (0.49)	1.03 (0.47)
Logis	2.09 (0.01)	0.97 (0.53)	1.54 (0.09)	0.99 (0.51)	1.00 (0.50)	1.01 (0.48)
LogisV	2.06 (0.01)	0.96 (0.55)	1.52 (0.10)	0.97 (0.53)	0.99 (0.52)	1.00 (0.50)

Tabel 10b. *F*-toetsen voor gewogen modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Japan						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.30 (1.00)	0.48 (0.99)	0.24 (1.00)	0.10 (1.00)	0.10 (1.00)
Exp	3.34 (0.00)	1.00 (0.50)	1.62 (0.07)	0.81 (0.73)	0.33 (1.00)	0.34 (1.00)
ExpV0	2.07 (0.01)	0.62 (0.93)	1.00 (0.50)	0.50 (0.98)	0.20 (1.00)	0.21 (1.00)
ExpV	4.11 (0.00)	1.23 (0.27)	1.99 (0.02)	1.00 (0.50)	0.41 (1.00)	0.42 (0.99)
Logis	10.09 (0.00)	3.02 (0.00)	4.88 (0.00)	2.45 (0.00)	1.00 (0.50)	1.02 (0.47)
LogisV	9.85 (0.00)	2.94 (0.00)	4.77 (0.00)	2.40 (0.01)	0.98 (0.53)	1.00 (0.50)

Tabel 10c. *F*-toetsen voor gewogen modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

In de Tabellen 10a–e staan voor deze modellen de *F*-ratio toetsen weergegeven. Telkens staat voor een land hoe het ene model (horizontaal) zich tot de andere modellen (kolommen) verhoudt. In iedere cel staat de *F*-ratio en tussen haakjes zijn overschrijdingskansen. Dit getal, in een schaal tussen 0 en 1, geeft aan hoe groot de kans is dat de ratio toevallig zo groot uitvalt als in de tabel staat, terwijl de ratio gemiddeld een waarde van één zou moeten hebben.

Uit Tabel 10a blijkt dat voor West-Duitsland het model LogisV goed doch statistisch niet overmatig beter fit dan het model Exp dat tenslotte twee parameters minder bevat. Daarmee is overigens rekening gehouden in de toets. De modellen Exp en ExpV verschillen blijkbaar nauwelijks zodat aan genomen mag worden dat de parameter *e* in ExpV niet significant van één verschilt.

Uit Tabel 10b blijkt dat voor Engeland het model Exp gekozen moet worden. De *F*-ratio van ExpV / Exp is zelfs groter dan één, door het verschil

Nederland						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.75 (0.81)	0.63 (0.93)	0.63 (0.92)	0.28 (1.00)	0.28 (1.00)
Exp	1.33 (0.19)	1.00 (0.50)	0.83 (0.72)	0.84 (0.71)	0.36 (1.00)	0.37 (1.00)
ExpV0	1.59 (0.07)	1.20 (0.28)	1.00 (0.50)	1.01 (0.49)	0.44 (0.99)	0.45 (0.99)
ExpV	1.58 (0.08)	1.19 (0.29)	0.99 (0.51)	1.00 (0.50)	0.44 (0.99)	0.45 (0.99)
Logis	3.63 (0.00)	2.74 (0.00)	2.28 (0.01)	2.30 (0.01)	1.00 (0.50)	1.03 (0.47)
LogisV	3.54 (0.00)	2.67 (0.00)	2.22 (0.01)	2.24 (0.01)	0.97 (0.53)	1.00 (0.50)

Tabel 10d. *F*-toetsen voor gewogen modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Verenigde Staten						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.99 (0.51)	1.01 (0.48)	1.00 (0.49)	1.01 (0.49)	1.02 (0.47)
Exp	1.01 (0.49)	1.00 (0.50)	1.02 (0.47)	1.01 (0.48)	1.02 (0.48)	1.03 (0.46)
ExpV0	0.99 (0.52)	0.98 (0.53)	1.00 (0.50)	0.99 (0.52)	0.99 (0.51)	1.00 (0.49)
ExpV	1.00 (0.51)	0.99 (0.52)	1.01 (0.48)	1.00 (0.50)	1.01 (0.49)	1.02 (0.48)
Logis	0.99 (0.51)	0.98 (0.52)	1.01 (0.49)	0.99 (0.51)	1.00 (0.50)	1.01 (0.48)
LogisV	0.98 (0.53)	0.97 (0.54)	1.00 (0.51)	0.98 (0.52)	0.99 (0.52)	1.00 (0.50)

Tabel 10e. *F*-toetsen voor gewogen modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

in vrijheidsgraden. Opvallend was de moeizame fit van de logistische modellen. Dit kan voor een deel veroorzaakt zijn doordat het buigpunt van de ontwikkeling (1957) dicht bij het begin van de gegevensreeks ligt (1950). In vergelijking: in Nederland was dat 1960, hetgeen de helft meer gegevens voor het buigpunt betekent.

Uit *Tabel 10c* blijkt dat voor Japan het model Logis gekozen moet worden. De *F*-ratio van LogisV / Logis is zelfs groter dan één door het verschil in vrijheidsgraden. Het verschil tussen deze beide modellen en de rest blijkt zeer groot. Verder blijkt de exponent op de verkeersprestatie bij de exponentiële ontwikkeling enige verbetering te leveren.

Uit *Tabel 10d* blijkt dat voor Nederland het model Logis gekozen moet worden. Hier blijkt de *F*-ratio van LogisV / Logis groter dan één te zijn. Het in Bijleveld & Oppe (1992) gebruikte model (Exp) kan zich met geen van de alternatieven meten. Deze conclusie geldt overigens uitsluitend voor de ontwikkeling van het totaal aantal doden. Voor andere, vooral kortere reeksen zal het om pragmatische redenen noodzakelijk blijven het in Bijleveld & Oppe (1992) gebruikte exponentiële model toe te passen.

Uit *Tabel 10e* blijkt dat voor de Verenigde Staten het model Exp0 gekozen moet worden. Geen van de andere modellen lijken voldoende aan het model toe te voegen. Ook onder deze omstandigheden lijkt geen van de modellen een betere verklaring te geven dan het Exp0-model.

Resumerend kan worden geconstateerd dat er niet één model is dat het duidelijk beter doet dan de andere modellen. Voor de landen met de snelste risicodaling, respectievelijk Japan, Nederland, en West-Duitsland lijkt een logistisch model

het beter te doen dan bij de landen met een wat langzamere daling, zoals de Verenigde Staten en Engeland. Hierbij lijkt de voorkeur uit te gaan naar een exponentieel model. Een nadere analyse, waarbij bijvoorbeeld voor de landen met een snelle risico-daling de laatste jaren worden weggelaten, is nodig om meer inzicht te krijgen in deze landelijke verschillen. Waarschijnlijk is het ook verstandig het aantal te vergelijken landen uit te breiden.

4.2.3. χ^2 -analyse

In deze paragraaf worden de analyses uit § 4.2.2 vanuit een enigszins gewijzigd perspectief uitgevoerd. In plaats van een gewogen analyse is hier gebruik gemaakt van een klassieke χ^2 -analyse, echter niet door deze waarde te minimaliseren, maar door de aannemelijkheid ervan te maximaliseren. Dit levert dezelfde uitkomsten. Het criterium is nu:

$$ssq_{\chi^2} = \sum \frac{(d_i - p_i)^2}{p_i}$$

In *Tabel 11* staan de aldus verkregen kwadratensommen weergegeven.

	BRD	GBR	JAP	NLD	USA
Exp0	3640.35 (37)	873.74 (40)	49143.18 (39)	833.55 (40)	24018.57 (67)
Exp	2711.62 (36)	399.20 (39)	10838.50 (38)	635.40 (39)	23233.45 (66)
ExpV0	3381.73 (36)	633.42 (39)	16850.74 (38)	538.14 (39)	23991.00 (66)
ExpV	2474.55 (35)	395.26 (38)	9450.98 (37)	530.38 (38)	23107.26 (65)
Logis	2711.62 (35)	399.20 (38)	3580.15 (37)	244.04 (38)	23233.45 (65)
LogisV	1544.08 (34)	395.26 (37)	3554.97 (36)	243.97 (37)	23107.26 (64)

Tabel 11. χ^2 -waarden en vrijheidsgraden (*df*) op basis van likelihood fit

De resultaten zijn duidelijk overeenkomstig met die uit *Tabel 9*. Het gaat hierbij weer om die oplossing waarvoor de kwadratensom minimaal is, bij gebruikmaking van zo min mogelijk parameters. Ook in *Tabel 11* valt het op dat de verschillen in kwadratensommen voor de Verenigde Staten onderling nauwelijks verschillen, vergeleken met de andere landen. Men kan ook hier weer in één oogopslag zien dat het model LogisV voor West-Duitsland het beste zal fitten, terwijl het model Logis dat doet voor Nederland (LogisV is slechts marginaal beter, zoals uit *Tabel 12d* ook zal blijken). Voor Engeland zal de keuze weer op het model Exp blijken te vallen.

West-Duitsland						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.77 (0.79)	0.95 (0.55)	0.72 (0.84)	0.79 (0.76)	0.46 (0.99)
Exp	1.31 (0.21)	1.00 (0.50)	1.25 (0.26)	0.94 (0.57)	1.03 (0.47)	0.60 (0.93)
ExpV0	1.05 (0.45)	0.80 (0.74)	1.00 (0.50)	0.75 (0.80)	0.82 (0.71)	0.48 (0.98)
ExpV	1.39 (0.16)	1.07 (0.43)	1.33 (0.20)	1.00 (0.50)	1.10 (0.39)	0.64 (0.90)
Logis	1.27 (0.24)	0.97 (0.53)	1.21 (0.29)	0.91 (0.61)	1.00 (0.50)	0.59 (0.94)
LogisV	2.17 (0.01)	1.66 (0.07)	2.07 (0.02)	1.56 (0.10)	1.71 (0.06)	1.00 (0.50)

Tabel 12a. *F*-toetsen voor likelihood modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Engeland						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.47 (0.99)	0.74 (0.82)	0.48 (0.99)	0.48 (0.99)	0.49 (0.98)
Exp	2.13 (0.01)	1.00 (0.50)	1.59 (0.08)	1.02 (0.48)	1.03 (0.47)	1.04 (0.45)
ExpV0	1.34 (0.18)	0.63 (0.92)	1.00 (0.50)	0.64 (0.91)	0.65 (0.91)	0.66 (0.90)
ExpV	2.10 (0.01)	0.98 (0.52)	1.56 (0.09)	1.00 (0.50)	1.01 (0.49)	1.03 (0.47)
Logis	2.08 (0.01)	0.97 (0.53)	1.55 (0.09)	0.99 (0.51)	1.00 (0.50)	1.02 (0.48)
LogisV	2.04 (0.02)	0.96 (0.55)	1.52 (0.10)	0.97 (0.53)	0.98 (0.52)	1.00 (0.50)

Tabel 12b. *F*-toetsen voor likelihood modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Japan						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.23 (1.00)	0.35 (1.00)	0.20 (1.00)	0.08 (1.00)	0.08 (1.00)
Exp	4.42 (0.00)	1.00 (0.50)	1.55 (0.09)	0.90 (0.63)	0.34 (1.00)	0.35 (1.00)
ExpV0	2.84 (0.00)	0.64 (0.91)	1.00 (0.50)	0.58 (0.95)	0.22 (1.00)	0.22 (1.00)
ExpV	4.93 (0.00)	1.12 (0.37)	1.74 (0.05)	1.00 (0.50)	0.38 (1.00)	0.39 (1.00)
Logis	13.02 (0.00)	2.95 (0.00)	4.58 (0.00)	2.64 (0.00)	1.00 (0.50)	1.02 (0.48)
LogisV	12.76 (0.00)	2.89 (0.00)	4.49 (0.00)	2.59 (0.00)	0.98 (0.52)	1.00 (0.50)

Tabel 12c. *F*-toetsen voor likelihood modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Nederland						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.78 (0.78)	0.66 (0.90)	0.67 (0.89)	0.31 (1.00)	0.32 (1.00)
Exp	1.28 (0.22)	1.00 (0.50)	0.85 (0.70)	0.86 (0.68)	0.39 (1.00)	0.40 (1.00)
ExpV0	1.51 (0.10)	1.18 (0.30)	1.00 (0.50)	1.01 (0.49)	0.47 (0.99)	0.48 (0.99)
ExpV	1.49 (0.11)	1.17 (0.32)	0.99 (0.51)	1.00 (0.50)	0.46 (0.99)	0.47 (0.99)
Logis	3.24 (0.00)	2.54 (0.00)	2.15 (0.01)	2.17 (0.01)	1.00 (0.50)	1.03 (0.47)
LogisV	3.16 (0.00)	2.47 (0.00)	2.09 (0.01)	2.12 (0.01)	0.97 (0.53)	1.00 (0.50)

Tabel 12d. *F*-toetsen voor likelihood modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

In de Tabellen 12a–e staan voor de op de loglikelihood gebaseerde modellen de *F*-ratio toetsen weergegeven. Telkens staat voor een land hoe het ene model (horizontaal) zich tot de andere modellen (kolommen) verhoudt. In iedere cel staat de *F*-ratio en tussen haakjes de overschrijdingskans hiervan. Dit is analoog aan Tabellen 10a–e.

Uit Tabel 12a blijkt weer dat voor West-Duitsland het model LogisV zich sterk doch statistisch niet overmatig beter verhoudt met het model Exp dat twee parameters minder bevat. De modellen Exp en ExpV verschillen ook hier nauwelijks, zodat aangenomen mag worden dat de parameter e in ExpV niet significant van één verschilt.

Uit Tabel 12b blijkt weer dat voor Engeland het model Exp gekozen moet worden. Ook nu levert ExpV te weinig extra boven Exp.

Verenigde Staten						
	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.98 (0.53)	1.01 (0.48)	0.99 (0.51)	1.00 (0.50)	1.01 (0.49)
Exp	1.02 (0.47)	1.00 (0.50)	1.03 (0.45)	1.01 (0.48)	1.02 (0.48)	1.03 (0.46)
ExpV0	0.99 (0.52)	0.97 (0.55)	1.00 (0.50)	0.98 (0.54)	0.98 (0.53)	0.99 (0.51)
ExpV	1.01 (0.49)	0.99 (0.52)	1.02 (0.46)	1.00 (0.50)	1.01 (0.49)	1.02 (0.48)
Logis	1.00 (0.50)	0.98 (0.52)	1.02 (0.47)	0.99 (0.51)	1.00 (0.50)	1.01 (0.48)
LogisV	0.99 (0.51)	0.97 (0.54)	1.01 (0.49)	0.98 (0.52)	0.99 (0.52)	1.00 (0.50)

Tabel 12e. *F*-toetsen voor likelihood modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

Net als in *Tabel 10b* volgt uit *Tabel 12b* dat voor Japan het model Logis gekozen moet worden. Ook nu levert LogisV te weinig extra boven Logis.

Uit *Tabel 12d* blijkt dat voor Nederland het model Logis gekozen moet worden. Het Expmodel uit Bijleveld & Oppe (1992) kan zich evenmin onder de likelihood-oplossing met één van de alternatieven meten.

Ook uit *Tabel 12e* blijkt dat voor de Verenigde Staten het model Exp0 gekozen moet worden. Geen van de modellen blijken voldoende aan dit model toe te voegen.

4.2.4. Likelihood-analyse onder Poisson-aanname

Om te beginnen zijn in *Tabel 13* weer de R^2 -waarden weergegeven. Deze waarden verschillen ook nu weer nauwelijks van de vergelijkbare waarden onder de aanname van de kleinste kwadraten.

	BRD	GBR	JAP	NLD	USA
Exp0	0.88	0.88	-0.12	0.90	0.85
Exp	0.91	0.94	0.75	0.93	0.86
ExpV0	0.89	0.91	0.56	0.94	0.85
ExpV	0.92	0.94	0.78	0.95	0.86
Logis	0.91	0.94	0.90	0.97	0.86
LogisV	0.95	0.94	0.90	0.97	0.86

Tabel 13. R^2 statistic voor likelihood modellen.

	BRD	GBR	JAP	NLD	USA
Exp0	-2040.20 (2)	-668.44 (2)	-22568.07 (2)	-631.88 (2)	-12307.98 (2)
Exp	-1575.51 (3)	-428.13 (3)	-5613.59 (3)	-526.04 (3)	-11964.98 (3)
ExpV0	-1912.13 (3)	-545.77 (3)	-8734.01 (3)	-474.76 (3)	-12295.35 (3)
ExpV	-1457.61 (4)	-426.00 (4)	-4780.18 (4)	-470.50 (4)	-11902.44 (4)
Logis	-1575.51 (4)	-428.13 (4)	-1998.05 (4)	-323.14 (4)	-11964.98 (4)
LogisV	-994.18 (5)	-426.00 (5)	-1987.94 (5)	-323.13 (5)	-11902.44 (5)

Tabel 14. Likelihood-waarden en (aantallen parameters) onder Poisson-aanname

In *Tabel 14* staan de log-likelihood waarden van de diverse modellen weergegeven. Naast ieder cijfer staat tussen haakjes het aantal geschatte parameters. In tegenstelling tot wat het geval is bij de kleinste kwadratenmodellen moet

hier gezocht worden naar de grootste waarde. Alle waarden zijn noodzakelijkerwijs negatief, dus moet gezocht worden naar de minst negatieve waarde.

Het patroon in *Tabel 14* vertoont sterke overeenkomsten met bovenstaande *Tabel 9* en *Tabel 11*. Blijkbaar maakt het niet zoveel uit van welk van de drie criteria (gewone kleinste kwadraten, gewogen kleinste kwadraten, meest aannemelijke schatters) uitgegaan wordt; de drie hier gebruikte criteria leveren dezelfde conclusies. Dit komt onder meer door de grote aantallen doden. De keuze om te delen door het aantal doden of de voorspelling daarvan in *Tabel 9* en *Tabel 11* maakt relatief niet zo veel meer uit op dit niveau. Om ongeveer gelijksoortige redenen bestaat er niet zoveel verschil tussen het maximum likelihoodcriterium en het daarvan afgeleide χ^2 - criterium.

4.3. Van gewogen analyse afgeleide resultaten

Gegeven de gewogen analyse is een aantal eigenschappen van de modellen afgeleid. De begrippen zijn beschreven in *Hoofdstuk 2* en *3*. Van de hier opgevoerde gegevens is geen statistische informatie afgeleid. De betreffende cijfers zullen dus onderworpen zijn aan enige onzekerheid.

4.3.1. Statistische significantie-parameters

	BRD	GBR	JAP	NLD	USA
Exp0					
Exp					c
ExpV0	e				e
ExpV	e, c	e		c	e, c
Logis	d	d			c, d
LogisV		e, d	e	e	e, c, d

Tabel 15. Volgens asymptotische-toets niet significant van nul (1 in het geval van e) verschillende parameters voor gewogen modellen

Hoewel volgens de 'parameters-effects curvature toetsen' (*Tabel 4*) er uiterst voorzichtig met de betrouwbaarheidsintervallen van de parameters moet worden omgegaan, is het toch nuttig om te onderzoeken welke parameters volgens de betrouwbaarheidsintervallen een zinnige bijdrage aan het model toevoegen. In *Tabel 15* staat voor ieder model aangegeven welke parameter wellicht weggelaten kan worden. Behalve voor de parameter *e*, die de machtsverheffing van de verkeersprestatie aangeeft, zou dat het geval kunnen zijn als de waarde 0 binnen het betrouwbaarheidsinterval ligt. Voor de parameter *e* moet dat voor de waarde 1 gelden. Het blijkt uit *Tabel 15* dat dit bij sommige modellen het geval is met betrekking tot de parameters *c* (de exponentiële asymptoot) of *d* (de logistische boven-asymptoot), maar voor geen van de parameters geldt bij de met de *F*-ratio toets geselecteerde modellen. Wat dat betreft voldoen de betrouwbaarheidsintervallen dus wel.

4.3.2. Nulpunten risico-krommen

Eén van de gevolgen van het uitbreiden van het model Exp0 naar Exp (en ExpV) is het verschijnsel dat het risico niet noodzakelijk meer niet-negatief is. Voor Engeland is het model Exp gekozen en het blijkt dat dit model een

	BRD	GBR	JAP	NLD	USA
Exp0					
Exp	2006.94	2010.59		2005.71	2049.66
ExpV0					
ExpV	2002.67	2009.39			2044.47
Logis	2006.94	2010.59			2049.66
LogisV		2009.39			2044.47

Tabel 16. *Nulpunten van risico ontwikkelingen*

negatieve waarde voor c heeft. Dit heeft tot gevolg dat men niet kan aannemen dat zelfs als het model nu de beste beschrijving geeft, dit altijd zo kan blijven. Blijkbaar heeft het model Exp een voorspellingsperiode die tot vóór het jaar 2010 reikt. Het zal dus voor slechts een beperkt aantal jaren (zeg tot het jaar 2000) een voorspellende waarde kunnen hebben.

Behalve de modellen voor de Verenigde Staten en Japan leiden de andere modellen tot een gelijksoortige conclusie.

4.3.3. *Relatieve afname van het risico*

	BRD	GBR	JAP	NLD	USA
Exp0					
Exp			1633		
ExpV0					
ExpV			1699	1526	
Logis			1967	1976	
LogisV	1982		1967	1976	

Tabel 17. *Minima verhouding risico-ontwikkelingen*

In *Tabel 17* staan de minima van de verhoudingen van het risico in een jaar en het jaar daarvoor opgetekend. Zie ook *Afbeelding 3* in § 2.1.4 voor een omschrijving. Opvallend is dat deze minpunten voor de Verenigde Staten en ten dele ook Japan veel eerder vallen in Europa, waarbinnen Nederland weer als eerste komt. Hoewel dit cijfer niet direct als maat in de ontwikkeling van het verkeer dan wel de verkeersveiligheid kan worden gezien, lijkt het wel een indicatie hiervoor. Wellicht speelt de (waarde van) de parameter a hier een rol in.

De cijfers voor de modellen ExpV en Exp in Nederland lijken aan een artefact te liggen.

De *Tabellen 18* en *19* geven het eveneens in § 2.1.4 besproken punt weer, waarin het risico wordt verondersteld slechts met 1% per jaar af te nemen (respectievelijk 0,1%).

Het hangt natuurlijk af van het pessimisme (of utopie) over de ontwikkeling van de verkeersveiligheid of het redelijk is aan te nemen dat de maatschappij ooit in een situatie komt waarin het niet meer mogelijk is of nodig wordt geacht om een redelijke verbetering in de verkeersveiligheid aan te brengen. Toch geeft een cijfer als in 1997, 1% verbetering voor het model LogisV

	BRD	GBR	JAP	NLD	USA
Exp0					
Exp			1996		
ExpV0					
ExpV			1991	2028	
Logis			1988	2009	
LogisV	1997		1988	2009	

Tabel 18. 99% punten van risico-ontwikkelingen

	BRD	GBR	JAP	NLD	USA
Exp0					
Exp			2011		
ExpV0					
ExpV			2002	2052	
Logis			1999	2028	
LogisV	2007		1999	2029	

Tabel 19. 99,9% punten van risico-ontwikkelingen

in West-Duitsland te denken. Ook deze gegevens lijken te duiden op een beperkte toepasbaarheid van dit model na het jaar 2000.

De oorzaak voor deze verschijnselen moet onder meer gezocht worden in het feit dat de modellen geschat moeten worden op basis van gegevens van een toch vrij korte periode in verhouding tot de verandering in ontwikkeling van het risico. Het kan dus zijn dat noch het logistische model, noch het exponentiële model al de juiste 'staart' heeft voor de toekomst.

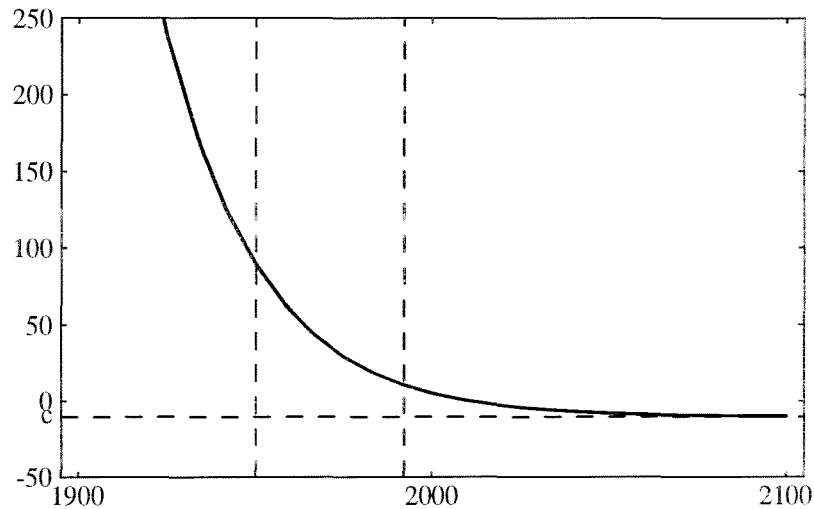
4.4. Vergelijking landen

Uit het bovenstaande kan geconcludeerd worden dat voor de Verenigde Staten met moeite een model gevonden kan worden. Waarschijnlijk is het de moeite waard verder te zoeken naar andere modellen. Ook zou geëxperimenteerd kunnen worden met splines of Wavelets zodat de resultaten daarvan vergeleken kunnen worden met de hier gevonden resultaten. Verder geldt dat de Verenigde Staten een vereniging is van zeer verschillende staten. Aparte analyses, bijvoorbeeld voor west en oost, zouden meer consistente uitkomsten kunnen vertonen.

Zowel voor Japan, West-Duitsland als Nederland heeft het logistische model (Logis of LogisV) duidelijk de voorkeur boven een exponentieel model.

In *Afbeelding 6* blijkt duidelijk dat als de bovengrens van de logistische ontwikkeling te groot wordt, deze ontwikkeling niet van de exponentiële is te onderscheiden.

Het is echter goed mogelijk dat de risico's in West-Europa in de periode vanaf het begin van de jaren zestig redelijk goed met een logistische kromme te beschrijven zijn.



Afbeelding 6. *Logistische ontwikkeling (Logis) voor Engeland*
Deze ontwikkeling is niet van een exponentiële te onderscheiden.

4.5. Golf-analyse

Gegeven de redelijke beschrijving van de modellen voor ieder land is een experiment gedaan met het zoeken naar periodieke afwijkingen van de risicokrommen. Deze afwijkingen worden ook wel slingers of golven genoemd. Uit de economie zijn dergelijke bewegingen bekend. Daar de economie een effect op de mobiliteit heeft en de mobiliteit een effect op de verkeersonveiligheid heeft, hetgeen weer een (zij het uitgesteld) effect heeft op het verkeersgedrag en op de bereidheid tot het nemen van (extra) maatregelen, is het redelijk aan te nemen dat er ook een periodiek effect in de ontwikkeling van het risico bestaat.

In de risico-adaptatie theorie, zoals ontwikkeld in Koomstra (1992), wordt uitgegaan van zo'n directe relatie tussen de economische golfbeweging op de mobiliteit en op het risico. Deze laatste vertoont dan een verschuiving in de tijd.

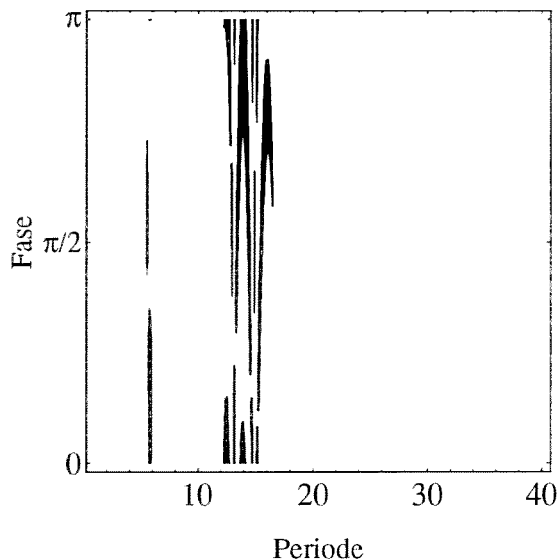
Om dit effect te onderzoeken is op het voor een land selecteerde model een extra analyse uitgevoerd. Daartoe is de risico-functie r_t uitgebreid met een extra term, Koomstra (1992):

$$r_t := r_t e^{p \cos(2t\pi/q+s)}$$

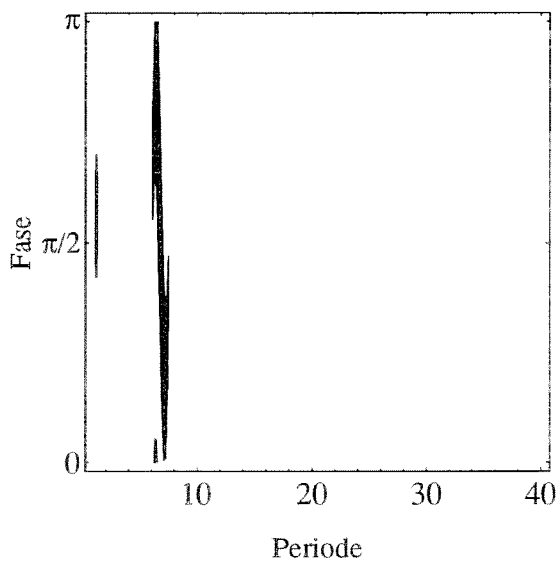
Als $p \equiv 0$ dan is dit model gelijk aan het eerder geselecteerde model zonder golf. Nu is voor verschillende waarden van de parameter q (de lengte van de golf in jaren) en de parameter s (de fase) met behulp van de Lagrange-Multiplier toets onderzocht of de lineaire hypothese $p \equiv 0$, $q \equiv q_i$, $s \equiv s_i$ verworpen moet worden. Dit is een zeer specifieke toets, en onderscheidt een dergelijk tijdsafhankelijk effect beter dan de algemenere Box-Ljung-Q toets uit § 3.2.3.

In de nu volgende grafieken staat voor ieder geanalyseerd land verticaal de fase (s) en horizontaal de golflengte (q) aangegeven. De fase loopt op tot π , hetgeen overeenkomt met 180 graden. Een fase groter dan 180 graden leidt tot een component die gelijkwaardig is aan een component met een

fase 180 graden kleiner, maar in teken verschilt. De relatieve grootte van de toetsingsgrootte is aangegeven met behulp van grijstinten; hoe donkerder, hoe groter de waarde van de toetsingsgrootte van de Langrange-multiplier toets en daarmee hoe significanter de toets. Het gebied waar de toets niet significant is gebleken, is wit gelaten.



Afbeelding 7. Multiplier-toets voor West-Duitsland

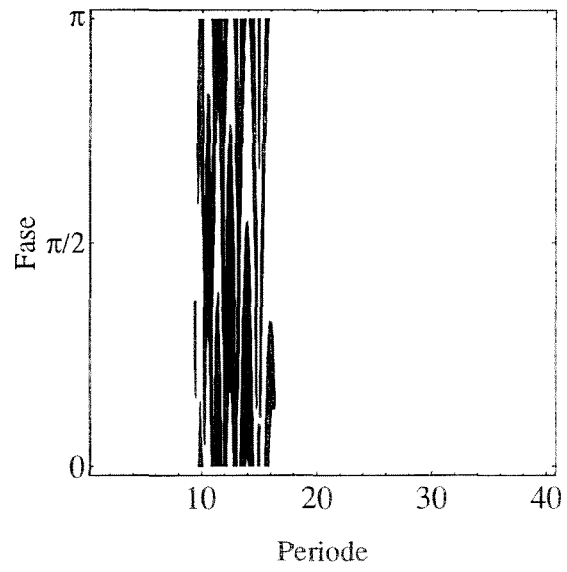


Afbeelding 8. Multiplier-toets voor Engeland

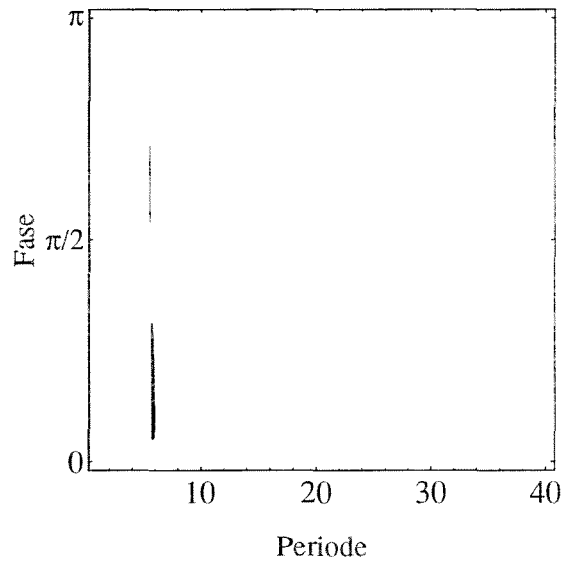
Uit de Afbeeldingen 7, 8 en 10 kan geconcludeerd worden dat er in Europa waarschijnlijk een periodiek effect van ongeveer zes à zeven jaar bestaat. Het lijkt raadzaam dit op te sporen.

In Japan lijkt een eventueel aanwezige golf een langere periode te hebben.

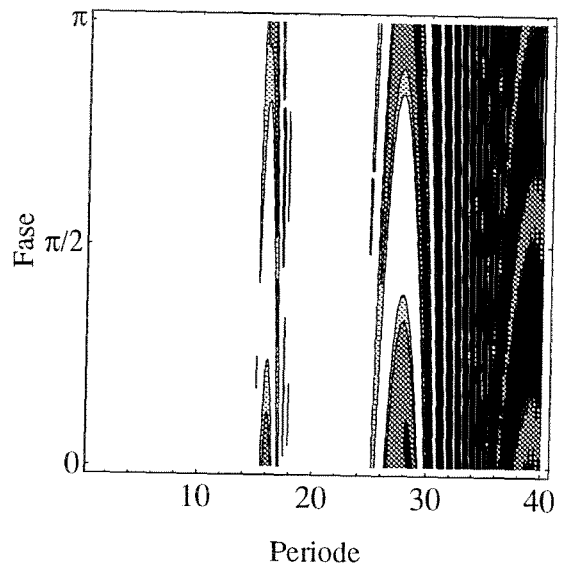
Voor Verenigde Staten lijkt er een belangrijke golf te bestaan met een veel langere golflengte. Dit kan ook het gevolg van een artefact zijn. Voorts zijn de modellen zelf ook niet best. Waarschijnlijk is het verstandig voor dit land nog enige alternatieve modellen te ontwikkelen.



Afbeelding 9. *Multiplier-toets voor Japan*



Afbeelding 10. *Multiplier-toets voor Nederland*



Afbeelding 11. *Multipliertoets voor de Verenigde Staten*

5. Conclusies

Gezien de meervoudige doelstelling van dit onderdeel zullen een aantal verschillende conclusies getrokken kunnen worden.

De eerste beperkte conclusie kan zijn dat het logistieke model Logis optimaal lijkt voor Nederland, in de zin dat dit model te verkiezen blijkt boven alle andere modellen. Hoewel ook aan dit model bezwaren kleven. Zo blijkt alleen het iets complexere model LogisV een iets betere fit te vertonen. Dit gaat echter ten koste van de eenvoud van het model, doordat een extra parameter geschat moet worden. Uit de *Tabellen 10d* en *12d* blijkt eenduidig dat het model Logis statistisch gezien te verkiezen is boven alle anderen. Zoals uit *Afbeelding 10* en *Tabel 5* blijkt, zit er nog enige systematiek in de residuen van Logis, zodat:

- óf er met een tijdsafhankelijk effect of golf rekening moet worden gehouden;
- óf er nog een component niet geïmplementeerd is;
- óf er bestaat een beter fittend model dat niet is toegepast;
- of een combinatie hiervan.

Het verschil met het in Bijleveld & Oppe (1992) gebruikte model Exp0 is dermate, dat dit model verworpen kan worden.

Het is gelukt een analyse-systeem op te zetten waarmee op betrekkelijk eenvoudige wijze modellen voor verschillende landen met elkaar kunnen worden vergeleken. Een rijke collectie technieken om modellen onderling te vergelijken is samengesteld, waaronder χ^2 en likelihood.

Voorts is een systeem opgezet om te onderzoeken of parameters (componenten) terecht in een model zijn opgenomen. Dit wordt gedaan met behulp van een asymptotische analyse. Ook kan onderzocht worden of bepaalde componenten waarschijnlijk een bijdrage aan de fit leveren. Dit wordt gedaan via de Lagrange Multiplier methode. Beide methoden benutten een linearisatie van het model rondom de oplossing. Dit veronderstelt een redelijke lineariteit van het model en of aan deze aanname is voldaan kan onderzocht worden met behulp van de curvatures-toetsen. Het blijkt overigens dat de meeste modellen dusdanig niet lineair zijn, dat enige tot soms grote voorzichtigheid geboden is met de interpretatie van de resultaten, zodat alleen de meest opvallende uitkomsten in aanmerking komen. Het is overigens mogelijk om technieken te ontwikkelen die de noodzaak van gebruikte parameters op een manier te testen die niet zoveel of zelfs niet afhangt van lineariteitsaannames. Het is aan te bevelen om dit uit te voeren, zodat in statistische zin beter naar de modellen zelf gekeken kan worden. Om te onderzoeken of een term aan een model moet worden toegevoegd, zal het waarschijnlijk het beste zijn een uitgebreid model op zich te schatten. Uiteraard blijft de Lagrange Multiplier methode zeer nuttig om kandidaten voor uitbreiding te selecteren.

Behalve de ontwikkelde methoden voor de parameters zijn ook methoden ontwikkeld om een model als geheel te beoordelen. De belangrijkste daarvan is de Portmanteau-toets, een vrij algemene toets voor afwijkingen. De toets kan gezien worden als een analyse voorafgaand aan de Lagrange Multiplier toets, waaruit gevonden kan worden welke aanpassing aan het model nodig is.

Ten slotte zal geconcludeerd moeten worden dat geen van de modellen een soort van globale optimaliteit vertoont. Voor West-Duitsland geldt dat in principe LogisV de beste kaarten heeft. Het fit veel beter dan Exp0 maar de verschillen met Exp en ExpV zijn niet groot te noemen. In het bijzonder valt op dat Exp beter fit dan Logis, zelfs als het aantal vrijheden niet in aanmerking wordt genomen. Toch moet LogisV gekozen worden, hoewel vooral Exp in het achterhoofd gehouden moet worden (zie *Tabel 11* in § 4.2.3).

In Engeland lijkt het model Exp de beste optie. Er blijkt geen model beter te fitten in statistische zin en het model fit statistisch beter dan het enige minder complexe model Exp0.

In Japan blijken de logistieke modellen Logis en LogisV de overhand te hebben. Het complexere model LogisV moet het statistisch gezien afleggen tegen Logis, daar het verschil in fit onvoldoende is om de extra parameter te verantwoorden. Van alle landen is hier de keuze eigenlijk het meest duidelijk. Afgezien van de logistische modellen zal ook hier Exp als tweede uit de strijd komen.

In Nederland blijkt de zaak iets anders te liggen, Logis en LogisV fitten wederom het best waarbij, zoals reeds vermeld, Logis geprefereerd moet worden. Echter, afgezien van deze twee zal Exp0 als optimale gekozen moeten worden, tenzij overschrijdingskansen van ongeveer 10% als significant mogen worden gezien. In dat geval zou ExpV0 en niet Exp verkozen moeten worden. De keuze voor Exp0 houdt in dat in dit geval schijnbaar de constante c niet van nul verschilt. Eigenlijk wordt slechts ten dele het model Exp verworpen.

In de Verenigde Staten ligt de situatie dermate dat eigenlijk alle modellen slecht fitten. Zie hiervoor de afbeeldingen in de bijlage.

Samenvattend lijkt er geen globaal model voor de ontwikkeling van het risico van het aantal doden per jaar te bestaan binnen de hier onderzochte groep. Het lijkt een zinnige gedachte de groep modellen uit te breiden. Hierbij wordt in het bijzonder gedacht aan het toepassen van een andere tijdsas zoals bij Minter (1987), waar de totale verkeersprestatie als tijdseenheid is gebruikt.

Literatuur

- Appel, H. (1982). *Strategische aspekten zur erhöhung der sicherheit im Straßenverkehr*. Automobil-Industrie 3, pagina's 347–356.
- Bijleveld, F.D. (1994). *Risicomodellen voor de module verkeersveiligheid bij de mobiliteitsverkenner*. R-94-24. SWOV, Leidschendam.
- Bijleveld, F.D. & Oppe, S. (1992). *Ontwikkeling van een model voor jaaranalyses ten behoeve van het beleidsinformatie systeem verkeersveiligheid (BIS-V)*. R-92-68. SWOV, Leidschendam.
- Blokpoel, A. (1982). *Note on the development of fatal accidents*. Internal SWOV note.
- Brüning, E., Ernst, R., Glaeser, K.P., Hundhausen, G., Klöckner, J.H., & Pfaffenroth, I. (1986). *Zum Rückgang der Getötetenzahlen im Straßenverkehr: Entwicklung in der Bundesrepublik Deutschland van 1970 bis 1984 (1985)*. Zeitschrift für Verkehrssicherheit 32(4), pagina's 154–163.
- Cox, D.R. (1983). *Some remarks on overdispersion*. Biometrika 70(1), pagina's 269–270.
- Fletcher, R. (1981). *Practical methods of optimization*, deel I. John Wiley & Sons, Chicester.
- Harvey, A.C. (1984). *A unified view of statistical forecasting procedures*. Journal of Forecasting 3, pagina's 245–275.
- Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge.
- Kendall, M. & Ord, J.K. (1990). *Time series*. Edward Arnold, London, derde uitgave.
- Koomstra, M.J. (1988). *Development of road safety in some European countries and USA*. R-88-33. SWOV, Leidschendam.
- Koomstra, M.J. (1992). *The evolution of road safety and mobility*. IATSS Research 16(2), pagina's 129–147.
- Minter, A.L. (1987). *Road casualties: Improvement by learning processes*. Traffic engineering and control 28(2), pagina's 74–79.
- Oppe, S. (1989). *Macroscopic models for traffic and traffic safety*. Accident analysis and Prevention 21, pagina's 225–232.
- Oppe, S. (1991a). *Development of traffic and traffic safety: Global trends and incidental fluctuations*. Accident analysis and Prevention 23(5), pagina's 413–422.
- Oppe, S. (1991b). *Development of traffic and traffic safety in six developed*

countries. Accident analysis and Prevention 23(5), pagina's 401–412.

Oppe, S. & Koomstra, M.J. (1990). *A mathematical theory for related long term developments of road traffic and safety*. In M. Koshi, redactie, Proceedings of the Eleventh International Symposium on Transportation and Traffic Theory, July 18-20, 1990 in Yokohama, Japan, pagina's 113–132. Elsevier, New York.

Oppe, S., Koomstra, 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.

Ross, G. (1990). *Nonlinear Estimation*. Springer-Verlag, Berlin, Heidelberg.

Seber, G.A.F. & Wild, C.J. (1988). *Nonlinear Regression*. John Wiley & Sons, New York.

Smeed, R.J. (1949). *Some statistical aspects of road safety research*. Journal of the Royal Statistical Society A(I), pagina's 1–34.

Tanner, J.C. (1958). *An analysis of increases in motor vehicles in Great Britain and the United States*. RN 3340. Road Research Laboratory, Crowthorne, U. K. Internal Research Note.

Lijst van tabellen

1.	Schema predictie-modellen	17
2.	Gebruikte modellen	17
3.	Intrinsieke krommings-waarden en significantie tegen 5% * voor gewogen modellen	24
4.	Parameter krommings-waarden en significantie tegen 5% * voor gewogen modellen	25
5.	Box-Ljung Q statistic voor gewogen modellen. * betekent significant onder aanname van ongecorrleerde residuen.	26
6.	R^2 statistic voor niet-gewogen modellen.	31
7.	R_D^2 statistic voor niet-gewogen modellen.	31
8.	R^2 statistic voor gewogen modellen.	32
9.	χ^2 -waarden en vrijheidsgraden (df) voor gewogen modellen	32
10a.	F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	33
10b.	F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	33
10c.	F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	33
10d.	F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	34
10e.	F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	34
11.	χ^2 -waarden en vrijheidsgraden (df) op basis van likelihood fit	35
12a.	F -toetsen voor likelihood modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	35
12b.	F -toetsen voor likelihood modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	36
12c.	F -toetsen voor likelihood modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	36
12d.	F -toetsen voor likelihood modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	36
12e.	F -toetsen voor likelihood modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	37
13.	R^2 statistic voor likelihood modellen.	37
14.	Likelihood-waarden en (aantallen parameters) onder Poisson-aanname	37
15.	Volgens asymptotische-toets niet significant van nul (1 in het geval van e) verschillende parameters voor gewogen modellen	38
16.	Nulpunten van risico ontwikkelingen	39
17.	Minima verhouding risico-ontwikkelingen	39
18.	99% punten van risico-ontwikkelingen	40
19.	99,9% punten van risico-ontwikkelingen	40
20.	Gebruikte gegevens voor West-Duitsland	53
21.	Gebruikte gegevens voor Engeland	55
22.	Gebruikte gegevens voor Nederland	57
23.	Gebruikte gegevens voor U.S.A	59
24.	Gebruikte gegevens voor Frankrijk	61

25.	Gebruikte gegevens voor Israel	63
26.	Gebruikte gegevens voor Japan	65
27.	Gebruikte gegevens voor Michigan	67
28a.	Parameter waarden OLS modellen	69
28b.	Parameter waarden OLS modellen	69
28c.	Parameter waarden OLS modellen	69
28d.	Parameter waarden OLS modellen	70
28e.	Parameter waarden OLS modellen	70
29a.	Parameter waarden WLS modellen	71
29b.	Parameter waarden WLS modellen	71
29c.	Parameter waarden WLS modellen	71
29d.	Parameter waarden WLS modellen	72
29e.	Parameter waarden WLS modellen	72
30a.	Parameter waarden Poisson modellen	73
30b.	Parameter waarden Poisson modellen	73
30c.	Parameter waarden Poisson modellen	73
30d.	Parameter waarden Poisson modellen	74
30e.	Parameter waarden Poisson modellen	74
31.	Parameter waarden OLS modellen	105
32.	Parameter waarden OLS modellen	105
33.	Parameter waarden WLS modellen	105
34.	Parameter waarden WLS modellen	106
35.	Parameter waarden Poisson modellen	106
36.	Parameter waarden Poisson modellen	106
37.	Volgens asymptotische-toets niet significant van nul (1 in het geval van e) verschillende parameters voor gewogen modellen	106
38.	R^2 statistic voor niet-gewogen modellen.	107
39.	R_D^2 statistic voor niet-gewogen modellen.	107
40.	R^2 statistic voor gewogen modellen.	107
41.	R^2 statistic voor likelihood modellen.	108
42.	χ^2 -waarden en vrijheidsgraden (df) voor gewogen modellen	109
43.	χ^2 -waarden en vrijheidsgraden (df) op basis van likelihood fit	109
44.	F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	109
45.	F -toetsen voor likelihood modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal	110
46.	Likelihood-waarden en (aantallen parameters) onder Poisson aanname	111
47.	Box-Ljung Q statistic voor gewogen modellen. * betekent significant onder aanname van ongecorreleerde residuen.	111
48.	Intrinsieke krommings-waarden en significantie tegen 5% * voor gewogen modellen	111
49.	Parameter krommings-waarden en significantie tegen 5% * voor gewogen modellen	112
50.	Nulpunten van risico ontwikkelingen	113
51.	Minima verhouding risico ontwikkelingen	113
52.	99% punten van risico ontwikkelingen	113
53.	99,9% punten van risico ontwikkelingen	114

Lijst van figuren

1.	Voorbeeld exponentiële risico kromme	14
2.	Voorbeeld logistische risico-kromme	16
3.	Voorbeeld verhoudingen logistische risico-kromme	16
4.	Twee reeksen data met precies dezelfde modelfouten.	28
5.	Ontwikkeling R^2 (verticaal) als functie van de steilheid (horizontaal) van de reeks in <i>Afbeelding 4</i> .	28
6.	Logistische ontwikkeling (Logis) voor Engeland	41
7.	Multiplier-toets voor West-Duitsland	42
8.	Multiplier-toets voor Engeland	42
9.	Multiplier-toets voor Japan	43
10.	Multiplier-toets voor Nederland	43
11.	Multiplier-toets voor de Verenigde Staten	44
12.	Verkeersprestatie West-Duitsland	54
13.	Verkeersdoden West-Duitsland	54
14.	Verkeersprestatie Engeland	56
15.	Verkeersdoden Engeland	56
16.	Verkeersprestatie Nederland	58
17.	Verkeersdoden Nederland	58
18.	Verkeersprestatie Verenigde Staten	60
19.	Verkeersdoden Verenigde Staten	60
20.	Verkeersprestatie Frankrijk	61
21.	Verkeersdoden Frankrijk	62
22.	Verkeersprestatie Israël	63
23.	Verkeersdoden Israël	64
24.	Verkeersprestatie Japan	65
25.	Verkeersdoden Japan	66
26.	Verkeersprestatie Michigan	68
27.	Verkeersdoden Michigan	68
28.	Doden West-Duitsland, Exp	75
29.	Doden West-Duitsland, Exp0	75
30.	Doden West-Duitsland, ExpV	76
31.	Doden West-Duitsland, ExpV0	76
32.	Doden West-Duitsland, Logis	77
33.	Doden West-Duitsland, LogisV	77
34.	Doden Engeland, Exp	78
35.	Doden Engeland, Exp0	78
36.	Doden Engeland, ExpV	79
37.	Doden Engeland, ExpV0	79
38.	Doden Engeland, Logis	80
39.	Doden Engeland, LogisV	80
40.	Doden Japan, Exp	81
41.	Doden Japan, Exp0	81
42.	Doden Japan, ExpV	82
43.	Doden Japan, ExpV0	82
44.	Doden Japan, Logis	83
45.	Doden Japan, LogisV	83

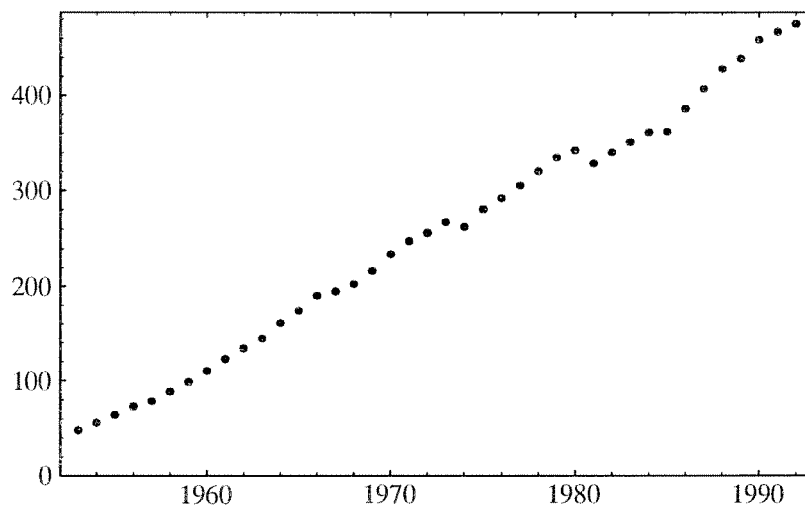
46.	Doden Nederland, Exp	84
47.	Doden Nederland, Exp0	84
48.	Doden Nederland, ExpV	85
49.	Doden Nederland, ExpV0	85
50.	Doden Nederland, Logis	86
51.	Doden Nederland, LogisV	86
52.	Doden Verenigde Staten, Exp	87
53.	Doden Verenigde Staten, Exp0	87
54.	Doden Verenigde Staten, ExpV	88
55.	Doden Verenigde Staten, ExpV0	88
56.	Doden Verenigde Staten, Logis	89
57.	Doden Verenigde Staten, LogisV	89
58.	Risico's West-Duitsland, Exp	90
59.	Risico's West-Duitsland, Exp0	90
60.	Risico's West-Duitsland, ExpV	91
61.	Risico's West-Duitsland, ExpV0	91
62.	Risico's West-Duitsland, Logis	92
63.	Risico's West-Duitsland, LogisV	92
64.	Risico's Engeland, Exp	93
65.	Risico's Engeland, Exp0	93
66.	Risico's Engeland, ExpV	94
67.	Risico's Engeland, ExpV0	94
68.	Risico's Engeland, Logis	95
69.	Risico's Engeland, LogisV	95
70.	Risico's Japan, Exp	96
71.	Risico's Japan, Exp0	96
72.	Risico's Japan, ExpV	97
73.	Risico's Japan, ExpV0	97
74.	Risico's Japan, Logis	98
75.	Risico's Japan, LogisV	98
76.	Risico's Nederland, Exp	99
77.	Risico's Nederland, Exp0	99
78.	Risico's Nederland, ExpV	100
79.	Risico's Nederland, ExpV0	100
80.	Risico's Nederland, Logis	101
81.	Risico's Nederland, LogisV	101
82.	Risico's Verenigde Staten, Exp	102
83.	Risico's Verenigde Staten, Exp0	102
84.	Risico's Verenigde Staten, ExpV	103
85.	Risico's Verenigde Staten, ExpV0	103
86.	Risico's Verenigde Staten, Logis	104
87.	Risico's Verenigde Staten, LogisV	104

A. Data

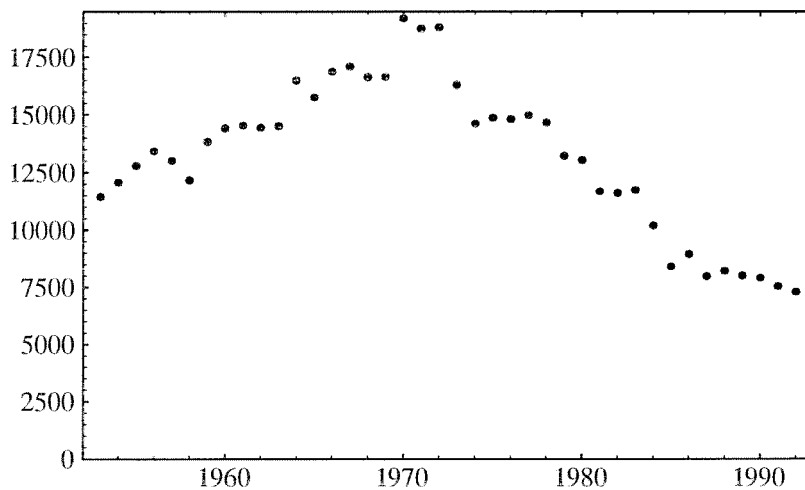
A.1. West-Duitsland

Jaar	Verkeers- prestatie	Doden	Jaar	Verkeers- prestatie	Doden
1953	48.2	11449	1973	266.7	16302
1954	56.0	12071	1974	262.1	14614
1955	64.7	12791	1975	280.2	14870
1956	73.4	13427	1976	292.2	14820
1957	78.6	13004	1977	305.1	14978
1958	88.7	12169	1978	320.3	14662
1959	98.7	13822	1979	334.7	13222
1960	110.1	14406	1980	342.0	13040
1961	122.6	14543	1981	328.2	11674
1962	134.1	14445	1982	340.2	11608
1963	144.6	14513	1983	350.8	11732
1964	161.1	16494	1984	360.8	10199
1965	174.1	15753	1985	361.5	8400
1966	190.3	16868	1986	385.8	8948
1967	194.7	17084	1987	406.6	7967
1968	202.2	16636	1988	427.4	8213
1969	216.3	16646	1989	438.5	7995
1970	234.2	19193	1990	458.5	7906
1971	248.1	18753	1991	466.9	7541
1972	256.0	18811	1992	475.2	7298

Tabel 20. Gebruikte gegevens voor West-Duitsland



Afbeelding 12. *Verkeersprestatie West-Duitsland*

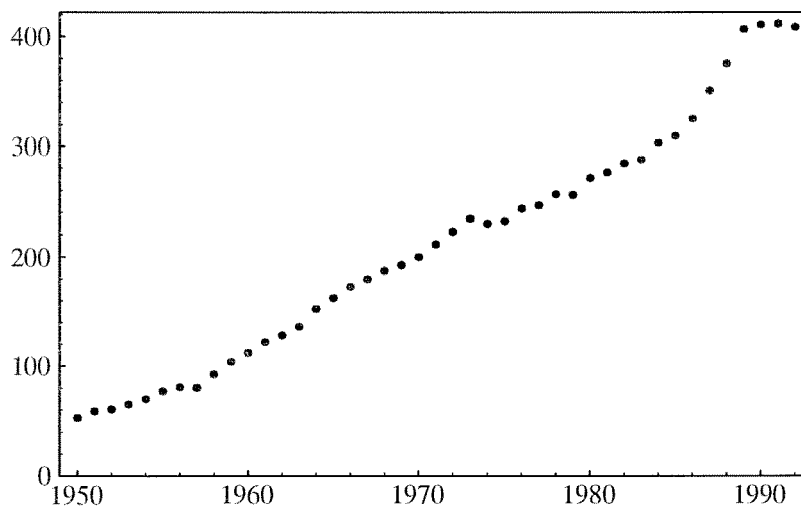


Afbeelding 13. *Verkeersdoden West-Duitsland*

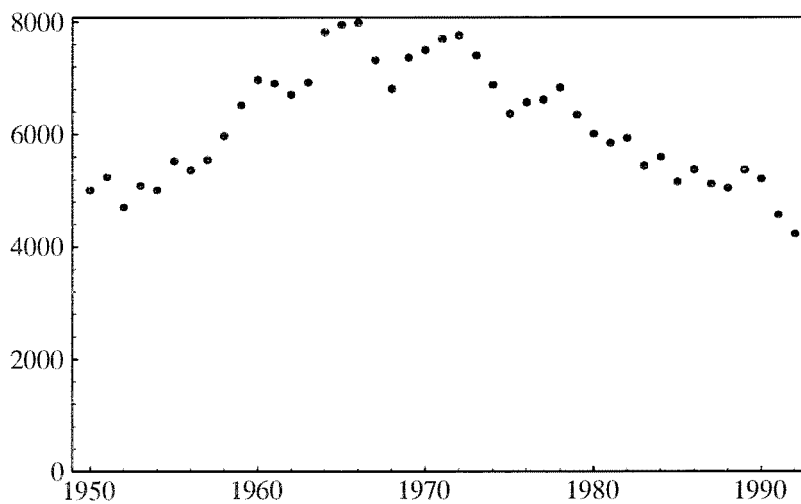
A.2. Engeland

Jaar	Verkeers- prestatie	Doden	Jaar	Verkeers- prestatie	Doden
1950	53.0	5012	1972	222.6	7763
1951	59.0	5250	1973	234.2	7406
1952	60.9	4706	1974	229.5	6883
1953	65.1	5090	1975	231.9	6366
1954	69.7	5010	1976	243.5	6570
1955	77.1	5526	1977	246.7	6614
1956	80.8	5367	1978	256.5	6831
1957	80.4	5550	1979	256.0	6352
1958	92.9	5970	1980	271.4	6010
1959	104.1	6520	1981	276.4	5846
1960	112.4	6970	1982	284.5	5934
1961	122.2	6908	1983	288.1	5445
1962	128.2	6709	1984	303.1	5599
1963	136.1	6922	1985	309.7	5165
1964	152.4	7820	1986	325.3	5382
1965	162.6	7952	1987	350.5	5125
1966	172.8	7985	1988	375.7	5052
1967	179.8	7319	1989	406.9	5373
1968	187.7	6810	1990	410.8	5217
1969	192.8	7365	1991	411.6	4568
1970	200.3	7499	1992	408.8	4229
1971	211.9	7699			

Tabel 21. *Gebruikte gegevens voor Engeland*



Afbeelding 14. *Verkeersprestatie Engeland*

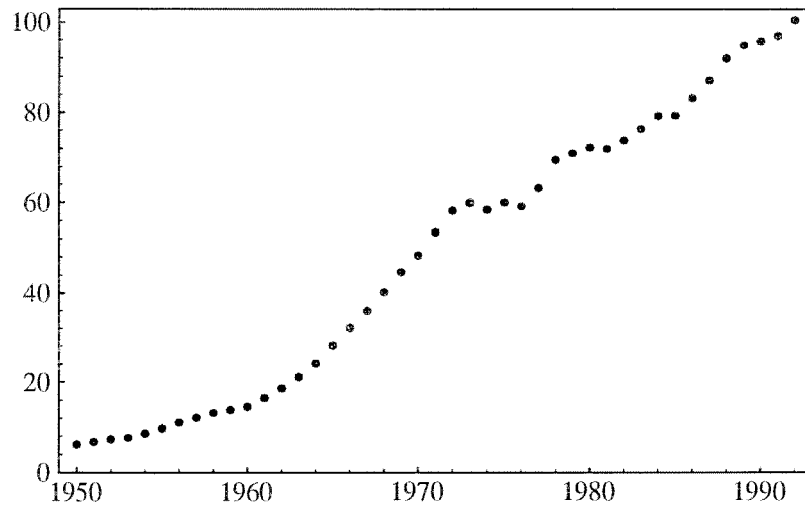


Afbeelding 15. *Verkeersdoden Engeland*

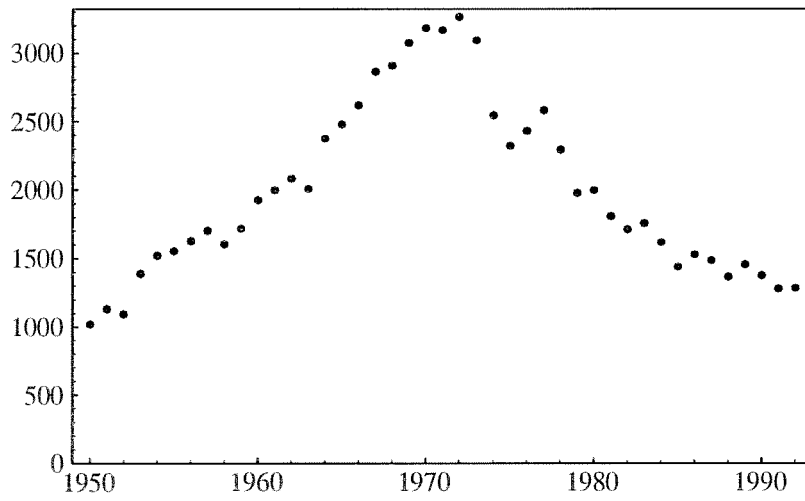
A.3. Nederland

Jaar	Verkeers- prestatie	Dodens		Jaar	Verkeers- prestatie	Dodens	
1950	6.2	1021	19533	1972	58.2	3264	70082
1951	6.8	1134	22159	1973	59.9	3092	70361
1952	7.3	1097	24133	1974	58.4	2546	66212
1953	7.7	1390	26945	1975	60.0	2321	59979
1954	8.6	1520	30441	1976	59.1	2432	62304
1955	9.7	1552	34097	1977	63.1	2583	64476
1956	11.1	1628	36635	1978	69.3	2294	62130
1957	12.2	1701	39514	1979	70.8	1977	56589
1958	13.2	1604	39782	1980	72.1	1997	56632
1959	13.8	1718	43526	1981	71.8	1807	53503
1960	14.6	1926	48358	1982	73.7	1710	52210
1961	16.5	1997	50171	1983	76.2	1756	52496
1962	18.6	2082	50520	1984	79.2	1615	50676
1963	21.1	2007	51216	1985	79.2	1438	48557
1964	24.1	2375	59187	1986	83.2	1528	50202
1965	28.2	2479	61887	1987	87.0	1485	49189
1966	32.2	2620	65304	1988	91.9	1366	47981
1967	36.0	2862	61102	1989	94.9	1456	50353
1968	40.2	2907	62098	1990	95.7	1376	52032
1969	44.6	3075	67599	1991	96.9	1281	47278
1970	48.4	3181	68225	1992	100.5	1285	48321
1971	53.4	3167	72167				

Tabel 22. Gebruikte gegevens voor Nederland



Afbeelding 16. *Verkeersprestatie Nederland*

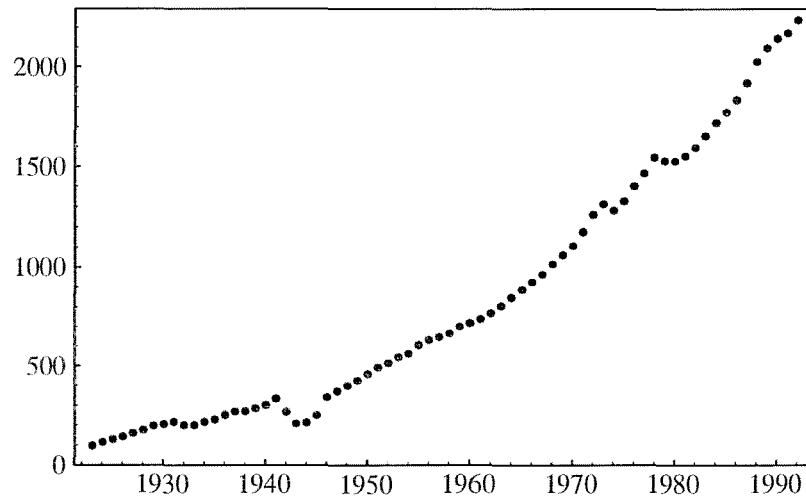


Afbeelding 17. *Verkeersdoden Nederland*

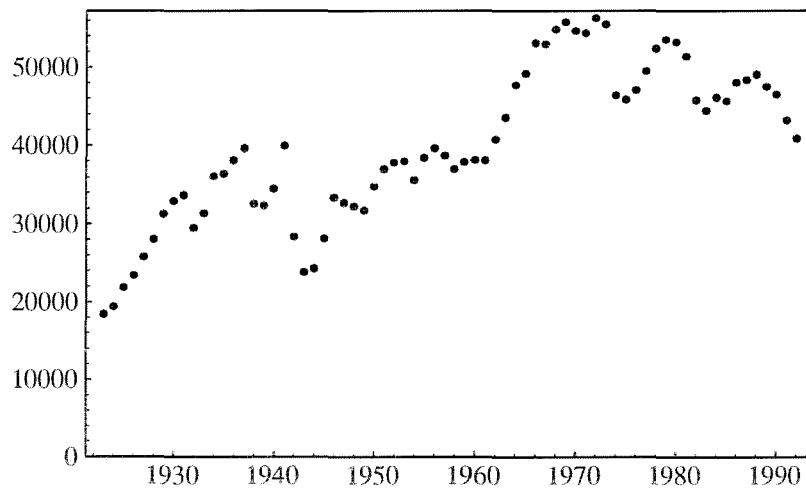
A.4. Verenigde Staten

Jaar	Verkeers- prestatie	Doden	Jaar	Verkeers- prestatie	Doden
1923	98	18394	1958	665	36981
1924	115	19380	1959	700	37910
1925	130	21877	1960	719	38137
1926	143	23431	1961	738	38091
1927	160	25796	1962	767	40804
1928	176	27996	1963	805	43564
1929	197	31215	1964	847	47700
1930	207	32929	1965	888	49163
1931	216	33675	1966	926	53041
1932	200	29451	1967	964	52924
1933	201	31363	1968	1016	54862
1934	216	36101	1969	1062	55791
1935	229	36369	1970	1110	54633
1936	252	38089	1971	1179	54381
1937	270	39643	1972	1260	56278
1938	271	32582	1973	1313	55511
1939	285	32386	1974	1281	46402
1940	302	34501	1975	1328	45853
1941	334	39969	1976	1402	47038
1942	268	28309	1977	1467	49510
1943	208	23823	1978	1545	52411
1944	213	24282	1979	1529	53524
1945	250	28076	1980	1527	53172
1946	341	33411	1981	1553	51385
1947	371	32697	1982	1595	45779
1948	398	32259	1983	1653	44452
1949	424	31701	1984	1720	46121
1950	458	34763	1985	1774	45671
1951	491	36996	1986	1835	48028
1952	514	37794	1987	1921	48344
1953	544	37955	1988	2026	49070
1954	562	35586	1989	2096	47502
1955	606	38426	1990	2144	46478
1956	631	39628	1991	2172	43208
1957	647	38702	1992	2237	40888

Tabel 23. Gebruikte gegevens voor U.S.A



Afbeelding 18. *Verkeersprestatie Verenigde Staten*

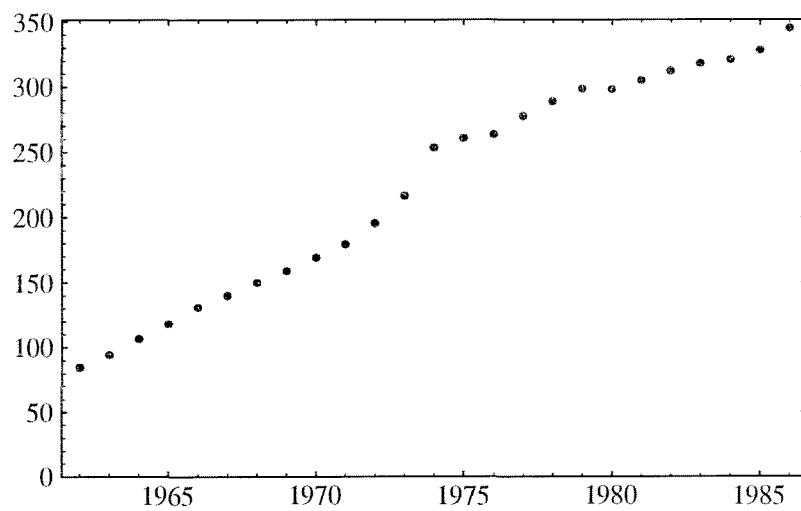


Afbeelding 19. *Verkeersdoden Verenigde Staten*

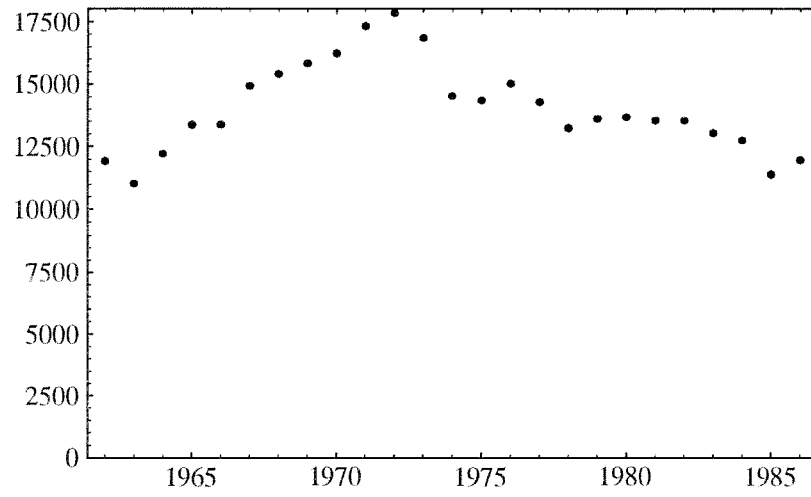
A.5. Frankrijk

Jaar	Verkeers- prestatie	Doden	Jaar	Verkeers- prestatie	Doden
1962	84.7	11921	1975	260.7	14355
1963	94.2	11030	1976	263.4	15027
1964	106.4	12216	1977	277.7	14283
1965	118.6	13365	1978	289.0	13229
1966	130.8	13374	1979	298.6	13603
1967	140.0	14944	1980	298.0	13672
1968	150.0	15416	1981	305.0	13546
1969	158.8	15837	1982	312.0	13526
1970	169.0	16237	1983	318.0	13021
1971	179.3	17346	1984	321.0	12735
1972	195.4	17869	1985	328.0	11388
1973	216.5	16861	1986	345.0	11947
1974	253.2	14526			

Tabel 24. Gebruikte gegevens voor Frankrijk



Afbeelding 20. Verkeersprestatie Frankrijk

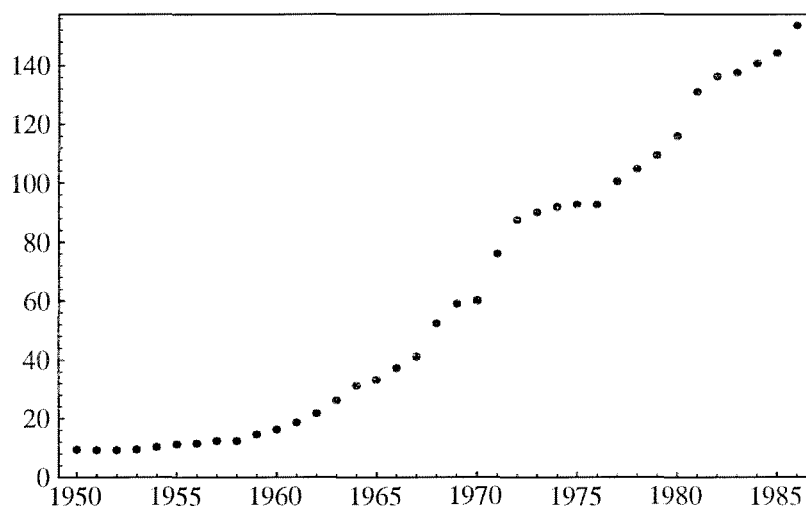


Afbeelding 21. *Verkeersdoden Frankrijk*

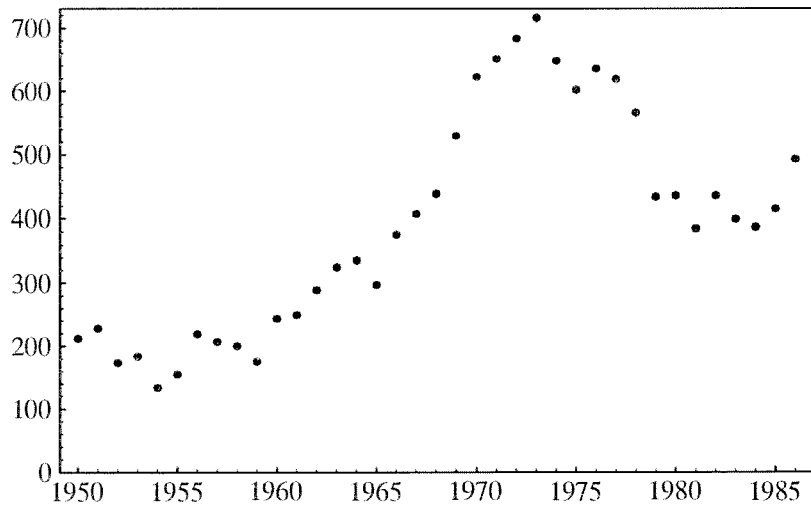
A.6. Israël

Jaar	Verkeersprestatie	Doden	Jaar	Verkeersprestatie	Doden
1950	9.5	212	1969	58.9	529
1951	9.3	228	1970	60.0	623
1952	9.2	174	1971	76.0	651
1953	9.5	184	1972	87.5	683
1954	10.4	134	1973	90.1	716
1955	11.3	155	1974	91.9	648
1956	11.6	219	1975	92.6	602
1957	12.4	207	1976	92.7	636
1958	12.4	200	1977	100.5	619
1959	14.6	176	1978	104.7	566
1960	16.2	244	1979	109.4	434
1961	18.6	250	1980	116.0	436
1962	21.7	289	1981	130.9	385
1963	26.3	325	1982	136.2	436
1964	31.3	336	1983	137.5	399
1965	33.3	297	1984	140.6	387
1966	37.2	376	1985	144.1	415
1967	41.2	407	1986	153.5	493
1968	52.3	439			

Tabel 25. Gebruikte gegevens voor Israël



Afbeelding 22. Verkeersprestatie Israël

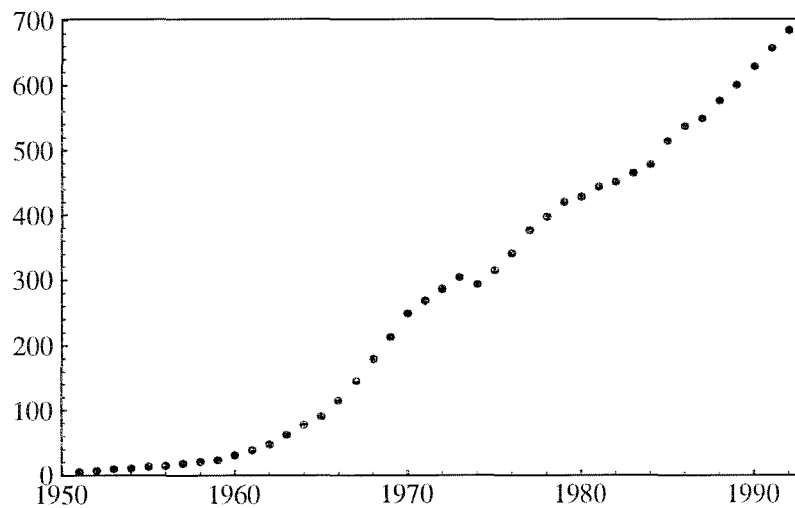


Afbeelding 23. *Verkeersdoden Israël*

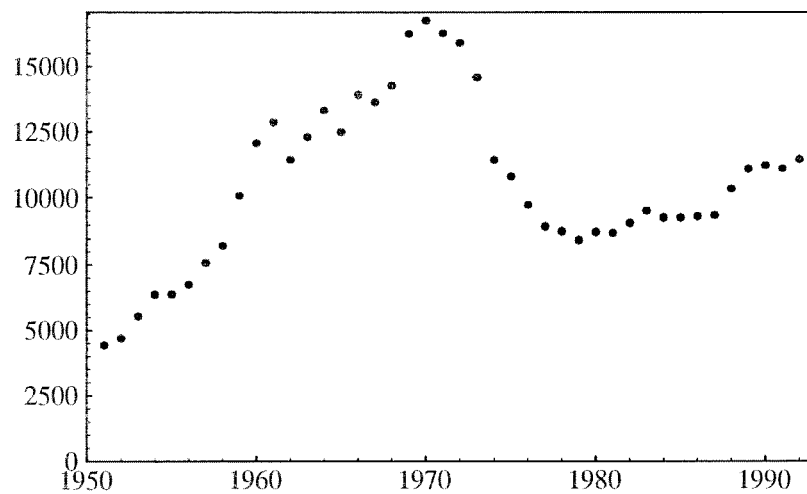
A.7. Japan

Jaar	Verkeers- prestatie	Doden	Jaar	Verkeers- prestatie	Doden
1951	5.5	4429	1972	285.5	15918
1952	7.4	4696	1973	303.8	14574
1953	9.8	5544	1974	293.1	11432
1954	11.0	6374	1975	314.9	10792
1955	13.2	6379	1976	340.6	9734
1956	15.3	6751	1977	376.5	8945
1957	18.4	7575	1978	397.4	8783
1958	21.1	8248	1979	420.2	8466
1959	23.6	10079	1980	428.0	8760
1960	31.0	12055	1981	444.1	8719
1961	39.0	12865	1982	451.2	9073
1962	47.4	11445	1983	465.2	9520
1963	61.9	12301	1984	478.1	9262
1964	77.3	13318	1985	514.3	9261
1965	90.4	12484	1986	536.9	9317
1966	115.1	13904	1987	548.8	9347
1967	145.4	13618	1988	575.6	10344
1968	179.7	14256	1989	600.2	11086
1969	212.9	16257	1990	628.6	11227
1970	248.6	16765	1991	657.3	11105
1971	267.8	16278	1992	685.0	11451

Tabel 26. Gebruikte gegevens voor Japan



Afbeelding 24. Verkeersprestatie Japan

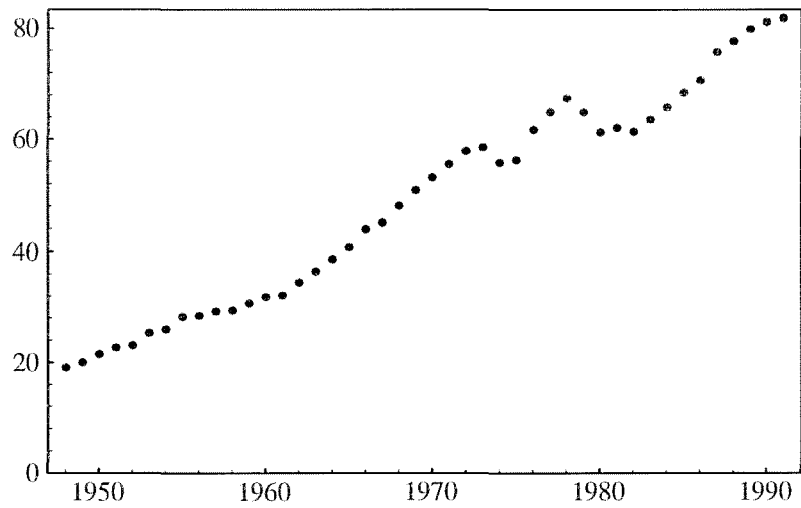


Afbeelding 25. Verkeersdoden Japan

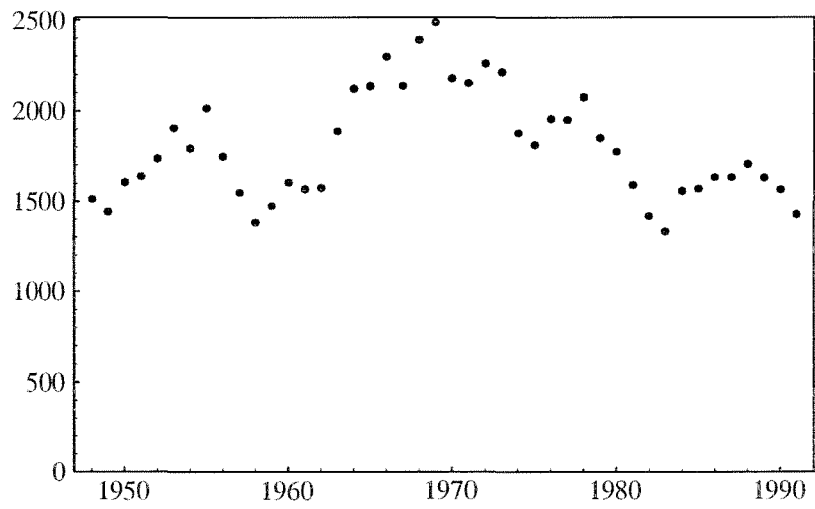
A.8. Michigan

Jaar	Verkeers- prestatie	Doden	Jaar	Verkeers- prestatie	Doden
1948	19.1	1512	1970	53.1	2177
1949	20.0	1441	1971	55.5	2152
1950	21.5	1605	1972	57.8	2258
1951	22.7	1640	1973	58.5	2213
1952	23.1	1736	1974	55.7	1875
1953	25.3	1905	1975	56.3	1811
1954	26.0	1793	1976	61.6	1955
1955	28.3	2016	1977	64.9	1950
1956	28.4	1746	1978	67.4	2076
1957	29.3	1548	1979	64.9	1849
1958	29.4	1382	1980	61.2	1774
1959	30.7	1473	1981	62.0	1589
1960	31.8	1604	1982	61.3	1417
1961	32.1	1567	1983	63.6	1331
1962	34.5	1574	1984	65.7	1556
1963	36.5	1887	1985	68.4	1569
1964	38.6	2122	1986	70.6	1632
1965	40.9	2136	1987	75.7	1632
1966	43.9	2298	1988	77.7	1704
1967	45.1	2137	1989	79.9	1630
1968	48.0	2392	1990	81.2	1563
1969	50.9	2487	1991	81.9	1425

Tabel 27. Gebruikte gegevens voor Michigan



Afbeelding 26. Verkeersprestatie Michigan



Afbeelding 27. Verkeersdoden Michigan

B. Geschatte parameterwaarden

B.1. Gewone kleinste kwadraten

West-Duitsland					
	a	b	c	d	e
Exp0	-65.03	132.38			
Exp	-52.66	108.21	-13.16		
ExpV0	-72.07	146.47			1.13
ExpV	-27.40	58.48	-38.10		0.77
Logis	-52.66	108.21	-13.16	0.00	
LogisV	-231.80	461.89	8.31	0.21	0.33

Tabel 28a. *Parameter waarden OLS modellen*

Engeland					
	a	b	c	d	e
Exp0	-49.50	101.09			
Exp	-36.96	76.68	-10.55		
ExpV0	-60.19	122.55			1.23
ExpV	-32.80	68.48	-12.84		0.95
Logis	-36.96	76.68	-10.55	0.00	
LogisV	-32.80	68.48	-12.84	0.00	0.95

Tabel 28b. *Parameter waarden OLS modellen*

Japan					
	a	b	c	d	e
Exp0	-111.07	223.12			
Exp	-161.98	323.06	13.56		
ExpV0	-49.53	100.91			0.46
ExpV	-216.56	431.13	18.38		1.32
Logis	-222.60	442.47	16.44	0.03	
LogisV	-221.26	439.80	16.22	0.04	0.98

Tabel 28c. *Parameter waarden OLS modellen*

Nederland					
	a	b	c	d	e
Exp0	-69.64	141.28			
Exp	-58.62	119.73	-9.01		
ExpV0	-87.59	177.48			1.25
ExpV	-96.63	195.48	4.71		1.33
Logis	-119.14	239.05	7.08	0.07	
LogisV	-117.68	236.28	7.82	0.06	1.05

Tabel 28d. *Parameter waarden OLS modellen*

Verenigde Staten					
	a	b	c	d	e
Exp0	-32.42	67.63			
Exp	-28.94	60.93	-6.32		
ExpV0	-36.98	76.63			1.11
ExpV	-27.62	58.35	-7.32		0.98
Logis	-28.94	60.93	-6.32	0.00	
LogisV	-27.62	58.35	-7.32	0.00	0.98

Tabel 28e. *Parameter waarden OLS modellen*

B.2. Gewogen kleinste kwadraten

West-Duitsland					
	a	b	c	d	e
Exp0	-66.27	134.81			
Exp	-54.54	111.89	-11.27		
ExpV0	-73.16	148.60			1.13
ExpV	-32.89	69.18	-27.47		0.79
Logis	-54.54	111.89	-11.27	0.00	
LogisV	-234.58	467.40	8.21	0.22	0.32

Tabel 29a. Parameter waarden WLS modellen

Engeland					
	a	b	c	d	e
Exp0	-49.60	101.29			
Exp	-36.67	76.12	-10.84		
ExpV0	-61.08	124.33			1.25
ExpV	-31.83	66.58	-13.55		0.95
Logis	-36.67	76.12	-10.84	0.00	
LogisV	-31.83	66.58	-13.55	0.00	0.95

Tabel 29b. Parameter waarden WLS modellen

Japan					
	a	b	c	d	e
Exp0	-114.18	229.26			
Exp	-156.83	312.90	13.06		
ExpV0	-48.92	99.69			0.48
ExpV	-225.36	448.63	19.33		1.40
Logis	-227.49	452.09	16.49	0.04	
LogisV	-225.36	447.82	16.02	0.04	0.95

Tabel 29c. Parameter waarden WLS modellen

Nederland					
	a	b	c	d	e
Exp0	-68.51	139.02			
Exp	-57.62	117.75	-8.86		
ExpV0	-87.38	177.11			1.27
ExpV	-98.28	198.83	5.64		1.36
Logis	-121.78	244.24	7.53	0.07	
LogisV	-121.50	243.72	7.64	0.07	1.01

Tabel 29d. *Parameter waarden WLS modellen*

Verenigde Staten					
	a	b	c	d	e
Exp0	-32.35	67.49			
Exp	-29.90	62.78	-4.44		
ExpV0	-33.46	69.67			1.03
ExpV	-25.43	54.04	-7.78		0.92
Logis	-29.90	62.78	-4.44	0.00	
LogisV	-25.43	54.04	-7.78	0.00	0.92

Tabel 29e. *Parameter waarden WLS modellen*

B.3. Maximum likelihood

West-Duitsland					
	a	b	c	d	e
Exp0	-66.16	134.61			
Exp	-55.01	112.82	-10.91		
ExpV0	-73.02	148.33			1.13
ExpV	-33.44	70.26	-26.74		0.79
Logis	-55.01	112.82	-10.91	0.00	
LogisV	-234.43	467.10	8.22	0.22	0.32

Tabel 30a. *Parameter waarden Poisson modellen*

Engeland					
	a	b	c	d	e
Exp0	-49.53	101.15			
Exp	-36.64	76.06	-10.92		
ExpV0	-60.88	123.94			1.24
ExpV	-32.22	67.34	-13.37		0.95
Logis	-36.64	76.06	-10.92	0.00	
LogisV	-32.22	67.34	-13.37	0.00	0.95

Tabel 30b. *Parameter waarden Poisson modellen*

Japan					
	a	b	c	d	e
Exp0	-106.44	214.09			
Exp	-159.48	318.15	13.41		
ExpV0	-45.92	93.76			0.45
ExpV	-221.70	441.36	18.87		1.37
Logis	-226.83	450.80	16.51	0.04	
LogisV	-224.00	445.12	15.91	0.04	0.94

Tabel 30c. *Parameter waarden Poisson modellen*

Nederland					
	a	b	c	d	e
Exp0	-69.13	140.26			
Exp	-58.65	119.77	-8.10		
ExpV0	-86.48	175.28			1.25
ExpV	-96.38	195.01	4.83		1.34
Logis	-120.64	242.00	7.38	0.07	
LogisV	-121.03	242.74	7.24	0.07	0.99

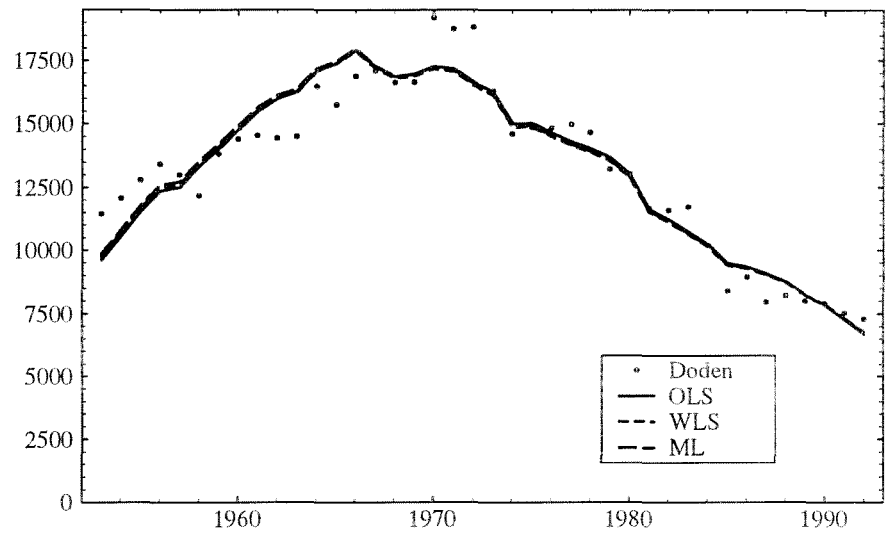
Tabel 30d. *Parameter waarden Poisson modellen*

Verenigde Staten					
	a	b	c	d	e
Exp0	-32.42	67.63			
Exp	-29.54	62.09	-5.40		
ExpV0	-33.59	69.94			1.03
ExpV	-25.24	53.70	-8.83		0.93
Logis	-29.54	62.09	-5.40	0.00	
LogisV	-25.24	53.70	-8.83	0.00	0.93

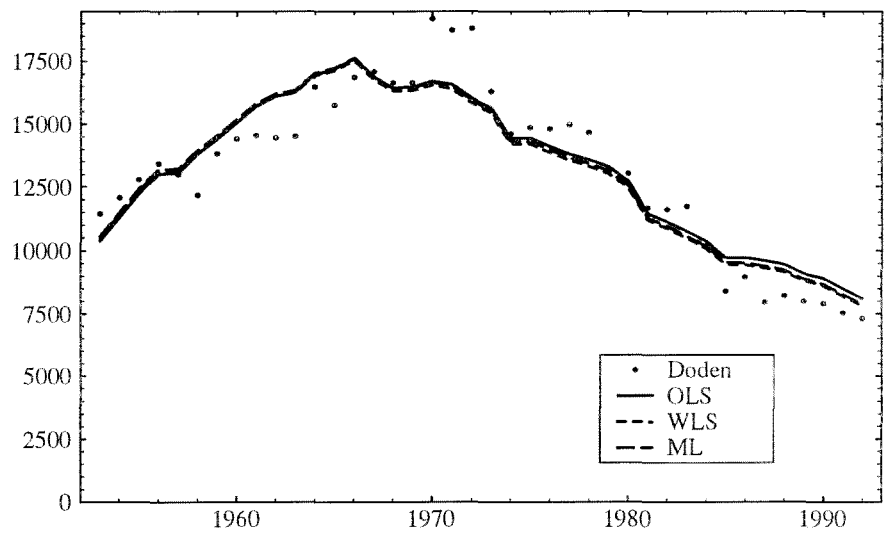
Tabel 30e. *Parameter waarden Poisson modellen*

C. Afbeeldingen modellen voor doden

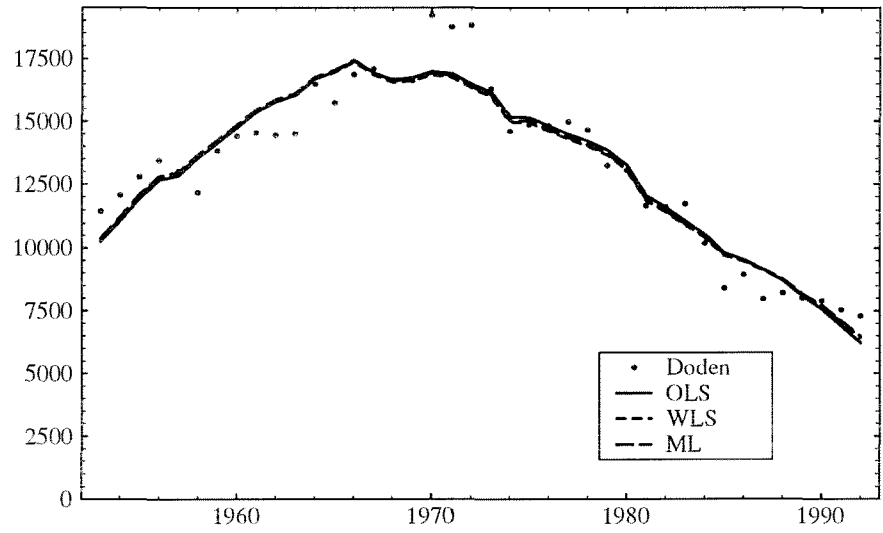
C.1. Doden West-Duitsland



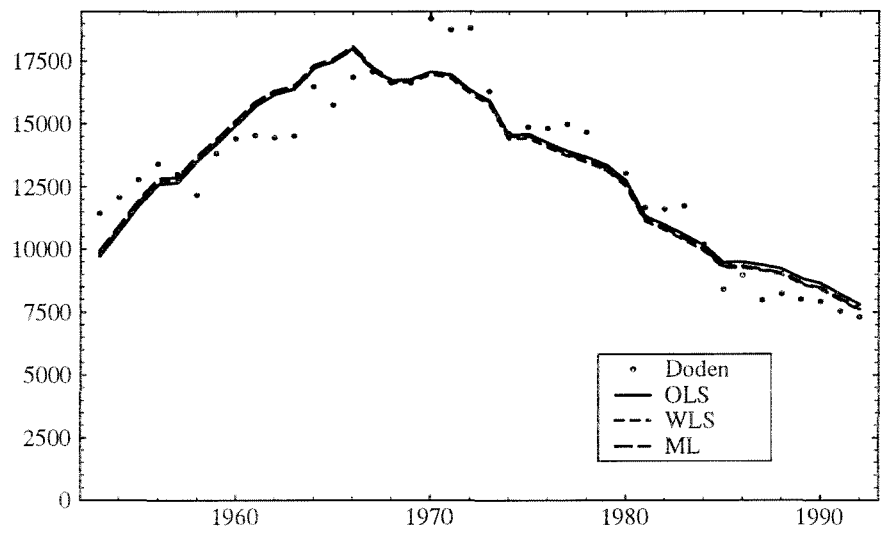
Afbeelding 28. *Doden West-Duitsland, Exp*



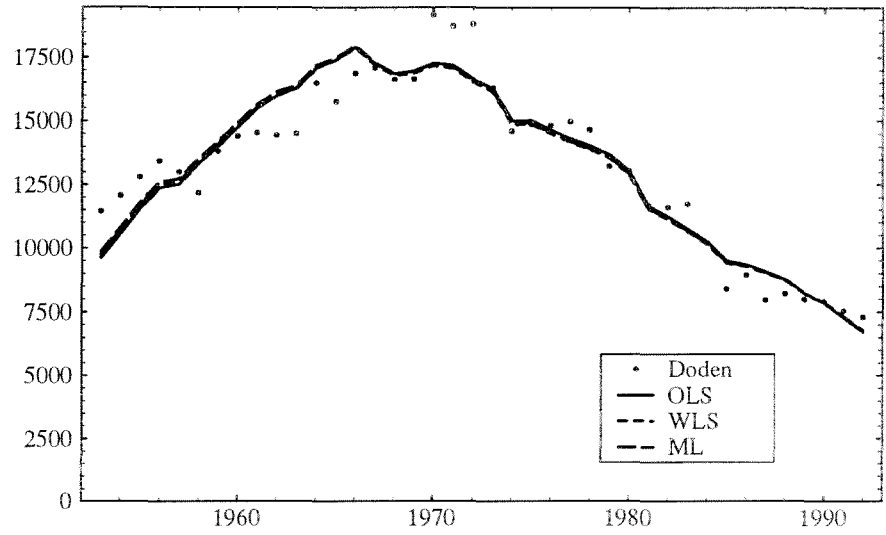
Afbeelding 29. *Doden West-Duitsland, Exp0*



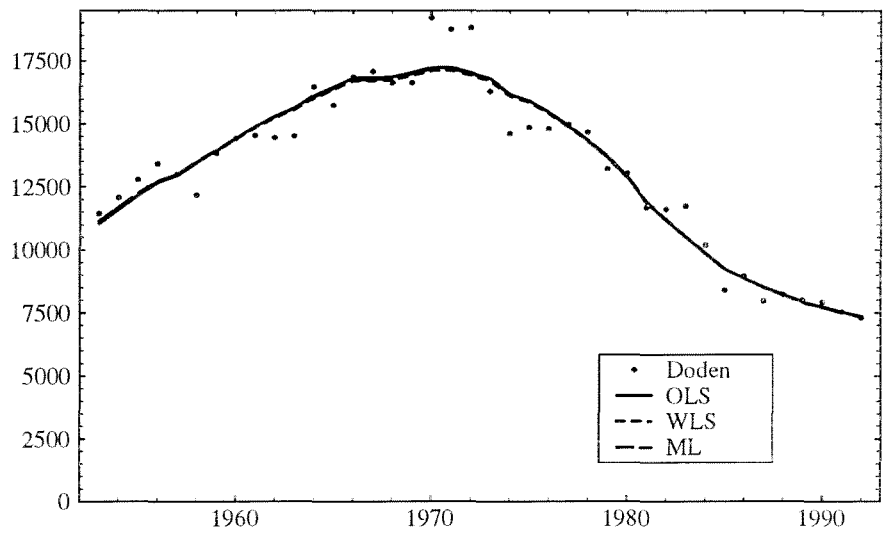
Afbeelding 30. Doden West-Duitsland, $ExpV$



Afbeelding 31. Doden West-Duitsland, $ExpV0$

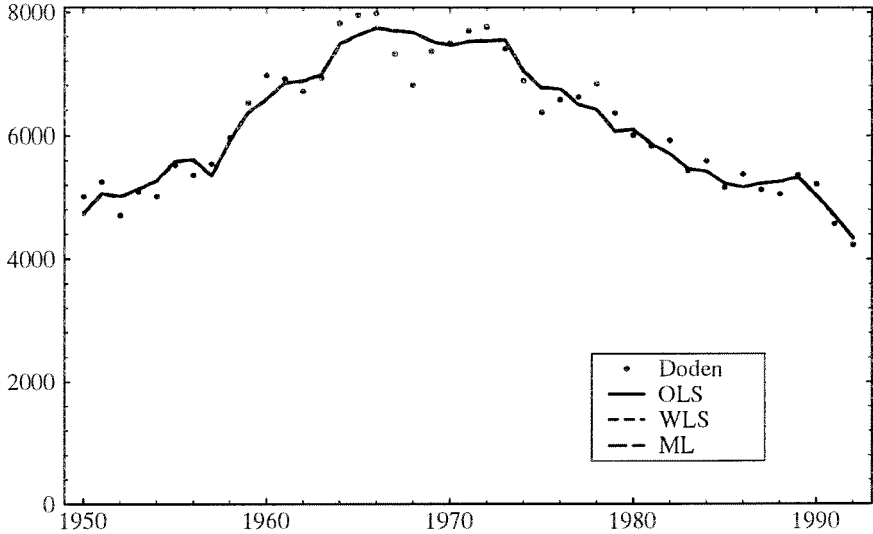


Afbeelding 32. *Doden West-Duitsland, Logis*

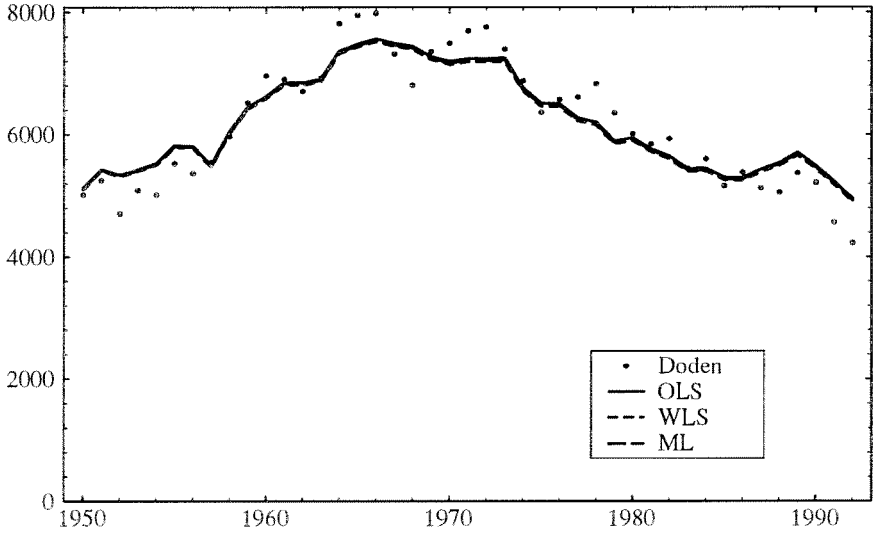


Afbeelding 33. *Doden West-Duitsland, LogisV*

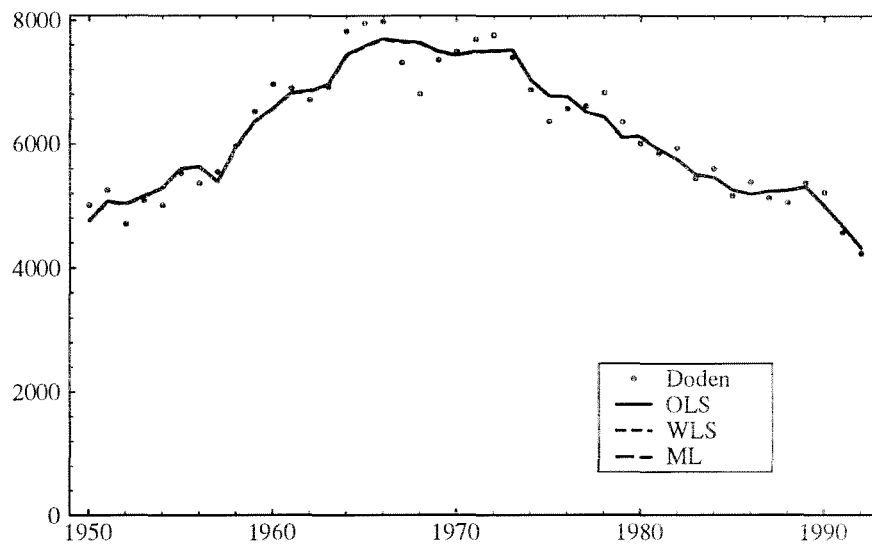
C.2. Doden Engeland



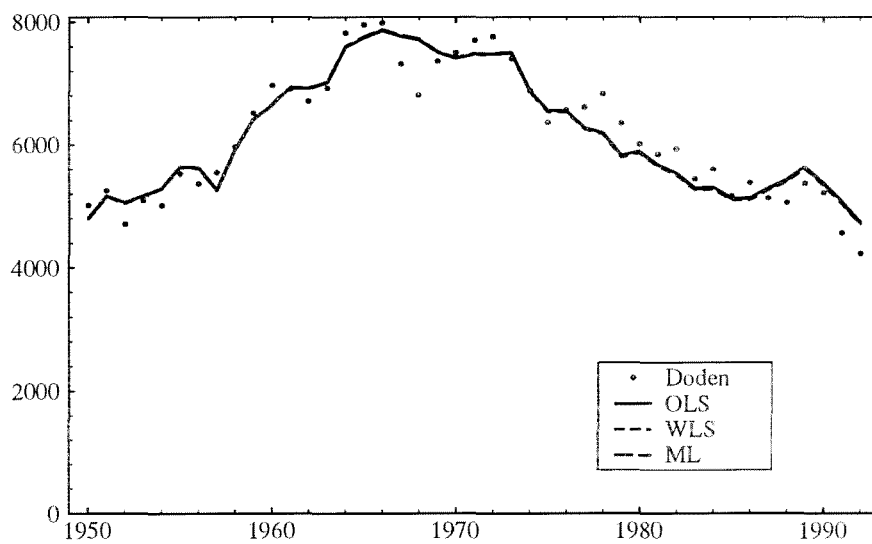
Afbeelding 34. Doden Engeland, Exp



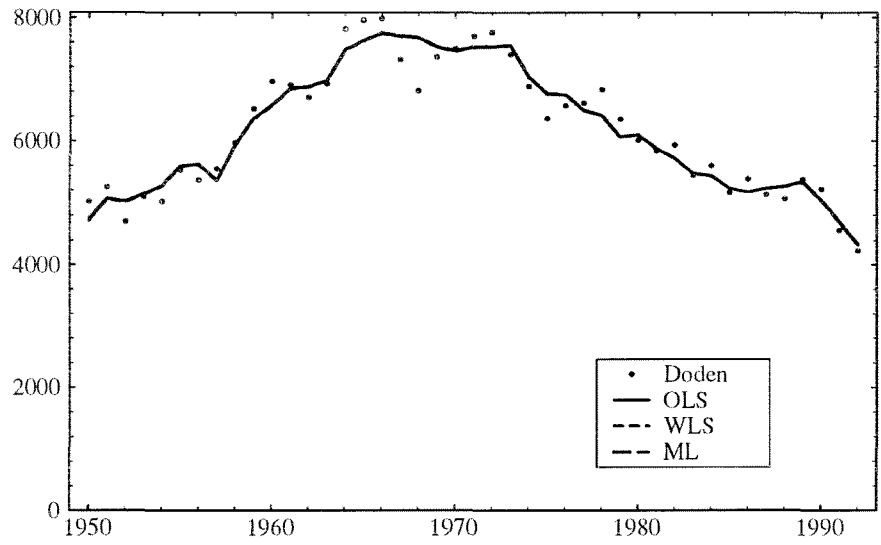
Afbeelding 35. Doden Engeland, Exp0



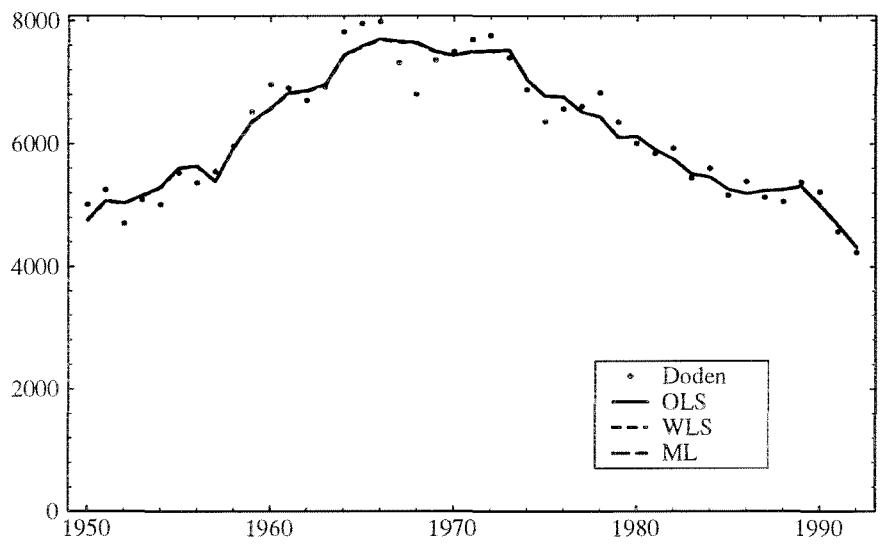
Afbeelding 36. Doden Engeland, ExpV



Afbeelding 37. Doden Engeland, ExpV0

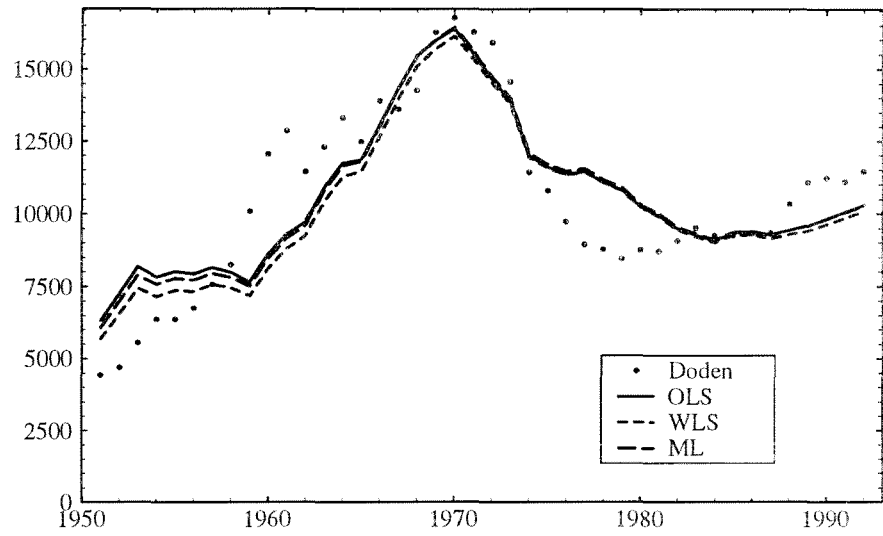


Afbeelding 38. Doden Engeland, Logis

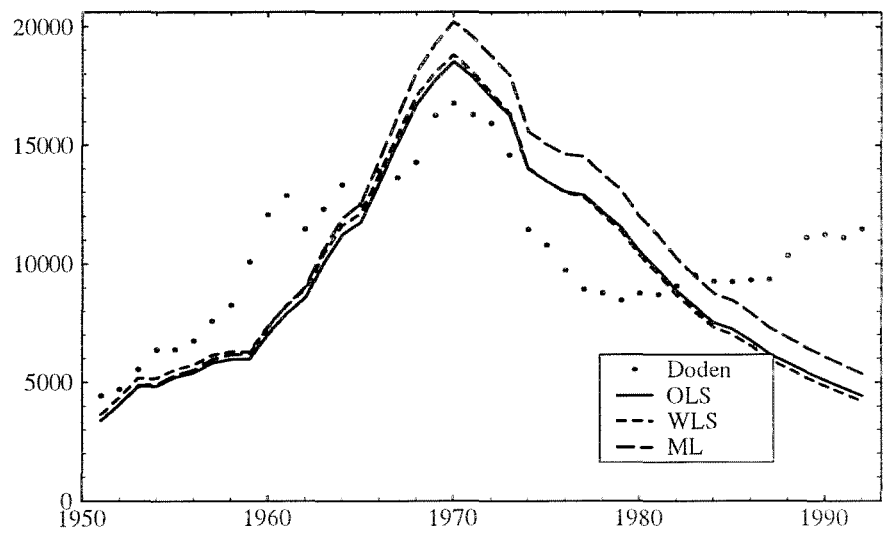


Afbeelding 39. Doden Engeland, LogisV

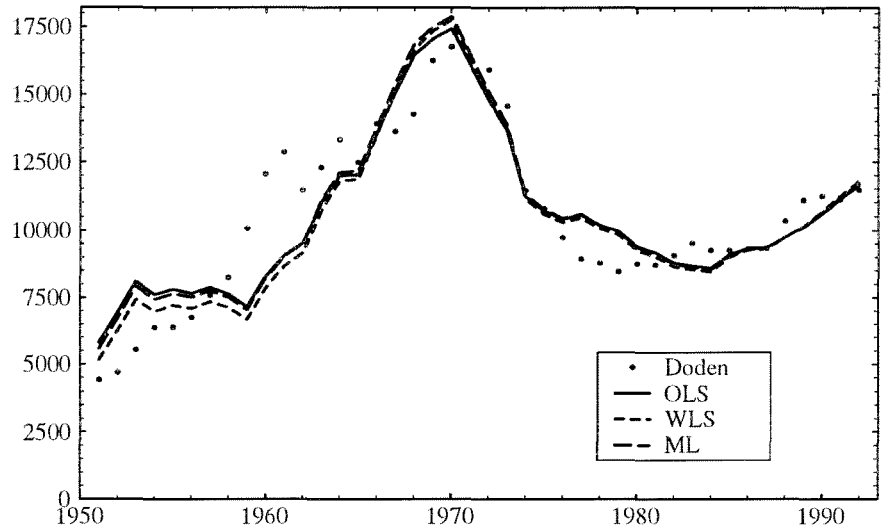
C.3. Doden Japan



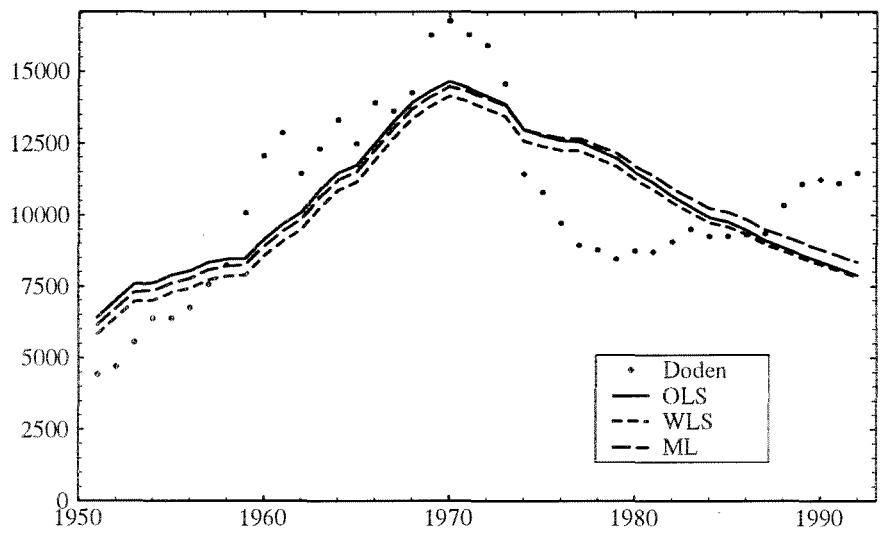
Afbeelding 40. *Doden Japan, Exp*



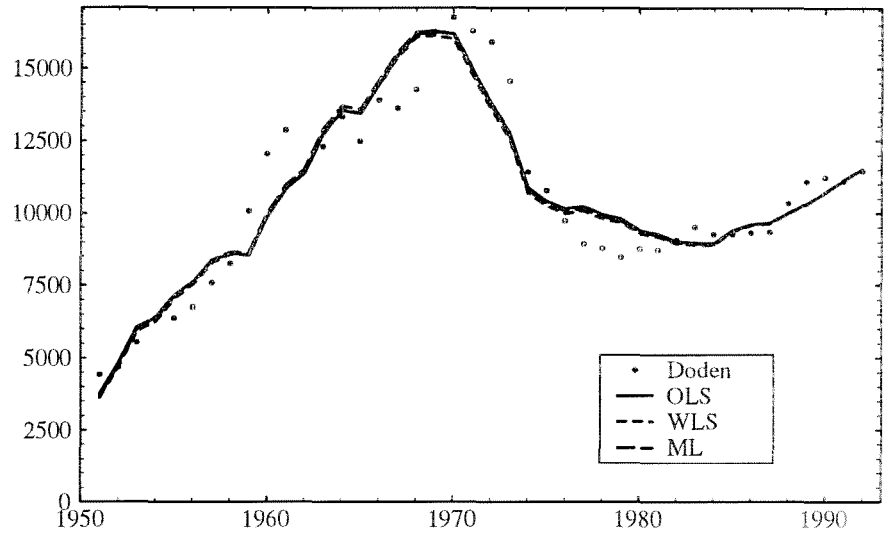
Afbeelding 41. *Doden Japan, Exp0*



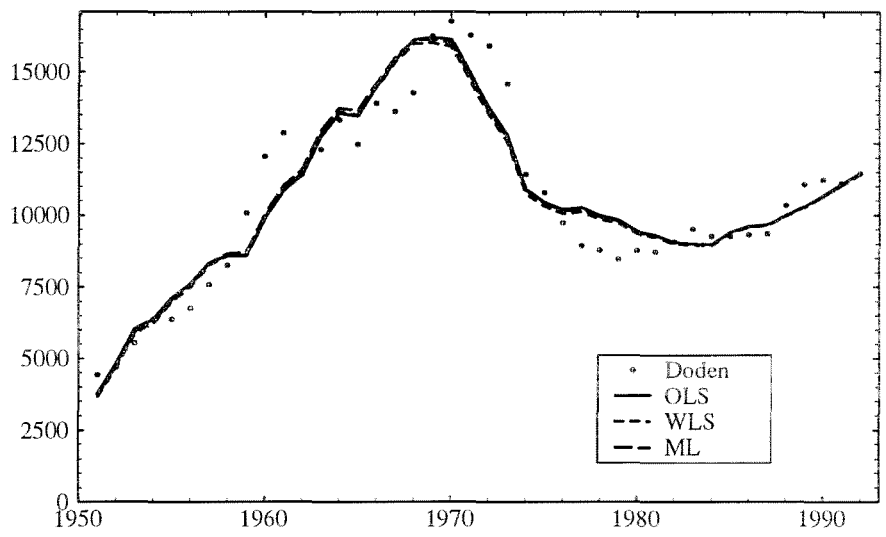
Afbeelding 42. Doden Japan, $ExpV$



Afbeelding 43. Doden Japan, $ExpV0$

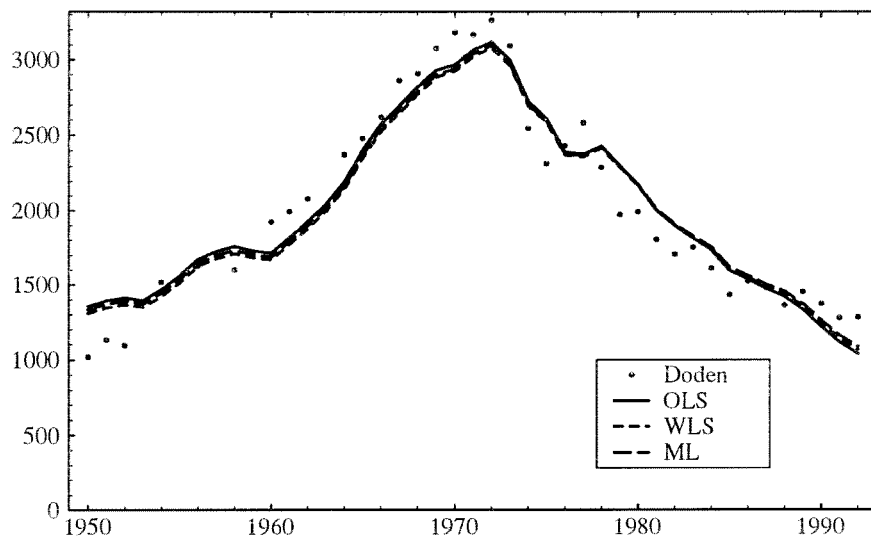


Afbeelding 44. *Doden Japan, Logis*

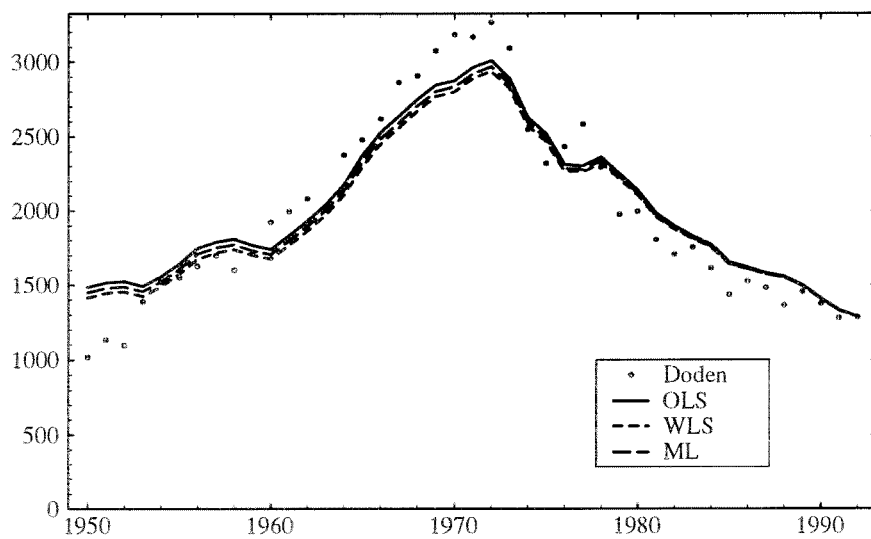


Afbeelding 45. *Doden Japan, LogisV*

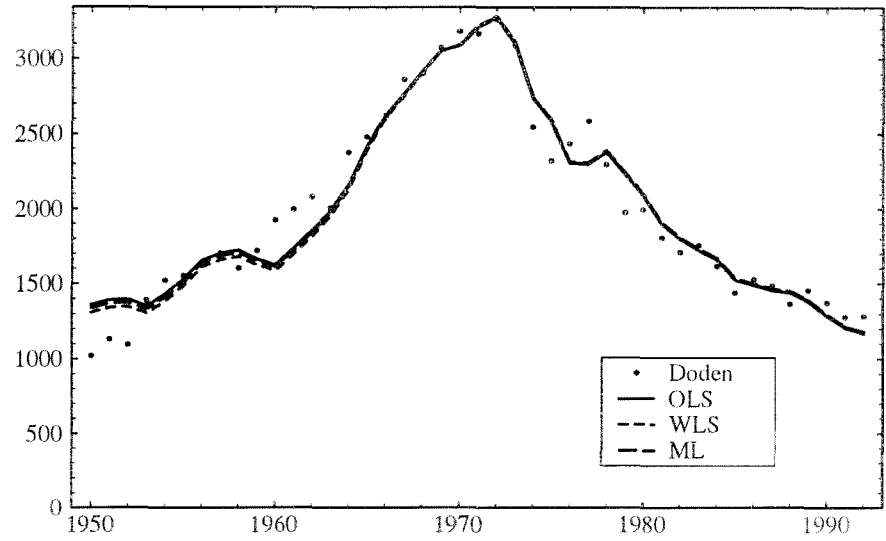
C.4. Doden Nederland



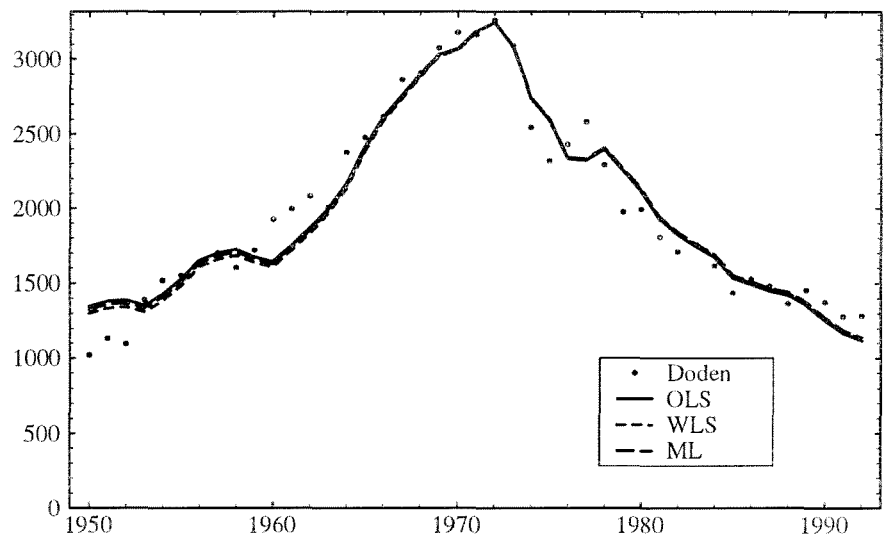
Afbeelding 46. Doden Nederland, Exp



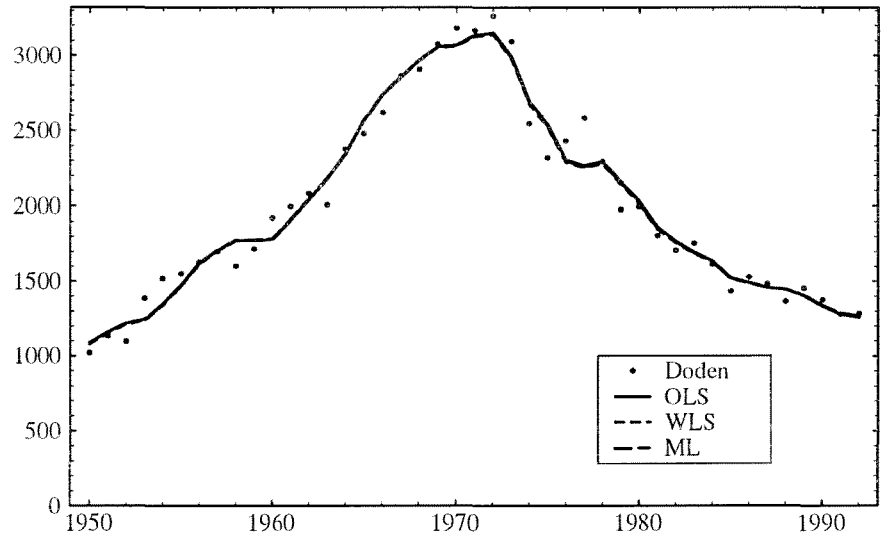
Afbeelding 47. Doden Nederland, Exp0



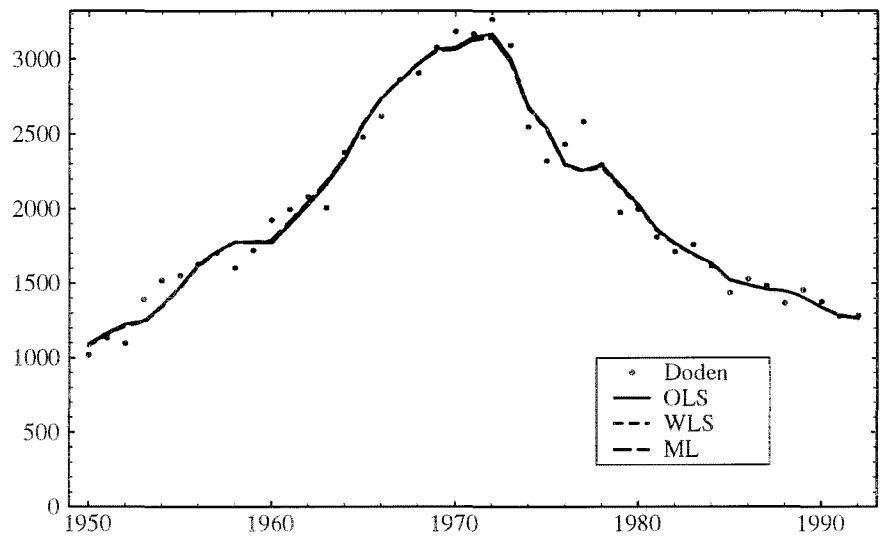
Afbeelding 48. *Doden Nederland, ExpV*



Afbeelding 49. *Doden Nederland, ExpV0*

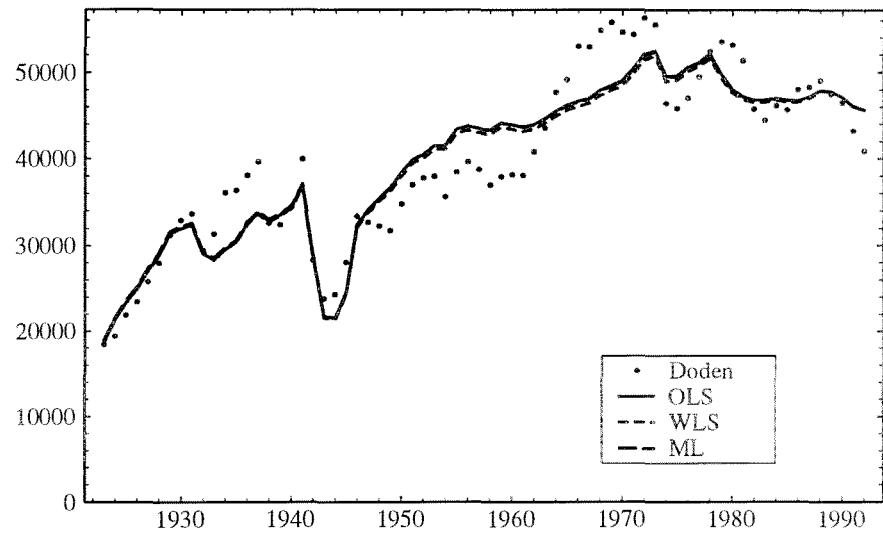


Afbeelding 50. Doden Nederland, Logis

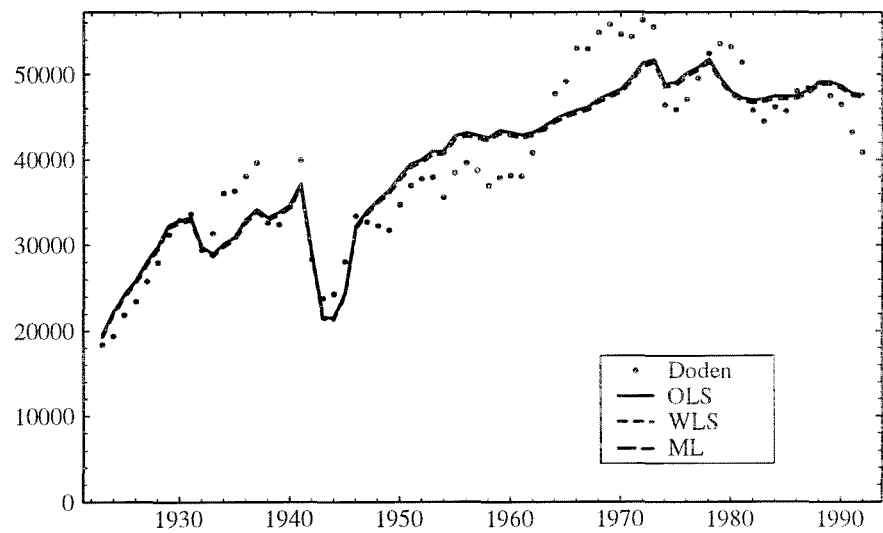


Afbeelding 51. Doden Nederland, LogisV

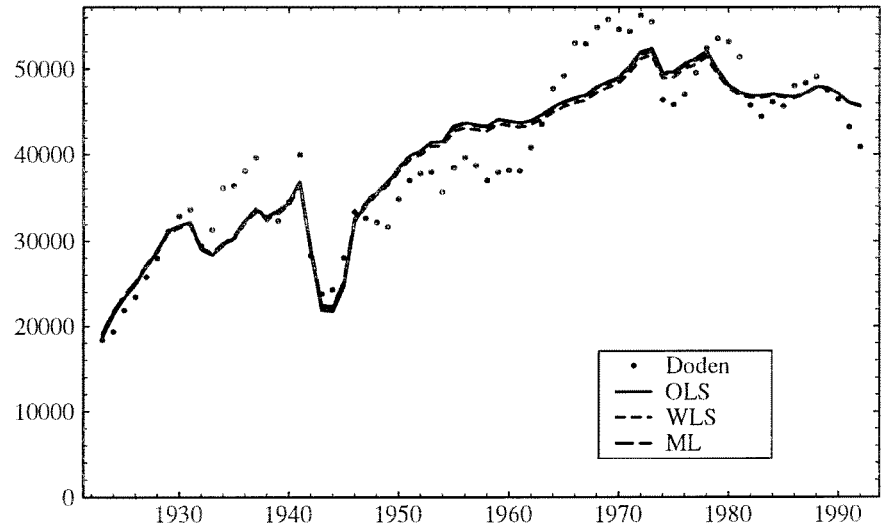
C.5. Doden Verenigde Staten



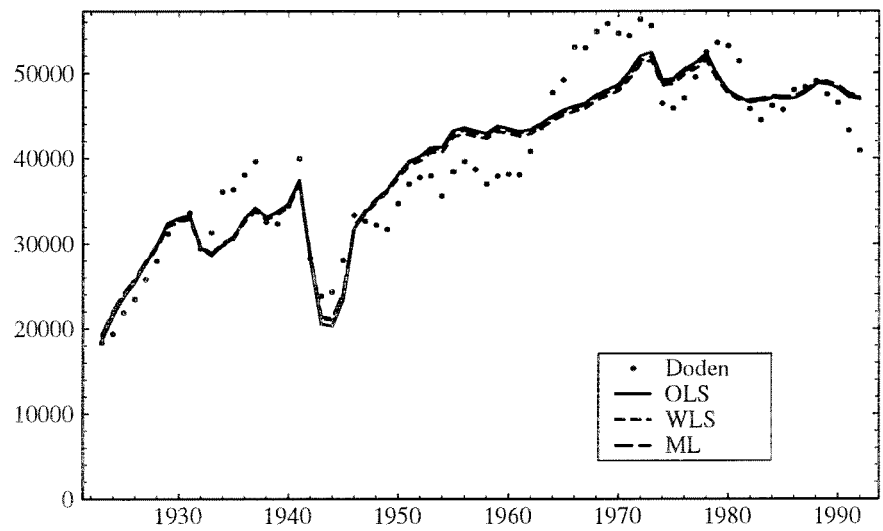
Afbeelding 52. *Doden Verenigde Staten, Exp*



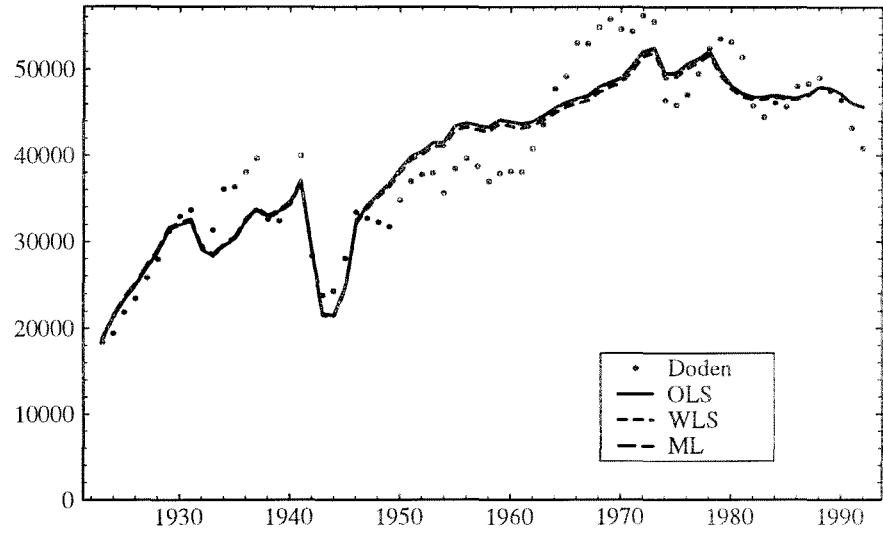
Afbeelding 53. *Doden Verenigde Staten, Exp0*



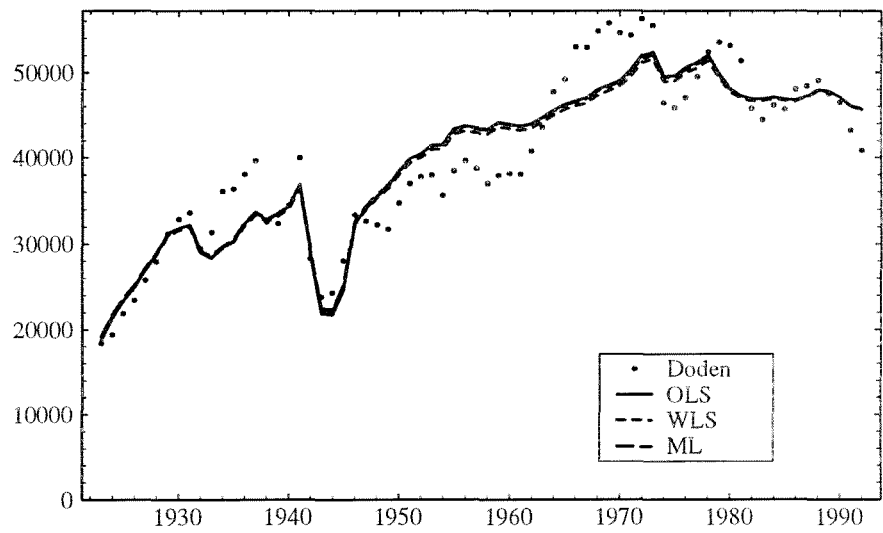
Afbeelding 54. Dodan Verenigde Staten, ExpV



Afbeelding 55. Dodan Verenigde Staten, ExpV0



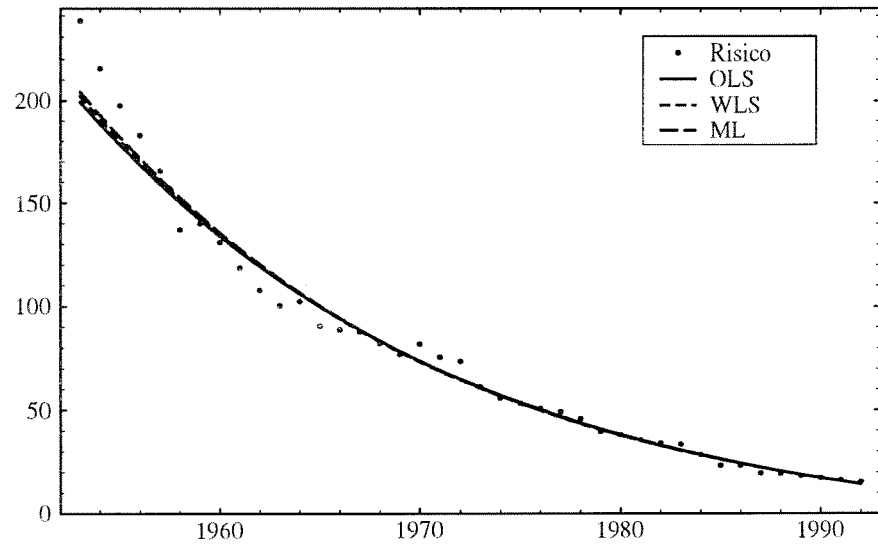
Afbeelding 56. Doden Verenigde Staten, Logis



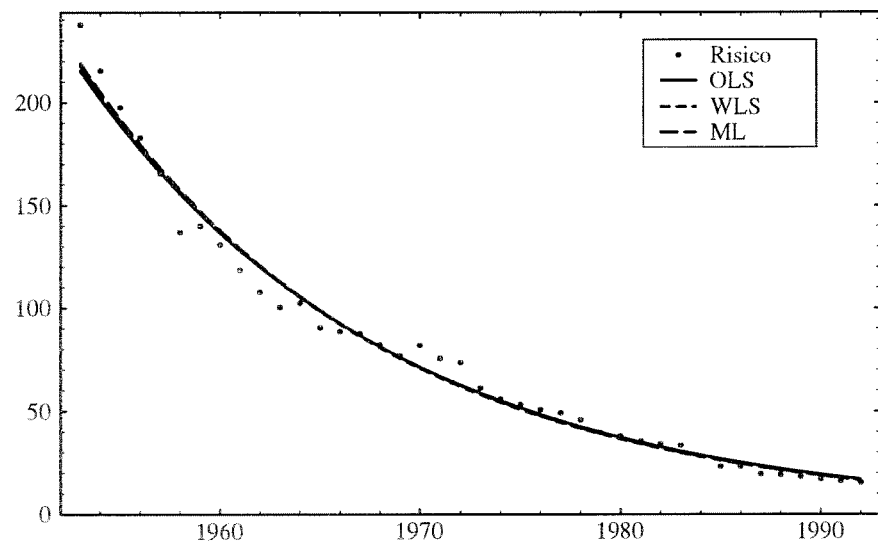
Afbeelding 57. Doden Verenigde Staten, LogisV

D. Afbeeldingen modellen voor Risico's

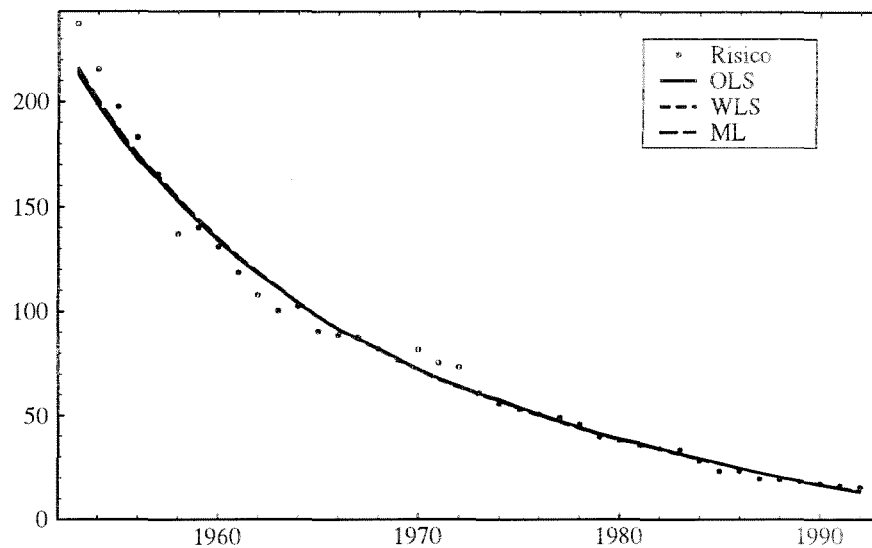
D.1. Risico's West-Duitsland



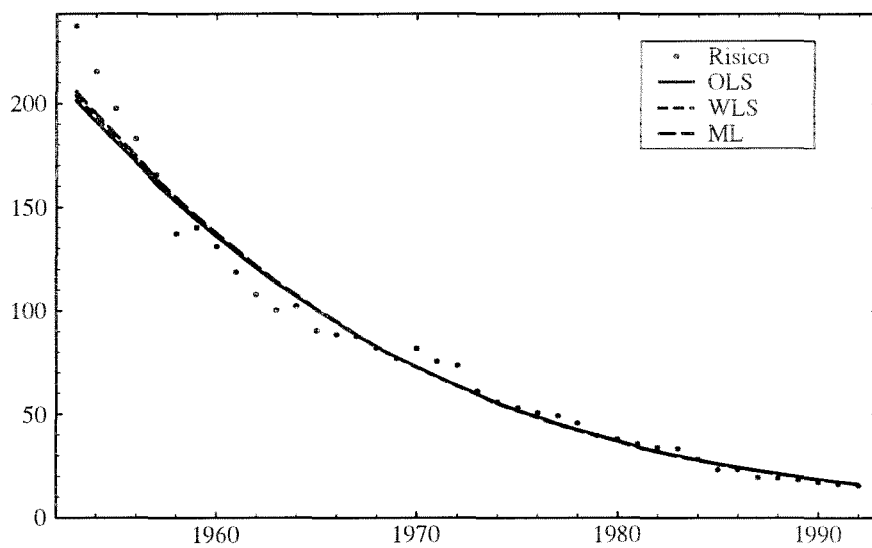
Afbeelding 58. *Risico's West-Duitsland, Exp*



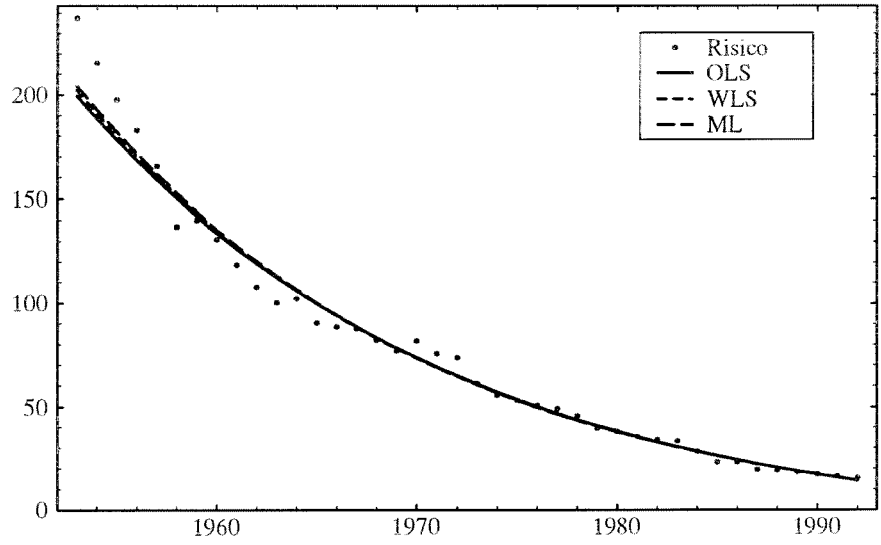
Afbeelding 59. *Risico's West-Duitsland, Exp0*



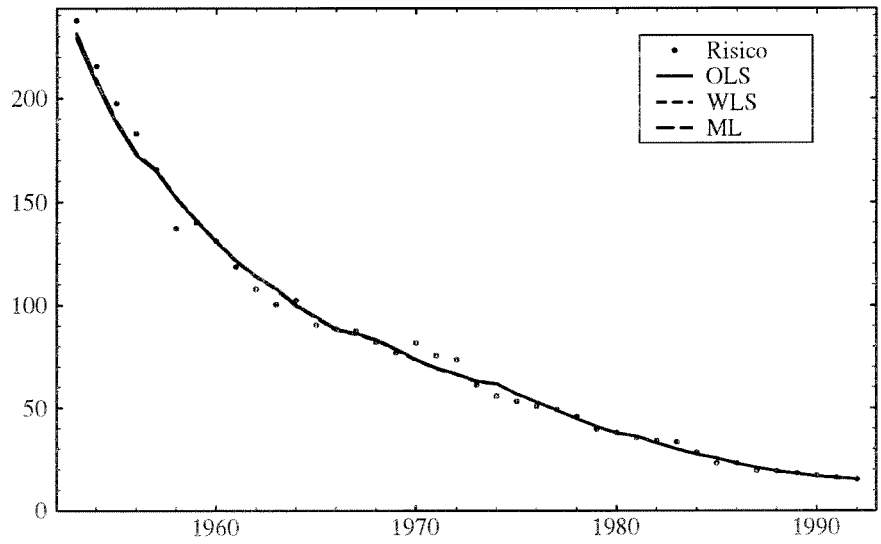
Afbeelding 60. *Risico's West-Duitsland, ExpV*



Afbeelding 61. *Risico's West-Duitsland, ExpV0*

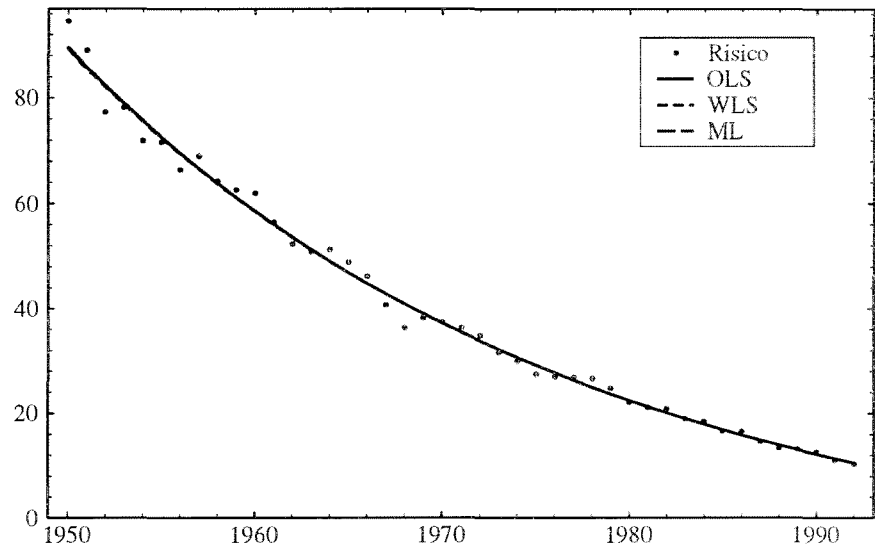


Afbeelding 62. *Risico's West-Duitsland, Logis*

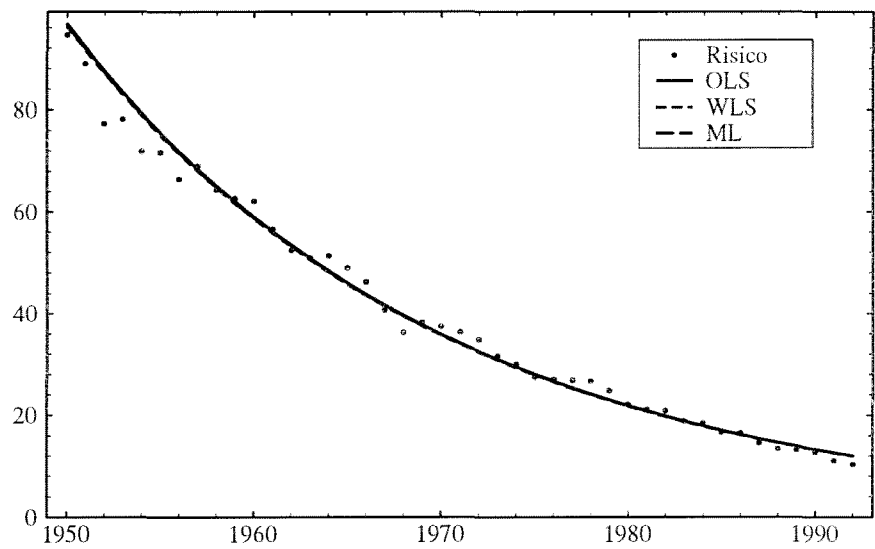


Afbeelding 63. *Risico's West-Duitsland, LogisV*

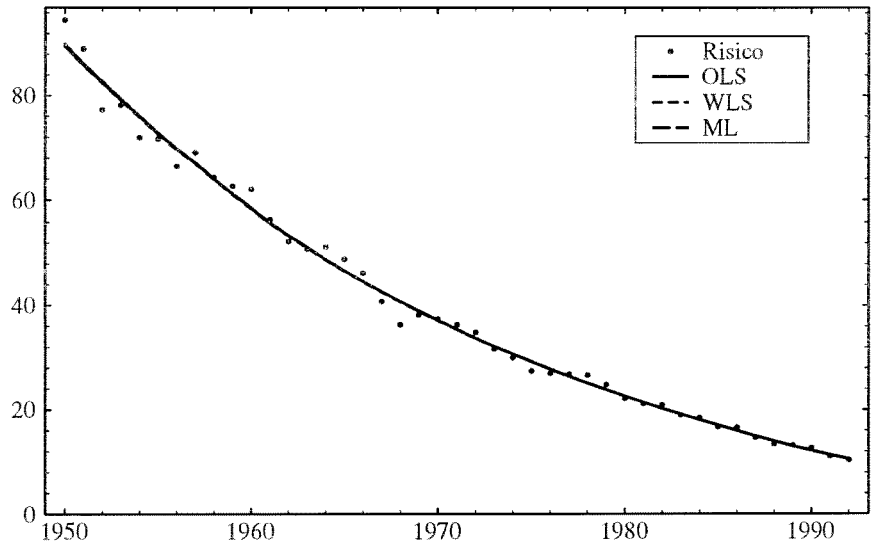
D.2. Risico's Engeland



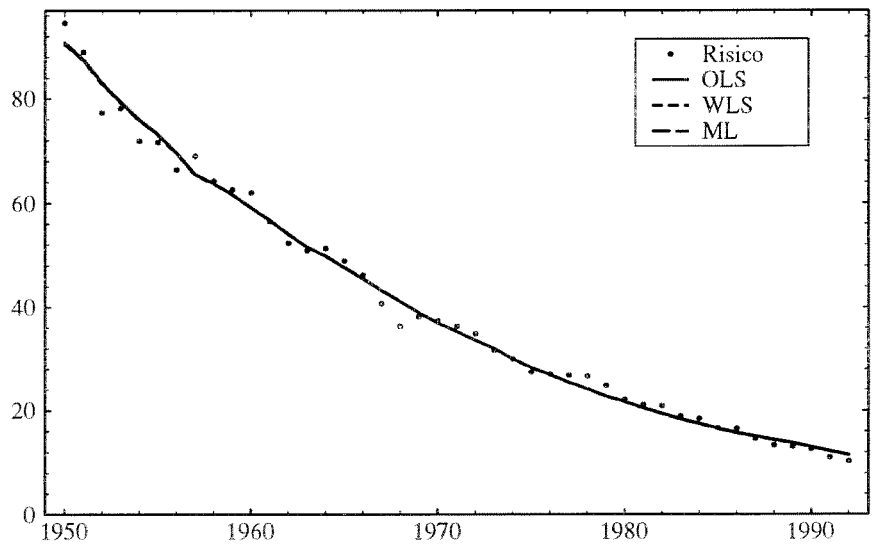
Afbeelding 64. *Risico's Engeland, Exp*



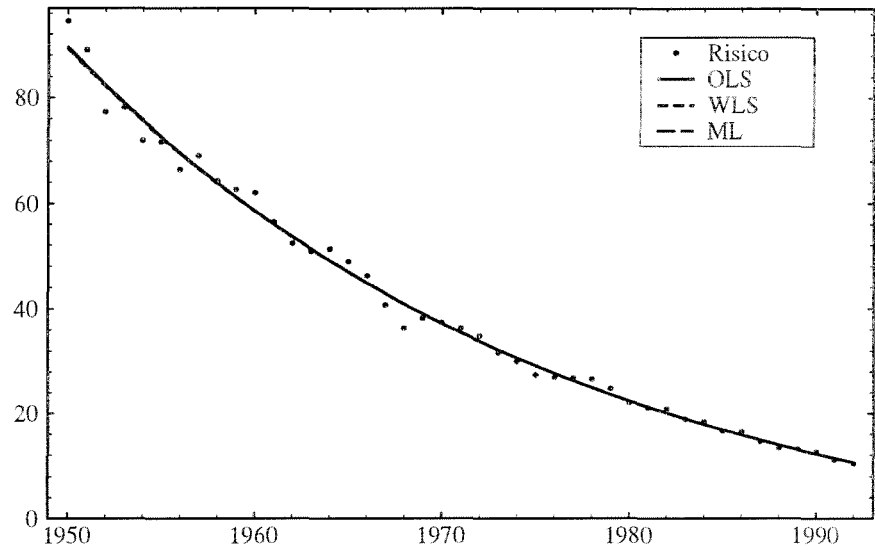
Afbeelding 65. *Risico's Engeland, Exp0*



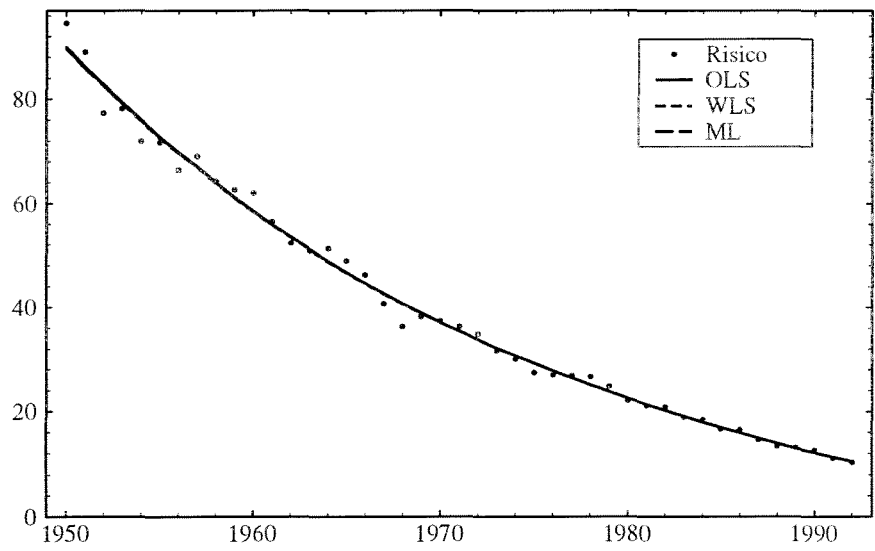
Afbeelding 66. *Risico's Engeland, ExpV*



Afbeelding 67. *Risico's Engeland, ExpV0*

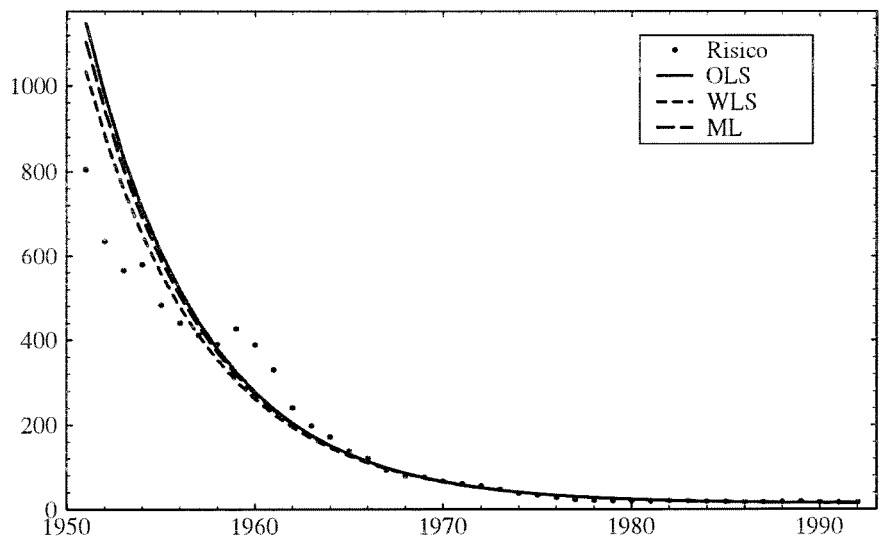


Afbeelding 68. *Risico's Engeland, Logis*

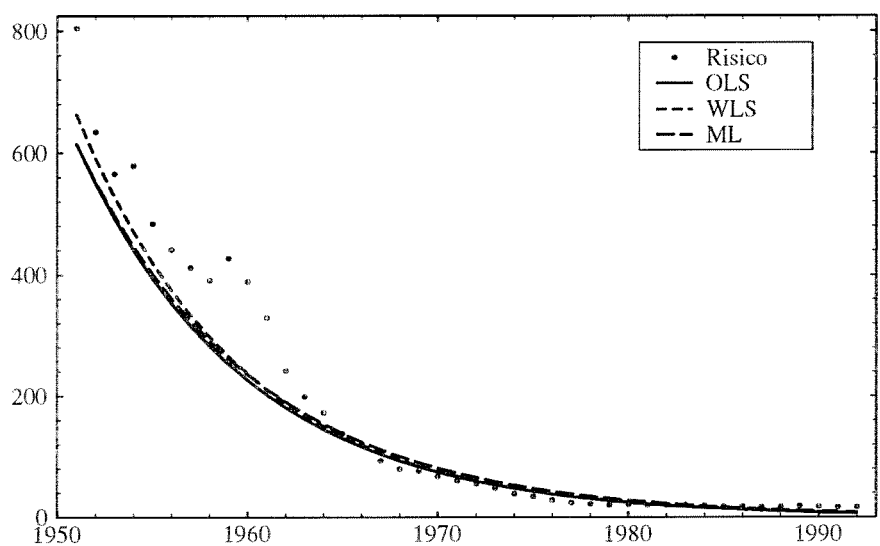


Afbeelding 69. *Risico's Engeland, LogisV*

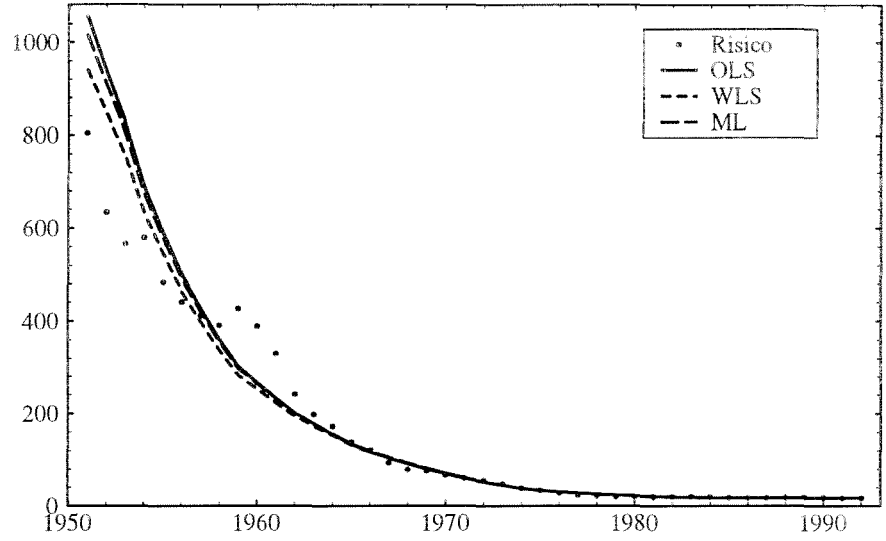
D.3. Risico's Japan



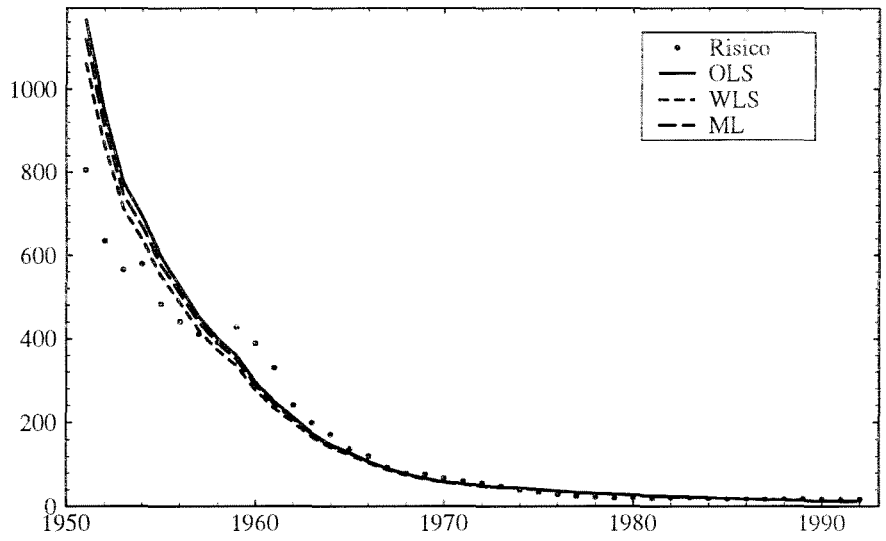
Afbeelding 70. *Risico's Japan, Exp*



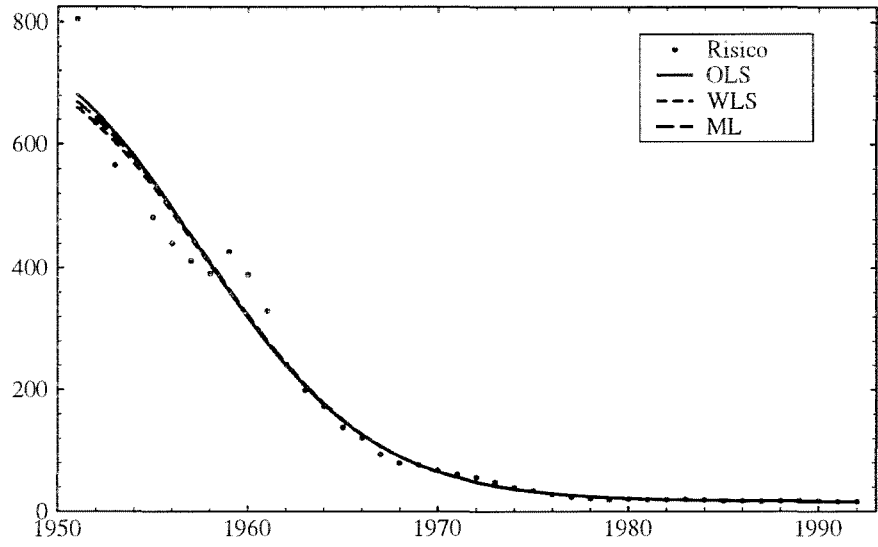
Afbeelding 71. *Risico's Japan, Exp0*



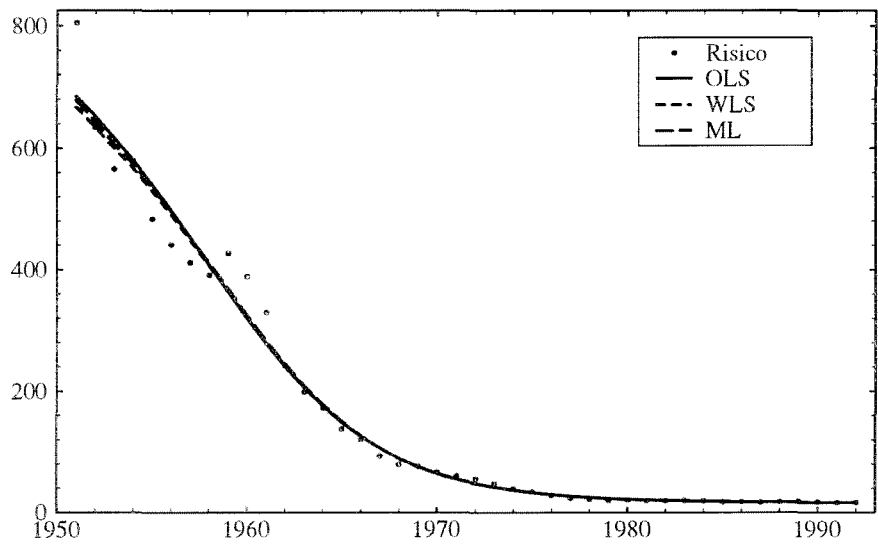
Afbeelding 72. *Risico's Japan, ExpV*



Afbeelding 73. *Risico's Japan, ExpV0*

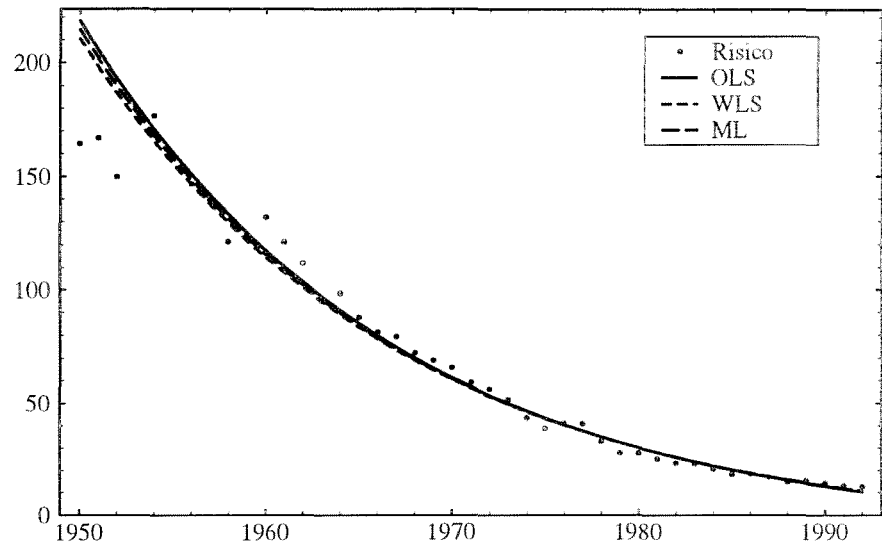


Afbeelding 74. *Risico's Japan, Logis*

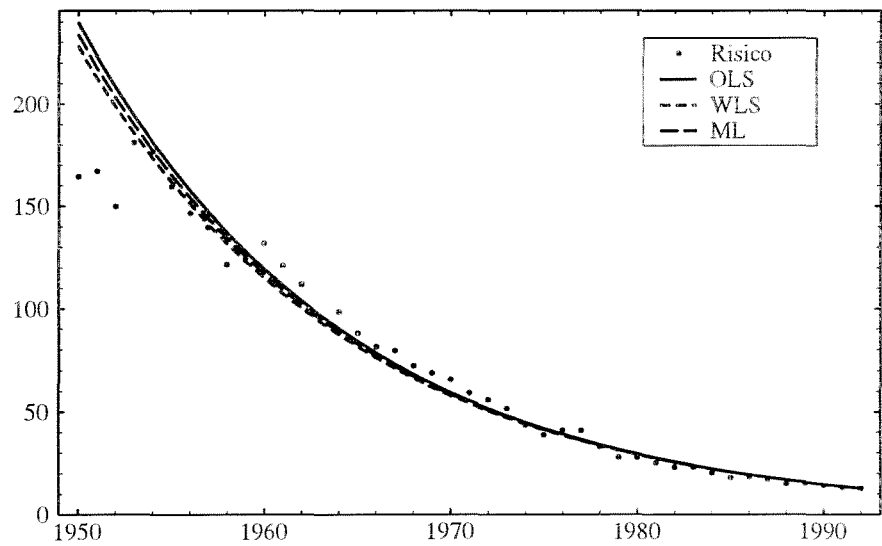


Afbeelding 75. *Risico's Japan, LogisV*

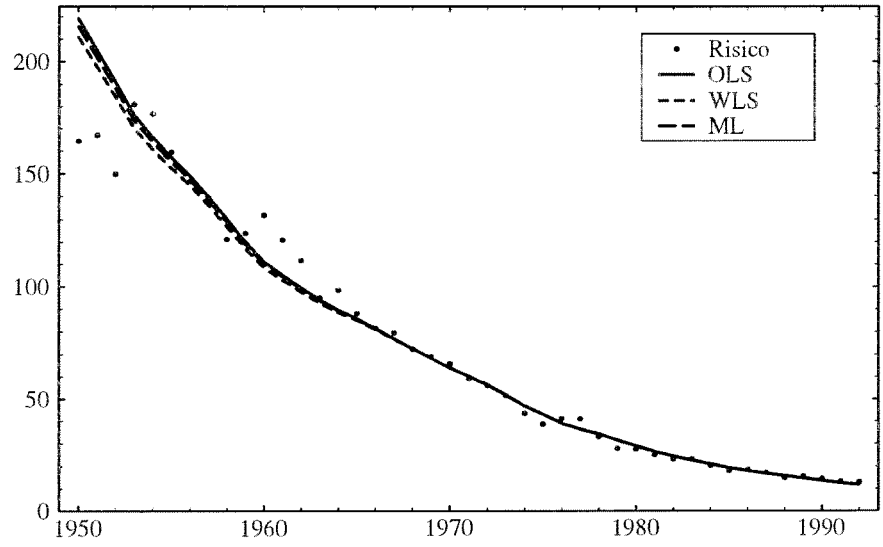
D.4. Risico's Nederland



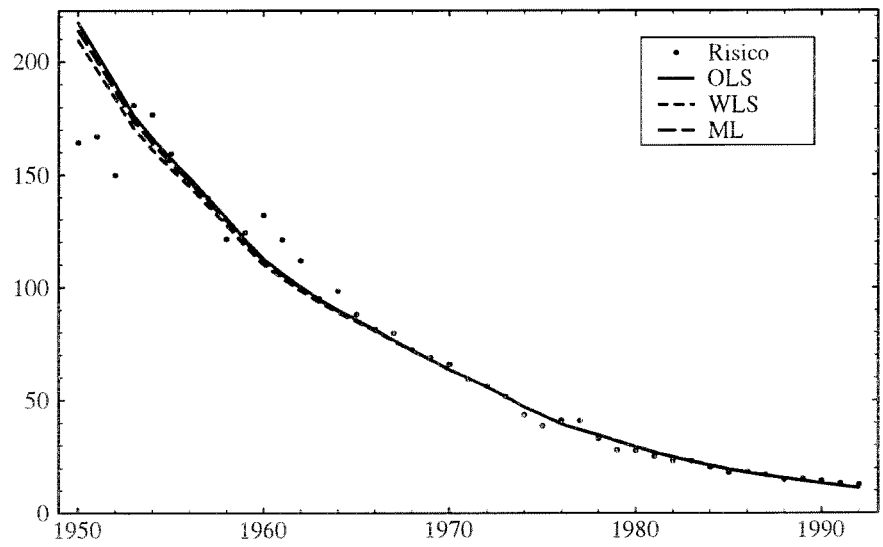
Afbeelding 76. *Risico's Nederland, Exp*



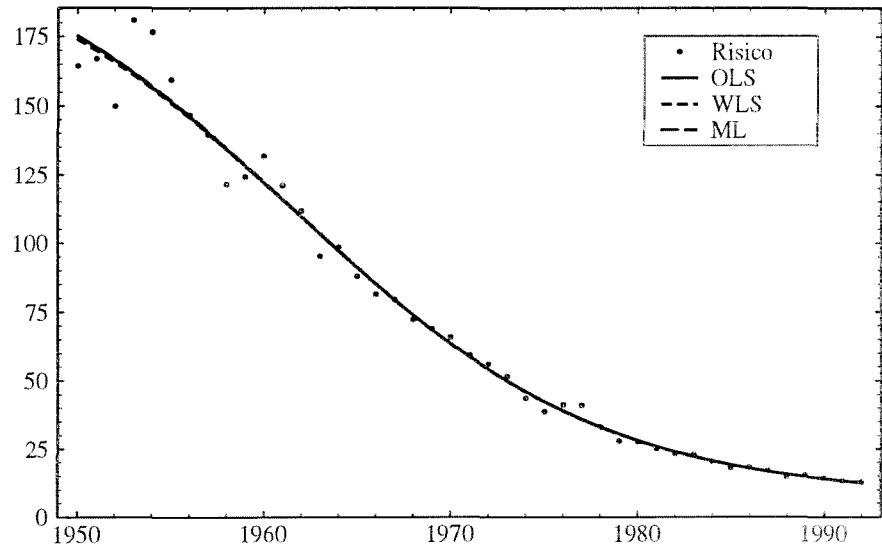
Afbeelding 77. *Risico's Nederland, Exp0*



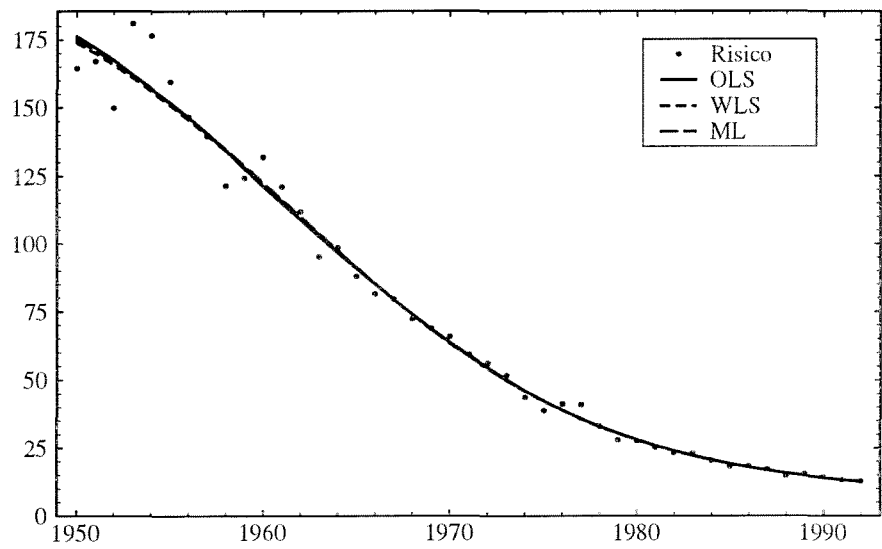
Afbeelding 78. *Risico's Nederland, ExpV*



Afbeelding 79. *Risico's Nederland, ExpV0*

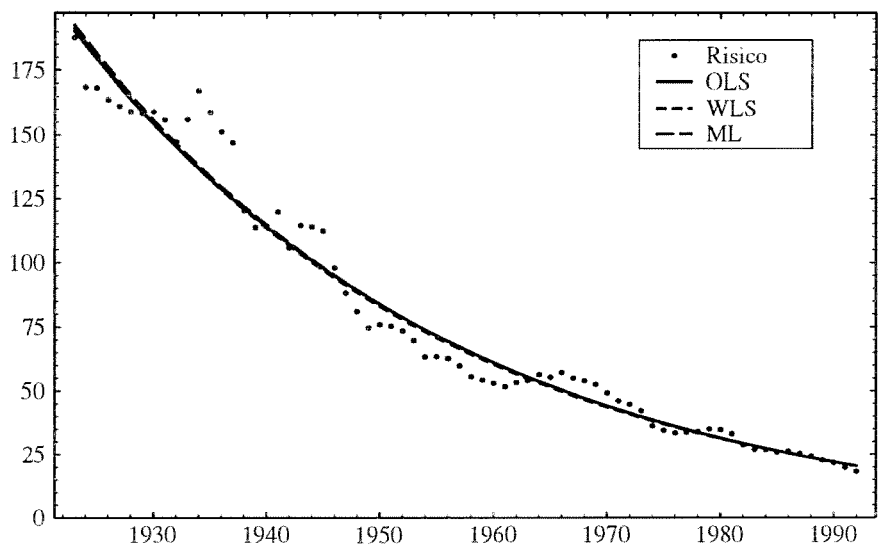


Afbeelding 80. *Risico's Nederland, Logis*

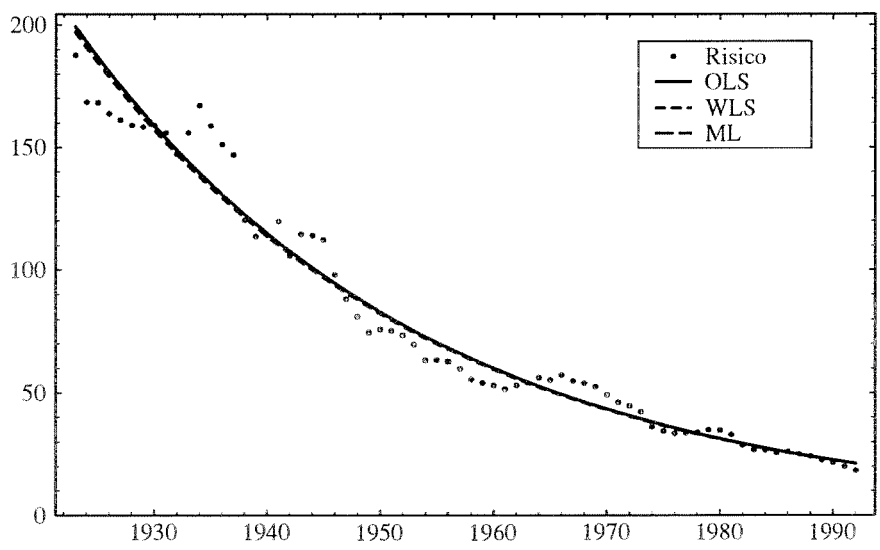


Afbeelding 81. *Risico's Nederland, LogisV*

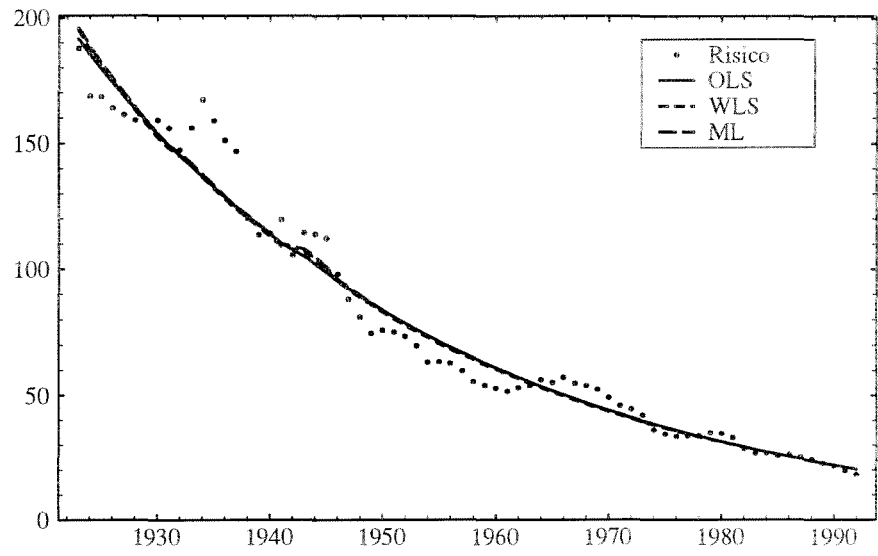
D.5. Risico's Verenigde Staten



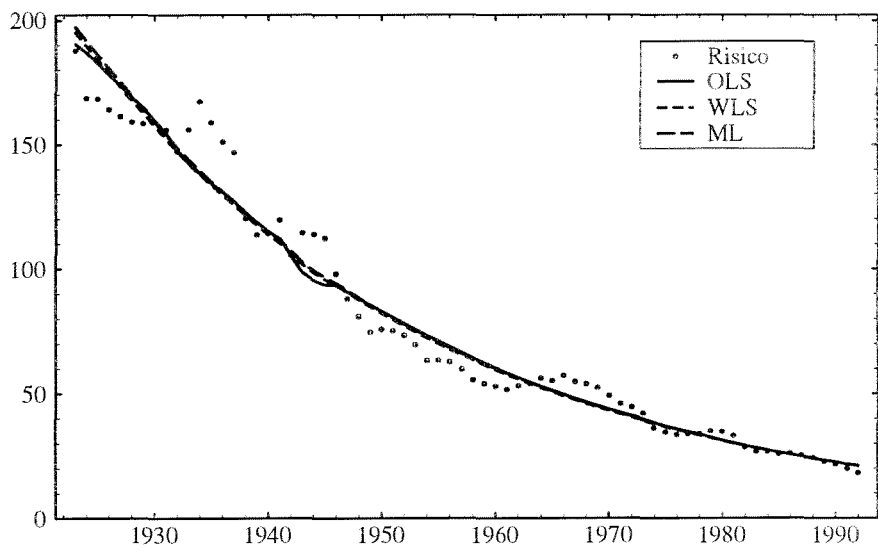
Afbeelding 82. *Risico's Verenigde Staten, Exp*



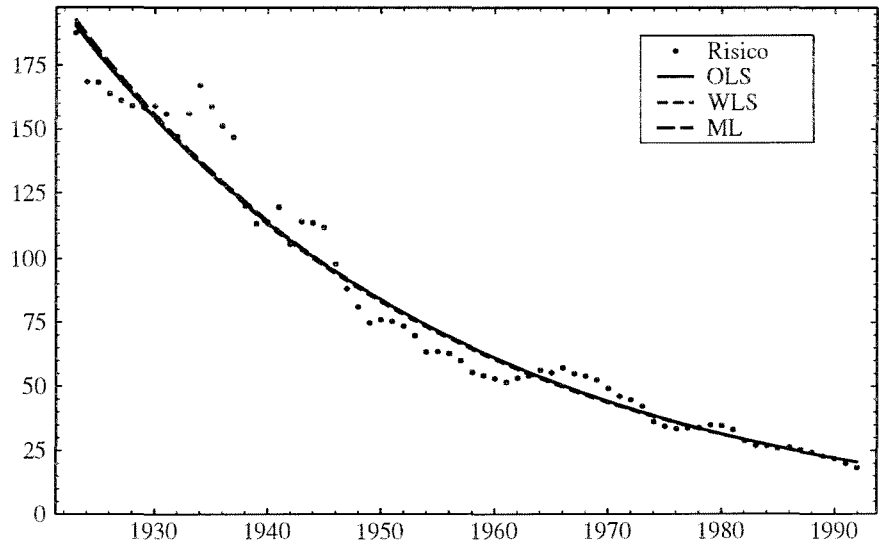
Afbeelding 83. *Risico's Verenigde Staten, Exp0*



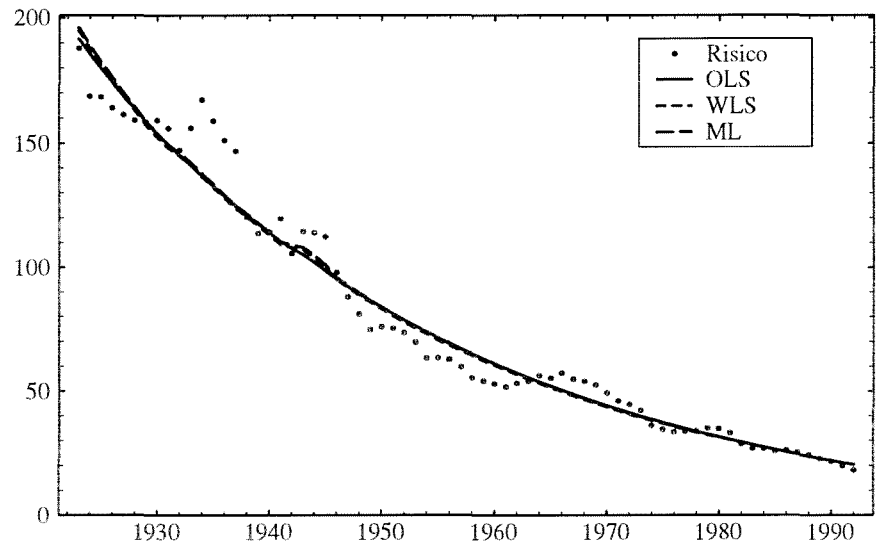
Afbeelding 84. *Risico's Verenigde Staten, ExpV*



Afbeelding 85. *Risico's Verenigde Staten, ExpV0*



Afbeelding 86. *Risico's Verenigde Staten, Logis*



Afbeelding 87. *Risico's Verenigde Staten, LogisV*

E. Voorlopige vergelijking van de Amerikaanse staat Michigan met de Verenigde Staten als geheel

E.1. Inleiding

Op het laatste moment is een vergelijking gemaakt tussen één staat uit de Verenigde Staten en de Verenigde Staten zelf. In deze bijlage wordt aan de hand van enige tabellen een vergelijking gemaakt die gelijk is aan de tabellen zoals die eerder zijn gebruikt in dit rapport. Voor een toelichting wordt naar het hoofdrapport verwezen.

E.2. Resultaten

E.2.1. Parameterwaarden

Verenigde Staten					
	a	b	c	d	e
Exp0	-32.42	67.63			
Exp	-28.94	60.93	-6.32		
ExpV0	-36.98	76.63			1.11
ExpV	-27.62	58.35	-7.32		0.98
Logis	-28.94	60.93	-6.32	0.00	
LogisV	-27.62	58.35	-7.32	0.00	0.98

Tabel 31. Parameter waarden OLS modellen

Michigan					
	a	b	c	d	e
Exp0	-34.07	70.78			
Exp	-20.39	44.36	-25.99		
ExpV0	-53.46	110.76			1.57
ExpV	-69.01	142.05	44.03		1.83
Logis	-62.50	127.14	5.12	0.10	
LogisV	-72.16	148.24	51.42	0.01	1.82

Tabel 32. Parameter waarden OLS modellen

Verenigde Staten					
	a	b	c	d	e
Exp0	-32.35	67.49			
Exp	-29.90	62.78	-4.44		
ExpV0	-33.46	69.67			1.03
ExpV	-25.43	54.04	-7.78		0.92
Logis	-29.90	62.78	-4.44	0.00	
LogisV	-25.43	54.04	-7.78	0.00	0.92

Tabel 33. Parameter waarden WLS modellen

Michigan					
	a	b	c	d	e
Exp0	-33.89	70.42			
Exp	-21.69	46.84	-21.42		
ExpV0	-53.15	110.13			1.57
ExpV	-70.43	144.91	51.62		1.86
Logis	-46.19	95.03	-0.83	0.08	
LogisV	-70.43	144.91	51.62	0.00	1.86

Tabel 34. Parameter waarden WLS modellen

Verenigde Staten					
	a	b	c	d	e
Exp0	-32.42	67.63			
Exp	-29.54	62.09	-5.40		
ExpV0	-33.59	69.94			1.03
ExpV	-25.24	53.70	-8.83		0.93
Logis	-29.54	62.09	-5.40	0.00	
LogisV	-25.24	53.70	-8.83	0.00	0.93

Tabel 35. Parameter waarden Poisson modellen

Michigan					
	a	b	c	d	e
Exp0	-34.08	70.80			
Exp	-21.35	46.22	-22.84		
ExpV0	-53.31	110.45			1.58
ExpV	-69.57	143.19	47.17		1.84
Logis	-56.49	115.28	3.48	0.09	
LogisV	-69.57	143.19	47.17	0.00	1.84

Tabel 36. Parameter waarden Poisson modellen

	USA	MIC
Exp0		
Exp	c	a, b, c
ExpV0	a, b, e	a, b, e
ExpV	a, b, e, c	a, b, e, c
Logis	c, d	a, b, c, d
LogisV	a, b, e, c, d	a, b, e, c, d

Tabel 37. Volgens asymptotische-toets niet significant van nul (1 in het geval van e) verschillende parameters voor gewogen modellen

E.2.2. R^2 -grootheden en aanverwanten

	USA	MIC
Exp0	0.85	0.68
Exp	0.86	0.72
ExpV0	0.85	0.79
ExpV	0.86	0.80
Logis	0.86	0.73
LogisV	0.86	0.80

Tabel 38. R^2 statistic voor niet-gewogen modellen.

	USA	MIC
Exp0	-0.58	-0.15
Exp	-0.52	-0.00
ExpV0	-0.56	0.23
ExpV	-0.52	0.27
Logis	-0.52	0.02
LogisV	-0.52	0.27

Tabel 39. R_D^2 statistic voor niet-gewogen modellen.

	USA	MIC
Exp0	0.85	0.67
Exp	0.86	0.72
ExpV0	0.85	0.79
ExpV	0.86	0.80
Logis	0.86	0.72
LogisV	0.86	0.80

Tabel 40. R^2 statistic voor gewogen modellen.

	USA	MIC
Exp0	0.85	0.68
Exp	0.86	0.72
ExpV0	0.85	0.79
ExpV	0.86	0.80
Logis	0.86	0.73
LogisV	0.86	0.80

Tabel 41. R^2 statistic voor likelihood modellen.

E.2.3. Kwadratensommen

	USA	MIC
Exp0	23202.61 (67)	668.87 (41)
Exp	22697.62 (66)	598.52 (40)
ExpV0	23181.43 (66)	452.64 (40)
ExpV	22577.76 (65)	423.52 (39)
Logis	22697.62 (65)	594.97 (39)
LogisV	22577.76 (64)	423.52 (38)

Tabel 42. χ^2 -waarden en vrijheidsgraden (df) voor gewogen modellen

	USA	MIC
Exp0	24018.57 (67)	665.64 (41)
Exp	23233.45 (66)	579.00 (40)
ExpV0	23991.00 (66)	442.73 (40)
ExpV	23107.26 (65)	419.61 (39)
Logis	23233.45 (65)	566.97 (39)
LogisV	23107.26 (64)	419.61 (38)

Tabel 43. χ^2 -waarden en vrijheidsgraden (df) op basis van likelihood fit

Michigan	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.92 (0.61)	0.69 (0.88)	0.67 (0.90)	0.94 (0.58)	0.68 (0.88)
Exp	1.09 (0.39)	1.00 (0.50)	0.76 (0.81)	0.73 (0.84)	1.02 (0.48)	0.74 (0.82)
ExpV0	1.44 (0.12)	1.32 (0.19)	1.00 (0.50)	0.96 (0.55)	1.35 (0.18)	0.98 (0.52)
ExpV	1.50 (0.10)	1.38 (0.16)	1.04 (0.45)	1.00 (0.50)	1.40 (0.15)	1.03 (0.47)
Logis	1.07 (0.42)	0.98 (0.52)	0.74 (0.82)	0.71 (0.85)	1.00 (0.50)	0.73 (0.83)
LogisV	1.46 (0.12)	1.34 (0.18)	1.02 (0.48)	0.97 (0.53)	1.37 (0.17)	1.00 (0.50)

Tabel 44. F -toetsen voor gewogen modellen. F -waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

	Exp0	Exp	ExpV0	ExpV	Logis	LogisV
Exp0	1.00 (0.50)	0.89 (0.64)	0.68 (0.89)	0.66 (0.90)	0.90 (0.63)	0.68 (0.88)
Exp	1.12 (0.36)	1.00 (0.50)	0.76 (0.80)	0.74 (0.82)	1.00 (0.49)	0.76 (0.80)
ExpV0	1.47 (0.11)	1.31 (0.20)	1.00 (0.50)	0.97 (0.53)	1.31 (0.20)	1.00 (0.50)
ExpV	1.51 (0.10)	1.35 (0.18)	1.03 (0.47)	1.00 (0.50)	1.35 (0.18)	1.03 (0.47)
Logis	1.12 (0.37)	1.00 (0.51)	0.76 (0.80)	0.74 (0.82)	1.00 (0.50)	0.76 (0.80)
LogisV	1.47 (0.12)	1.31 (0.20)	1.00 (0.50)	0.97 (0.53)	1.32 (0.20)	1.00 (0.50)

Tabel 45. *F*-toetsen voor likelihood modellen. *F*-waarde is χ^2 waarde horizontaal / χ^2 waarde verticaal

E.2.4. *Likelihood-resultaten*

	USA	MIC
Exp0	-12307.98 (2)	-538.05 (2)
Exp	-11964.98 (3)	-497.56 (3)
ExpV0	-12295.35 (3)	-427.98 (3)
ExpV	-11902.44 (4)	-415.45 (4)
Logis	-11964.98 (4)	-493.13 (4)
LogisV	-11902.44 (5)	-415.45 (5)

Tabel 46. *Likelihood-waarden en (aantallen parameters) onder Poisson aanname*

E.2.5. *Diagnostische resultaten*

	USA	MIC
Exp0	*219.55	*65.92
Exp	*214.33	*101.30
ExpV0	*216.13	*97.07
ExpV	*224.15	*87.65
Logis	*214.33	*102.87
LogisV	*224.15	*87.65

Tabel 47. *Box-Ljung Q statistic voor gewogen modellen. * betekent significant onder aannname van ongecorrleerde residuen.*

	USA	MIC
Exp0	0.02	0.02
Exp	0.05	0.07
ExpV0	0.04	0.02
ExpV	0.09	0.08
Logis	*38.57	0.21
LogisV	*658220.91	*18.27

Tabel 48. *Intrinsieke krommings-waarden en significantie tegen 5% * voor gewogen modellen*

	USA	MIC
Exp0	0.03	0.03
Exp	*1.83	*18.46
ExpV0	0.06	0.04
ExpV	*10.68	*18.03
Logis	* 3.91×10^{16}	*75.27
LogisV	* 8.47×10^{20}	* 1.94×10^{16}

Tabel 49. *Parameter krommings-waarden en significantie tegen 5% * voor gewogen modellen*

E.2.6. Afgeleide resultaten

	USA	MIC
Exp0		
Exp	2049.66	2018.70
ExpV0		
ExpV	2044.47	
Logis	2049.66	2061.23
LogisV	2044.47	

Tabel 50. Nulpunten van risico ontwikkelingen

	USA	MIC
Exp0		
Exp		
ExpV0		
ExpV		1245
Logis		
LogisV		1701

Tabel 51. Minima verhouding risico ontwikkelingen

	USA	MIC
Exp0		
Exp		
ExpV0		
ExpV		2027
Logis		
LogisV		2027

Tabel 52. 99% punten van risico ontwikkelingen

	USA	MIC
Exp0		
Exp		
ExpV0		
ExpV		2062
Logis		
LogisV		2062

Tabel 53. 99,9% punten van risico ontwikkelingen