

Oorzaken van verkeersonveiligheid

Een methodologische literatuurstudie

R-93-63
Frank Poppe
Leidschendam, 1993
Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
Postbus 170
2260 AD Leidschendam
Telefoon 070-3209323
Telefax 070-3201261

Samenvatting

Het project *Mobiliteit en Veiligheid* probeert op geaggregeerd niveau een relatie te leggen tussen ontwikkelingen in de mobiliteit en de verkeersonveiligheid, met het doel daar ook vooruitberekeningen mee te kunnen doen. Om de effecten van specifieke maatregelen in dergelijke vooruitberekeningen mee te kunnen nemen is belangstelling ontstaan voor methoden die op geaggregeerd niveau de mogelijke relaties in beeld trachten te brengen. Het gaat daarbij enerzijds om maatschappelijke ontwikkelingen en overheidsmaatregelen en anderzijds om veranderingen in de omvang van het verkeer, het aantal ongevallen en de ernst van die ongevallen. Daarmee kunnen mogelijke 'oorzaken' van de verkeersonveiligheid worden opgespoord.

In dit rapport is een aantal buitenlandse onderzoeken gezien. De aandacht is daarbij gericht geweest op de toegepaste methoden en statistische analysetechnieken. De inhoudelijke resultaten van de onderzoeken komen hier dus niet aan de orde. De centrale vraag was of hieruit methoden van aanpak konden worden afgeleid die ook in de Nederlandse situatie kunnen worden toegepast.

Geconstateerd is dat met de interpretatie van de met dergelijke methoden gevonden verbanden zeer voorzichtig moet worden omgesprongen. Niettemin kunnen deze methoden, toegepast bij exploratief onderzoek, interessante resultaten opleveren. Het onderscheid tussen de verschillende fasen in het verkeersysteem, zoals dat met name bij Gaudry consequent gebeurt, sluit goed aan bij de wijze waarop men in Nederland verkeer en verkeersonveiligheid plaatst binnen de maatschappelijke ontwikkelingen.

Er zijn in Nederland voldoende tijdreeksen met gegevens over de verschillende niveaus (fasen) in het verkeersysteem om een dergelijk onderzoek te starten. Dat zou het beste relatief kleinschalig kunnen beginnen, gebruikmakend van direct beschikbare gegevens. Ook lijkt het voor een eerste stap niet wenselijk elders ontwikkelde statistische analysemodellen eerst te vertalen naar de Nederlandse situatie. Met gebruikmaking van bestaande en beschikbare analysemethoden moet een dergelijk model op te stellen zijn. Wel is het goed de internationale samenwerking, die op dit terrein van de grond komt, te benutten en te versterken.

Abstract

The SWOV-project *Mobility and Safety* tries to link, at an aggregate level, the developments in mobility and those in traffic safety, with the goal to make prognoses possible. To incorporate the effects of specific countermeasures in such prognoses interest is growing in methods that try to make these relations explicit. This concerns social and economical developments and legislation on the one hand, and changes in the amount of traffic, the number of accidents and the seriousness of those accidents. In this way possible 'causes' of traffic unsafety can be detected.

In this report a number of foreign studies has been looked at. The focus has been at the used methodology and the applied statistical analysis techniques. The results in the field of traffic safety itself will not be discussed. The central question was whether it was possible to deduce analysis methods that could be applied successfully in the Netherlands.

It is concluded that one has to be very carefully with the interpretation of the relations that can be found with such methods. Nevertheless these methods can lead to interesting results when they are applied in the context of exploratory research.

The distinction between the different phases in the traffic system fits very well with the way the traffic system is looked at in the Netherlands. Particularly in the model of Gaudry these phases are handled consistently.

In the Netherlands ample time series of data on the different levels (phases) of the traffic system are available. Research as described in this report can therefore be started. It would be best to start this on a relatively small scale, using immediate available data. Being a first step, it does not seem advisable to transform models for statistical analysis that have been developed elsewhere to the Dutch situation. It will be possible to formulate such a model using existing and available methods for analysis.

At the same time it would be good to use and strengthen the international cooperation that is developing in this field.

Inhoud

1.	<i>Inleiding</i>	6
1.1.	Achtergrond	6
1.2.	Wijze van beschrijven	6
1.3.	Geselecteerde modellen	6
2.	<i>De context</i>	8
3.	<i>Het model</i>	9
3.1.	Aggregatie en hiërarchie	9
3.2.	Expliciete, impliciete en latente verbanden	9
4.	<i>De werkwijze</i>	11
4.1.	Vorbereidingen	11
4.2.	Na de analyse	11
5.	<i>De techniek</i>	14
5.1.	Autocorrelatie	14
5.2.	Poisson-proces	14
6.	<i>Een kritische beschouwing</i>	15
6.1.	Het model zelf	15
6.2.	Het gebruik van de resultaten	16
6.3.	Conclusies	17
7.	<i>Mogelijkheden in de Nederlandse situatie</i>	18
7.1.	Exploratief en toetsend onderzoek	18
7.2.	Consequenties voor toepassingen in Nederland	18
7.3.	Gegevens	19
7.4.	Het voorstel concreet	20
Bijlage A.	<i>Enkele samenvattingen</i>	23
A.1.	Oudere verkeersslachtoffers in Kroatië	23
A.2.	Pleidooi voor econometrische modellen	23
A.3.	De vier Noordse landen	24
A.4.	De verantwoording van het DRAG-model	24

1. Inleiding

1.1. Achtergrond

Binnen het project *Mobiliteit en Veiligheid* zijn de activiteiten er op gericht om op een geaggregeerd niveau het verband tussen mobiliteitsgegevens en onveiligheidsgegevens in kaart te brengen, om zo prognoses voor de onveiligheid (gegeven een prognose voor de mobiliteit) te kunnen maken. 'Geaggregeerd' wil hierbij zeggen dat de prognoses totalen voor een gebied betreffen, voor Nederland of (bijv.) voor een provincie. De verschillende lijnen die daarbij gevolgd (kunnen) worden hebben als gemeenschappelijk kenmerk dat het 'intermediair' tussen mobiliteit en onveiligheid gevormd wordt door het *risico*.

Daarbij wordt wel uitgesplitst naar een aantal factoren die (de ontwikkeling van) het risico beïnvloeden, zoals vervoerwijze, leeftijd bestuurder en/of slachtoffer, wegtype, enz. het doel hiervan is enerzijds het bereiken van een betere voorspelling, anderzijds wordt daarbij tegemoet gekomen aan beleidsmatige wensen.

Daarnaast bestaat de wens ook het effect van meer specifieke maatregelen op de prognoses voor de verkeersonveiligheid door te kunnen rekenen. Daarvoor is echter nodig dat op een dergelijk geaggregeerd niveau de effecten van maatregelen kwantitatief bepaald kunnen worden. Anders gezegd, kunnen de oorzaken van verkeersonveiligheid opgespoord worden door op macro-niveau een groot aantal mogelijk verklarende variabelen te relateren aan variabelen die iets zeggen over de verkeersonveiligheid?

Onder meer als gevolg van het werk van Gaudry [1984, 1989, 1990] (resp. de complete modelbeschrijving, een aantal analyseresultaten, een kort overzicht) heeft deze benadering de laatste tijd weer meer belangstelling gekregen. Een aantal van dergelijke benaderingen wordt hier besproken. De deelstudie kan gekarakteriseerd worden als een 'methodologische literatuurstudie'. Het gaat hier dus niet in de eerste plaats om de inhoudelijke conclusies die in de literatuur aangetroffen worden, maar om de toegepaste methoden en technieken. Daarbij zal de aandacht ook gericht worden op de vraag of dergelijke methoden en/of technieken ook in Nederland toegepast zouden kunnen worden.

1.2. Wijze van beschrijven

De bestudeerde analyses hebben uiteraard een groot aantal verschillen. Ze hebben echter ook een gemeenschappelijke noemer. Daarom wordt eerst de gemeenschappelijke context beschreven die deze benaderingen delen. Vervolgens wordt in de volgende paragrafen van deze notitie een theoretisch kader aangegeven. De bestudeerde literatuur wordt daarbij ingedeeld naar o.a. het gehanteerde *model*, de *werkwijze* waarbinnen het model toegepast wordt, en de *techniek* die gebruikt wordt om het model op de gegevens toe te passen. Die begrippen, model, werkwijze en techniek, worden eerst uitgewerkt.

1.3. Geselecteerde modellen

Getracht is om ten behoeve van deze rapportage een beperkt aantal publikaties te selecteren aan de hand waarvan de in de praktijk voorkomende benaderingen kunnen worden beschreven. Voor het zoeken naar dergelijke literatuurreferenties is gebruik gemaakt van dezelfde 'keywords' als voor de andere in dit project uitgevoerde literatuurstudie [Bos, 1993]: *traffic* en *safety*, in combinatie met *model of analysis*.

Na een verdere selectie op relevantie, is vervolgens ook gekeken naar in de geselecteerde publikaties nog genoemde literatuur ('cascade-systeem'). Een aantal

van de op deze wijze gevonden beschrijvingen van onderzoeken viel buiten het kader van deze literatuurstudie, bijv. omdat het uitsluitend 'cross-sectioneel' onderzoek betrof, waar geen metingen op verschillende tijdmomenten plaatsvond.

Ook publikaties die uitsluitend de resultaten van een onderzoek beschrijven, en niet ingaan op de onderzoek- en/of analysemethode, zijn verder terzijde gelegd. Overigens kwamen sommige van dergelijke onderzoeken wel in een andere publikatie aan de orde.

Uiteindelijk worden de volgende onderzoeken verder beschreven.

In Bijlage A wordt de geselecteerde literatuur kort samengevat. Alleen de rapportage van Gaudry wordt uitgebreider behandeld (A.4), enerzijds omdat de aandacht voor deze publikatie één van de aanleidingen voor dit literatuuronderzoek was, anderzijds om de toegankelijkheid van die rapportage te vergroten, aangezien deze slechts in het Frans beschikbaar is.

Vorko-Jović & Jović [1992] gaan uit van het jaarlijkse aantal verkeersdoden en -gewonden *onder ouderen* in de periode 1971–1986. Dit wordt in verband gebracht met de samengestelde variabelen brandstofverkoop per weglengte (benadering van verkeersintensiteit) en aantal motorvoertuigen per inwoner (de motorisatie). Ook de jaarlijkse verandering in die variabelen wordt in het model geïntroduceerd. De analyse wordt uitgevoerd als een rangorde-regressie,

Fridstrøm [1991] houdt een algemeen pleidooi voor het toepassen van (geaggregeerde) econometrische modellen in de verkeersveiligheidsanalyse. Deze zijn immers bedoeld om te pogen 'oorzakelijke' verbanden te leggen bij gegevens (vnl. tijdreeksen) die vrijwel uitsluitend op niet-experimentele wijze verkregen worden. In veel verkeersveiligheidsanalyses wordt met hetzelfde probleem geworsteld. Bij wijze van voorbeeld worden analyses uitgevoerd op maandelijkse gegevens. Verschillende verkeersonveiligheidsmaten worden in verband gebracht met gegevens over expositie, weer, daglicht, weglengte, registratie, voertuiginspectie, politietoezicht, gordelgebruik, rijervaring en alcoholverkoop. Statistisch gezien zijn de resultaten bevredigend. Expositie blijkt veruit de belangrijkste verklarende variabele.

Een gezamenlijk onderzoek van onderzoeksinstituten uit Noorwegen, Zweden, Finland en Denemarken [Fridstrøm e.a., 1992] heeft geprobeerd de variatie in het maandelijkse aantal geregistreerde ongevallen te relateren aan expositie, weer en lichtomstandigheden, en registratiesysteem, d.m.v. een gegeneraliseerd Poisson regressiemodel. De bedoeling was om na eliminatie van deze factoren te kunnen bepalen wat de (lange termijn) daling in het risico is. De auteurs zijn tevreden over de prestaties van het model. Hoewel het aantal te analyseren ongevallen soms klein was, 'degenereerde' de oplossing niet snel. Overigens was ook hier expositie de belangrijkste verklarende variabele, verantwoordelijk voor ca. 50% van de variatie.

In de bespreking wordt overigens nog verwezen naar elementen uit een aantal andere publikaties.

2. De context

Het gemeenschappelijke van de verschillende benaderingen die in de geselecteerde literatuur gevonden worden zit in het gebruik van tijdreeksen van geaggregeerde gegevens. Op het niveau van een groot geografisch gebied (een staat of een geheel land) worden verklarende, en te verklaren factoren met elkaar in verband gebracht.

Verschillen zijn te constateren in de mate van uitsplitsing die op dat niveau worden toegepast. Sommige benaderingen beperken zich tot het totaal aantal ongevallen (of slachtoffers, of doden) [Fridstrøm e.a., 1992], andere beperken zich tot een deelcategorie daarbinnen (bijv. alleen ouderen: [Vorko-Jović & Jović, 1992]). Het komt ook voor dat naast het totaal ook een aantal deelcategorieën afzonderlijk ‘verklaard’ wordt [Fridstrøm, 1991].

Hiervoor is het woord *verklaard* gebruikt. In sommige referenties wordt vrij uitgebreid aandacht besteed aan het feit dat het vinden van correlaties nooit kan aantonen dat één waargenomen aspect de oorzaak van een ander aspect is. Sterker nog, dat men oorzaak en gevolg nooit zal kunnen bewijzen. Wanneer een aantal voorwaarden vervuld wordt kan de hypothese wel zeer aannemelijk worden. De neiging om een gevonden correlatie ook direct te verklaren in termen van de op dat niveau geanalyseerde variabelen blijkt steeds groot.

Het valt bovendien nooit met zekerheid aan te nemen dat een dergelijke relatie in de toekomst blijft bestaan. Een statistische bevredigende verklaring (die zich wellicht ook goed laat interpreteren) van het verband tussen verschillende tijdreeksen kan berusten op een verband dat in latere jaren niet meer bestaat [Partyka, 1991].

De behandeling van het toeval en waarschijnlijkheid blijft soms zeer impliciet. In een enkel geval wordt onderscheid gemaakt tussen fluctuaties die ontstaan doordat het beschouwde systeem in zichzelf een mate van onvoorspelbaarheid heeft, fluctuaties in de waarnemingen doordat deze waarnemingen de werkelijkheid niet nauwkeurig vastleggen, en fluctuaties in het model doordat het model niet voldoende nauwkeurig de werkelijke relaties vastlegt [Fridstrøm, 1991, Gaudry, 1984]

3. Het model

Aan het model zoals dat geformuleerd wordt ten behoeve van de analyse is een aantal aspecten te onderscheiden. In de navolgende paragrafen gaan we enerzijds in op de wijze waarop de *gegevens* worden beschreven en worden vastgelegd, en anderzijds op de wijze waarop de gegevens met elkaar in verband worden gebracht.

3.1. Aggregatie en hiërarchie

Vrijwel alle analyses gaan (vaak impliciet) uit van een situatie waarbij de gegevens bestaan uit een bestand waarin enerzijds de kenmerken onderscheiden worden en anderzijds de waarnemingen. De gegevens zijn dus vastgelegd in een zgn. *rechthoekig* bestand.

Dit lijkt wellicht triviaal, maar heel vaak zijn de gegevens *niet* vastgelegd in zo'n rechthoekig bestand; zeker in onderzoek naar verkeersongevallen is veel vaker sprake van hiërarchische bestanden. Een hiërarchisch bestand is dan in het algemeen te beschouwen als een verzameling rechthoekige bestanden die via één of meer gemeenschappelijke kenmerken gekoppeld kunnen worden. Bij ongevallenanalyses gaat het dan meestal om de volgende bestanden:

- locaties;
- ongevallen;
- betrokken voertuigen;
- slachtoffers.

Ten opzichte van andere hiërarchische bestanden hebben deze bestanden nog de bijzonderheid dat voor sommige niveaus het aantal bijbehorende observaties van het lagere niveau ook 'nul' kan zijn (locaties zonder ongevallen in een onderzoekjaar, voertuigen betrokken in een botsing maar zonder slachtoffers). Voor de toepassing van sommige technieken kan dat problemen geven.

Hier gaat het echter steeds om geaggregeerde gegevens, dus gedisaggregeerde gegevens en gegevens van lagere niveaus zijn opgeteld of anderszins tot kenmerken van een hoger niveau gebracht, en zijn dus niet meer als gedisaggregeerde gegevens beschikbaar. Er blijft echter vaak nog een keuze te maken met betrekking tot het precieze niveau van de aggregatie. Het gaat daarbij meestal om de geografische aggregatie en/of om de tijdeenheid. In geografische zin kan men optellen tot een geheel land, of enige onderverdeling (staat, provincie, county) toelaten. De tijdvakken waarover men telt kunnen jaren zijn, maar maandelijkse waarnemingen zijn ook mogelijk. Fridstrøm [1991] geeft enige overwegingen gegeven op grond waarvan men deze keuze beargumenteerd kan maken. Enkele valkuilen, zoals de 'ecological fallacy' [Robinson, 1950] die bij het interpreteren van geaggreerde analyses opdoemen, worden daar aangestipt.

3.2. Expliciete, impliciete en latente verbanden

Onderdeel van 'het model' is ook de vraag hoe het verband tussen de verschillende kenmerken gedacht wordt. Daarbij kunnen verschillende stromingen onderscheiden worden. Wordt er alleen een onderscheid gemaakt tussen de verklarende en de te verklaren kenmerken, of worden ook indirecte variabelen gebruikt? Zo ja, worden deze tussenliggende variabelen expliciet gedefinieerd en ook gemeten, of worden deze als latente variabelen beschouwd.

In sommige benaderingen (bijvoorbeeld Gaudry [1990] en Fridstrøm [1991]) loopt de invloed van externe kenmerken via een tevoren geformuleerd 'pad' naar de afhankelijke kenmerken. Hier is sprake van *expliciete* interveniërende

kenmerken.

Bij een definiëring als dynamisch tijd-ruimte systeem [Bijleveld, 1990] wordt niet bij voorbaat gedefinieerd hoe dit systeem functioneert. Hier is dus wel sprake van kenmerken die een vertaling van de waargenomen verklarende naar de waargenomen afhankelijke kenmerken verzorgen, maar deze zijn *latent*.

Het meest eenvoudige model houdt in het geheel geen rekening met tussenliggende kenmerken: er wordt direct gezocht naar de relatie tussen de waargenomen kenmerken. Het 'meest eenvoudige' heeft hier overigens uitsluitend betrekking op het uitvoeren van de analyse, voor het interpreteren van de resultaten van de analyse hoeft dat in het geheel niet op te gaan.

4. De werkwijze

Hiervoor is ingegaan op de manier waarop het model gedefinieerd wordt, aan de hand van enerzijds de beschrijving en vastlegging van de gegevens. In het volgende hoofdstuk zal ingegaan worden de technische (vnl. statistische) kanten van het rekenen aan dat model.

In dit hoofdstuk wordt kort ingegaan op de plaats die dat binnen de werkwijze van de onderzoeker kan hebben.

Dit heeft vooral betrekking op hetgeen er vóórafgaand aan 'het rekenen' plaatsvindt, en wat er direct daarna gebeurt.

4.1. Voorbereidingen

Vrijwel alle hier beschreven onderzoekingen behoren tot hypothese-vormend onderzoek. Dat wil zeggen dat voorafgaand aan de analyse geen mogelijke verbanden tussen de variabelen geformuleerd zijn die met behulp van de verzamelde gegevens getoetst kunnen worden. Alleen in het werk van Gaudry [1984] wordt een aantal veronderstellingen geformuleerd. Ook daar is het echter kennelijk niet de bedoeling tot een expliciete en formele toetsing daarvan te komen, maar was het doel meer gericht op het aantonen van het nut om bepaalde econometrische noties in het verkeersveiligheidsonderzoek te introduceren.

4.2. Na de analyse

Naarmate minder sprake is van formeel hypothese-toetsende analyses zijn de mogelijkheden voor het interpreteren van de gevonden verbanden groter. Bij de hier besproken analyses is dat dus het geval. Omdat de dan geformuleerde interpretaties niet meer getoetst (kunnen) worden, zijn in dit stadium de risico's van verkeerde interpretaties groot. Hier zullen een aantal problemen aangeduid worden. Later (Hoofdstuk 6) zal dit getoetst worden aan een aantal van de besproken onderzoekingen.

Hiervoor (par. 3.1) werd al de 'ecological fallacy' genoemd. Deze werd voor het eerst onder die naam beschreven door Robinson [1950]. In zijn artikel wordt onder *ecologie* zoveel als groeps eigenschappen verstaan. Hij geeft eerst aan hoe vaak correlaties tussen eigenschappen van groepen in diverse studies gebruikt worden, en dat daar vaak impliciet of zelfs expliciet interpretaties aan worden gekoppeld t.a.v. dergelijke correlaties bij individuele leden van die groepen. Het Amerikaanse *Census Bureau* verdeelt de Amerikaanse staten in negen groepen. Vervolgens laat hij (zie Afbeelding 1) de relatie zien voor die negen gebieden tussen het percentage analfabeten en het percentage negers (de gegevens zijn van 1930, nu zou dat met 'zwarten' of 'van Afrikaanse afkomst' heten). De Pearson-correlatie is .946, hetgeen ook duidelijk uit de figuur te zien is. Vervolgens wordt voor de bijna 100 000 individuen waarop de gegevens gebaseerd zijn de 2×2 -tabel gegeven van analfabetisme en ras. Hieruit blijkt een correlatie van .203!

Waar op basis van groepsgemiddelde een bijna perfecte correlatie geconstateerd werd, blijkt dit op individueel niveau een veel zwakker verband te zijn.

Eenzelfde voorbeeld wordt gegeven voor de relatie tussen analfabetisme en het al dan niet geboren zijn binnen de VS: voor de 9 gebieden is de zgn. ecologische correlatie tussen percentage analfabeten en percentage buitenlands geboren - .619 (minder analfabeten bij meer import), de individuele correlatie echter is +.118. Hier is de relatie op individueel niveau zelfs tegengesteld aan die op groepsniveau.

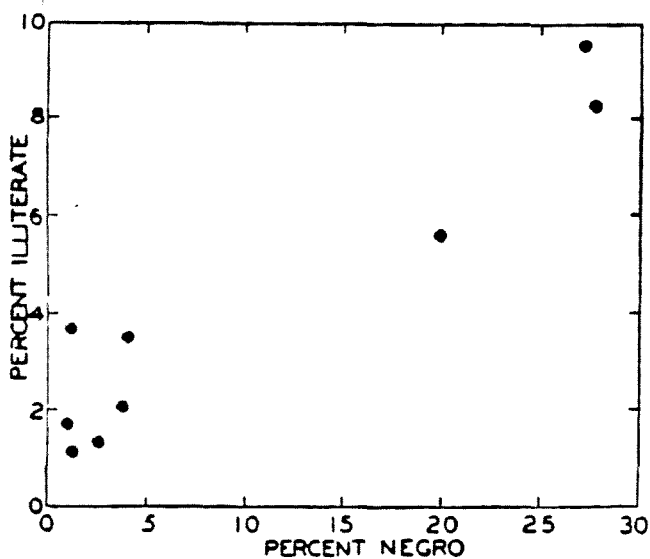


TABLE I. THE INDIVIDUAL CORRELATION BETWEEN COLOR AND ILLITERACY FOR THE UNITED STATES, 1930 (for the population 10 years old and over)¹⁷

	Negro	White	Total
Illiterate	1,512	2,406	3,918
Literate	7,780	85,574	93,354
Total	9,292	87,980	97,272

Afbeelding 1. Verband tussen analfabetisme en ras, in percentages per geografisch gebied (figuur) en per individu (tabel) (Bron: Robinson, 1950)

Door Fridstrøm [1991] wordt nog een ander voorbeeld aangehaald: uit het gegeven dat zelfmoorden relatief vaak voorkomen in vnl. protestante gebieden [Durkheim, 1951] mag niet de conclusie worden getrokken dat protestanten vaker zelfmoord plegen: het kunnen de enkele katholieken in die gebieden zijn die juist zelfmoord plegen.

Waar, zoals we gezien hebben, vrijwel alle analyses op geaggregeerde aantallen betrekking hebben, is het gevaar van interpretaties die ongefundeerd de correlaties vertalen naar de achterliggende individuen nadrukkelijk aanwezig.

Ook wanneer geen expliciete hypothesen zijn opgesteld maken onderzoekers bij het presenteren van de resultaten, en m.n. de correlaties, onderscheid tussen 'intuïtief het goede teken' en 'in de verkeerde richting'. Het eerste geval wordt meestal verder niet behandeld, terwijl het tweede geval voorzien wordt van een min of meer omslachtige redenering die het gevonden resultaat toch aannemelijk maakt. Zo vond Poppe [1986] dat nachtelijke ongevallen relatief vaak voorkwamen op kruispunten *mèt* verlichting. Verondersteld werd vervolgens dat vooral de om allerlei verschillende redenen onveilige kruispunten al van verlichting voorzien waren. Omdat het 'allerlei verschillende redenen' waren konden die elk voor zich geen belangrijke rol in de analyse spelen.

Maar ook wanneer met het 'goede teken' vindt, kan er in werkelijkheid sprake zijn van een dergelijke omslachtige redenering die de implicaties van het gevonden verband toch geheel anders doet zijn.

Elvik [1992] wijst er op dat ongevallen meer en meer een (zeldzaam) toevalsproces worden doordat mensen leren van hun fouten. Enerzijds worden daardoor de verschillen tussen individuen kleiner, anderzijds moet men er rekening mee houden dat sommige groepen meer kans krijgen te leren dan andere groepen. Bijv., in het zuiden van Zweden is de ongevallenkans op besneeuwde wegen ca. 10 maal zo groot als op onbesneeuwde. In het noorden van Zweden is die verhouding ca. $2\frac{1}{2}$. In het noorden komen besneeuwde wegen dan ook veel vaker voor: men rijdt daar ca. 40% op besneeuwde wegen, tegen het zuiden zo'n 6%.

Ook de keuze van de onderzoekseenheden kan van belang zijn. Bijleveld [1993] analyseerde o.a. de gegevens over de effecten van de invoering van de gordel-draagplicht. Wanneer de gegevens per kalenderjaar geanalyseerd worden is er geen effect, wanneer maandelijks verschoven 12-maandsperioden geanalyseerd worden is er wèl sprake van een effect!

5. De techniek

Onder techniek verstaan we in dit verband het soort algoritmes (en de daarbij behorende aannames over de eigenschappen van de gegevens) dat gebruikt wordt om de hiervoor beschreven modellen te 'fitten' op de gegevens.

5.1. Autocorrelatie

Een eerste onderscheid daarbij is de vraag of gebruik gemaakt wordt van de wetenschap dat de gegevens een tijdreeks vormen of niet. Men kan de waarnemingen altijd beschouwen als een aantal losse observaties van een proces, en ze als zodanig onderzoeken. Wanneer men echter weet dat de observaties in de tijd gezien met elkaar verband houden, dus wanneer de uitkomst op een later tijdstip gedeeltelijk voorspeld kan worden uit de waarneming op een eerder tijdstip, dan is sprake van 'autocorrelatie'. Men kan autocorrelatie vaststellen door bijvoorbeeld een reeks observaties te correleren met dezelfde reeks, één plaats verschoven in de tijd. Soms echter vindt men op die manier géén autocorrelatie, maar zou die wel gevonden zijn wanneer de reeks bijvoorbeeld twee plaatsen verschoven zou zijn in de tijd. Er is dan een hogere-orde autocorrelatie. Dergelijke extra informatie (*autocorrelatie*) kan in de analyse meegenomen worden.

Vorko-Jović & Jović [1992] schakelen dit effect juist expliciet uit, door van rangordes gebruik te maken. In het algemeen is bij tijdreeksanalyse, en ook bij Gaudry [1984], het bestaan van autocorrelaties en het onderzoek daarnaar een belangrijk element in de theorie.

5.2. Poisson-proces

Het wel of niet plaatsvinden van een ongeval kan in hoge mate beschouwd worden als een toevalsproces (daarbij kan hier in het midden gelaten worden of dat ligt aan de eigenschappen van het proces zelf, of aan onze onbekendheid met dat proces). Uit veel onderzoek is gebleken dat een dergelijk proces beschreven kan worden als een Poisson-proces. Waar een statistische analyse van het materiaal plaatsvindt, lijkt daar ook steeds rekening mee te worden gehouden. In een aantal gevallen [Fridstrøm, 1991] wordt ook expliciet aandacht besteed aan het feit dat *wel* het aantal ongevallen als een Poisson-proces beschreven kan worden, maar het aantal slachtoffers *niet* (doordat er meer slachtoffers per ongeval kunnen vallen). Voor elke geografische eenheid die beschouwd wordt is de verwachte waarde voor dat Poisson-proces echter verschillend. Wanneer de analyse zich uitstrekt over meerdere eenheden dient daar rekening mee gehouden te worden. Zo is bijvoorbeeld de som van ongevallen die voortkomen uit een aantal Poisson-processen (eventueel met verschillende Poisson-parameters) zelf ook weer een Poisson-proces, maar die geldt niet voor de verdeling van de ongevallen zelf. Als de Poisson-parameters verschillen, dan is in dat geval de variantie van dat aantal ongevallen groter dan het gemiddelde aantal. Bij een Poisson-verdeling zijn gemiddelde en variantie aan elkaar gelijk.

Niet duidelijk is of in alle gevallen met dergelijke effecten rekening gehouden is, en op welke wijze dat plaatsgevonden heeft.

6. Een kritische beschouwing

In de voorgaande paragrafen zijn al een aantal kanttekeningen geplaatst die in het algemeen aandachtspunten aangeven bij het type analyses waarbij hier sprake van is. Hier zal dit explicieter t.a.v. de bestudeerde literatuur gebeuren. Omdat het werk van Gaudry in theoretisch opzicht het meest uitgebreid beschreven is wordt deze beschouwing aan de hand daarvan gedaan.

6.1. Het model zelf

De verdeling in verschillende niveaus in het DRAG-model van Gaudry wordt logisch beargumenteerd en wordt ook consequent doorgevoerd. Deze sluit ook goed aan bij de fasering zoals deze bij het SWOV-prognosemodel voor de verkeersveiligheid is gedacht. Overigens gaat ook het al langer geleden gepresenteerde SWOV-fasemodel voor de verkeersveiligheid uit van dit zelfde onderscheid. In dit (theoretische) model werden wel meer fasen onderscheiden.



Afbeelding 2. De conceptuele overeenkomst tussen het DRAG-model (links), het SWOV-fasemodel (midden) en het concept-SWOV-prognosemodel (rechts)

Andere modellen (zoals bij Fridstrøm [1991] en Fridstrøm e.a. [1992]) moeten noodgewongen pas bij de interpretatie bepaalde effecten onderscheiden in een effect op bijv. de hoeveelheid verkeer en een (evt. tegengesteld gericht) effect op de ongevals- of letselkans.

De statistische onderbouwing van het DRAG-model lijkt gezond. Wel is e.e.a. veel ingewikkelder dan bij andere toepassingen, maar dat is onvermijdelijk wanneer men zowel rekening moet houden met autoregressieve karakter van de gegevens, als met de verschillende onderscheiden niveaus.

Wel moet worden opgemerkt dat, hoewel op een correcte (en optimale) manier rekening lijkt te worden gehouden met de onderlinge correlaties tussen verklarende variabelen, die onderlinge correlaties bij de gegeven verklaringen nog wel eens uit het oog schijnen te worden verloren. Ook rijst de vraag of bij het grote

aantal te schatten parameters het model nog wel stabiel blijft (m.a.w., of andere steekproeven uit dezelfde populatie wel tot dezelfde geschatte parameters zouden leiden).

6.2. Het gebruik van de resultaten

In tegenstelling tot bij een aantal andere studies echter lijkt het onderscheid tussen gedisaggregeerde verbanden en geaggregeerde gegevens bij Gaudry weinig aandacht te krijgen. Bij de interpretaties lopen die een aantal keren door elkaar. In de voorbereidende hoofdstukken wordt getracht een basis te leggen door op *het individuele niveau* gebaseerde econometrische nutsmodellen toe te passen. Overigens is dat minder uniek voor verkeersonderzoek dan gesuggereerd wordt – in de jaren '70 deed Hamerslag dat al volop, terwijl Volmuller eind jaren '60 daar al voorzichtig mee begonnen was (al ging het daar om prognoses van de hoeveelheid verkeer, en niet om de verkeersveiligheid). Zonder verdere onderbouwing worden echter individuele verbanden vertaald in verbanden die dan ook op het geaggregeerde niveau zouden moeten gelden.

Het omgekeerde vindt plaats bij de interpretatie van de gevonden verbanden. Zo wordt er een relatie gevonden tussen de verschillende te verklaren variabelen en het percentage vrouwen dat zwanger is. Uit de grafieken (blz. 147 in [Gaudry, 1984], zie par. A.4.5) blijkt dat de variatie in deze variabele in hoge mate een één-jaarsvariatie is. Onduidelijk blijft toch waarom ervan uitgegaan kan worden dat juist deze variabele verantwoordelijk is voor de te verklaren variabelen die ook een belangrijke één-jaarsvariatie kennen (blijkens de grafieken op blz. 28 t/m 38 in [Gaudry, 1984], zie par. A.4.5).

Gaan we er nu verder even van uit dat er inderdaad sprake is van een relatie tussen het percentage zwangere vrouwen in de populatie en de hoeveelheid verkeer, het aantal en de ernst van de ongevallen, dan is het opvallend dat Gaudry van dit resultaat op geaggregeerd niveau onmiddellijk overstapt naar een veronderstelling dat het hormonale evenwicht van individuele vrouwen daar verantwoordelijk voor is. Door Fridstrøm [1991] wordt er al op gewezen dat dan nagegaan zou moeten worden of een dergelijk effect relatief sterk optreedt bij ongevallen waarbij 15–45 jarige vrouwen betrokken zijn, en vrijwel afwezig bij andere ongevallen.

Er blijkt inderdaad enig ondersteunend medisch materiaal te vinden voor de veronderstelling dat het zwanger zijn invloed kan hebben op de rijvaardigheid. Overigens gaan vijf van de genoemde onderzoeken in op het onomstreden feit dat tijdens de zwangerschap belangrijke hormonale veranderingen plaatsvinden (wat opzich niets over de rijvaardigheid zegt), één onderzoek gaat over het effect van hormonen op de uitvoering van mechanische taken door ratten, en één onderzoek over het effect van 'orale anti-conceptiva' op de kans op ongevallen, wat iets anders is dan een zwangerschap.

Maar waarom is niet eerst een poging gedaan te onderzoeken of onder de slachtoffers inderdaad een relatief groot aantal zwangere vrouwen te vinden is? En waarom is niet verondersteld dat zenuwachtige aanstaande *vaders* een gevaar op de weg vormen? Of dat een grotere activiteit van medische diensten een rol speelt?

Met andere woorden, op het gevonden statistische resultaat kunnen, met even veel validiteit, heel andere hypothesen worden gebaseerd. Dezelfde vragen rijzen bij de interpretaties van een aantal andere gevonden relaties, met name waar het de effecten van overheidsmaatregelen betreft, zoals wetten en politietoezicht.

In veel van de besproken publikaties wordt er bovendien impliciet van uit ge-

gaan dat een eenmaal gevonden verband ook in de toekomst zijn geldigheid blijft behouden. Dat dat niet zonder meer op hoeft te gaan heeft Partyka geïllustreerd met een evaluatie van een eigen onderzoek. In 1984 publiceerde zij een lineair regressiemodel op een tijdreeks van 22 jaar, waarmee het aantal verkeersdoden in de Verenigde Staten gerelateerd werd aan het aantal werklozen, het aantal werkenden, en de overige bevolking. Met een dummyvariabele voor de jaren na 1974 (de oliecrisis en de invoering van de maximum snelheid in de VS) werd een correlatiecoëfficiënt van 0,98 bereikt. Destijds waarschuwde zij zelf al voor het maken van voorspellingen. In een later artikel [Partyka, 1991] komt zij op de analyse terug. Het model zoals dat in 1982 geschat was geeft voor de jaren daarna een snel groeiende afwijking van de werkelijkheid. Voor het eerste jaar is de afwijking al groter dan voor enig jaar daarvoor (9%), voor 1990 is de afwijking al 58%): het aantal doden is veel lager gebleven.

6.3. Conclusies

Met enkele vraagpunten over de statistische stabiliteit lijken de modellen zoals die beschreven worden door Gaudry [1984], Fridstrøm [1991] en Fridstrøm e.a. [1992], verantwoorde resultaten te kunnen opleveren over de mogelijke correlaties van een aantal achtergrondvariabelen. In veel gevallen betreft het variabelen die op zich betrekkelijk autonoom zijn en zich niet lenen voor overheidsingrijpen. Het gaat om bevolkingsomvang, weer, seizoen e.d. Waar het gaat variabelen die zich meer lenen voor overheidsingrijpen rijst de vraag in hoeverre de op geaggregeerd niveau gevonden verbanden hun geldigheid behouden op het individuele niveau waarop ze geacht te worden werken. Harde bewijzen kunnen daar uiteraard niet verwacht worden, maar enige evidentie uit andere bronnen zou daar nuttig kunnen zijn.

7. Mogelijkheden in de Nederlandse situatie

De theoretische (statistische) onderbouwing van modellen die op geaggregeerd niveau tijdreeksen van een aantal 'verklarende' variabelen in verband brengen met verkeersveiligheidsmaten lijkt op zich ver genoeg gevorderd om ook in de Nederlandse situatie een toepassing mogelijk te maken. De toepassingen op gegevens uit Quebec en uit Noorwegen laten zien dat het ook voor kleinere geografische eenheden zoals Nederland het in principe wel mogelijk is significante verbanden aan te tonen.

7.1. Exploratief en toetsend onderzoek

Alvorens de vraag de beantwoorden of de benaderingen zoals die in dit rapport beschreven zijn ook in Nederland tot relevante onderzoektoepassingen kunnen leiden moet eerst bekeken worden wat het doel daarvan zou moeten zijn. Daarvoor maken we onderscheid tussen de verschillende stappen die bij het onderzoeken van een probleemveld gemaakt kunnen worden, en de onderzoeksmethoden die daarbij horen. Hier kunnen we volstaan met het onderscheid tussen 'exploratief' en 'toetsend' onderzoek.

Bij exploratief onderzoek werkt men in het algemeen niet met expliciet uitgewerkte hypothesen, maar probeert men in het verzamelde materiaal verbanden op het spoor te komen. Uiteraard is de scheiding tussen deze benadering en het toetsend onderzoek niet 100% scherp te trekken. Ook bij exploratief onderzoek worden tevoren veronderstellingen geformuleerd, al was het maar om te bepalen welke gegevens verzameld dienden te worden.

Wanneer er tussen de verzamelde gegevens verbanden gevonden zijn (dat wil zeggen, er waren in statistische zin correlaties), dan is de volgende stap het opstellen van een mogelijke verklaring daarvoor. Die verklaring dient in de eerste plaats gericht te zijn op het (aggregatie)niveau waarop ook de variabelen bepaald zijn. Daarnaast kunnen ook hypothesen geformuleerd worden over effecten op andere aggregatieniveaus, deze zullen echter aan de hand van andere gegevens nader onderzocht moeten worden. De geformuleerde verklaringen (interpretaties) dienen vervolgens zo goed mogelijk getoetst te worden. Daarmee komen we in de tweede van de hiervoor onderscheiden fasen.

Bij het toetsende onderzoek worden de veronderstellingen over de in de gegevens aanwezige verbanden veel meer expliciet vooraf geformuleerd. Vervolgens wordt gekeken of dit verband ook in het materiaal aanwezig is, en hoe groot de kans is dat dit slechts een toevalstreffer is. Voor dit laatste is een aantal veronderstellingen over de aard van het materiaal noodzakelijk. Deze hebben te maken met de omvang en de representativiteit van de steekproef, en de verdelingseigenschappen van de verschillende kenmerken.

7.2. Consequenties voor toepassingen in Nederland

Een aantal van de besproken onderzoeken lijkt zowel exploratief als toetsend te willen werken. Zowel in het werk van Gaudry [1984] als dat van Fridstrøm e.a. [1992] is géén sprake van expliciet geformuleerde hypothesen, worden aan de hand van de gevonden correlaties mogelijke verklaringen geformuleerd maar is óók een belangrijk deel van het werk gericht op het bepalen van statistische significanties. Het is goed om bij mogelijke toepassingen op Nederlands materiaal een onderscheid tussen deze twee stappen van het analyseproces te maken. Het zoeken naar mogelijke interessante verbanden wordt dan niet belast met een overmaat aan aandacht voor de toetsing van een en ander.

Een aantal van de besproken onderzoeken onderscheiden zich op één aspect duidelijk van veel in Nederland gebruikelijk onderzoek: *het analysetechnische model sluit aan bij het conceptuele model*. Hiermee wordt het volgende bedoeld. In Nederland is het gebruikelijk onderscheid te maken tussen de verschillende fasen die bij verkeer en bij het ontstaan van een ongeval voorkomen: maatschappelijke activiteiten — verplaatsingsbehoefte — vervoerwijzekeuze — routekeuze — verkeer — ontmoetingen — conflicten — ongevallen — slachtoffers. Het door de SWOV geformuleerde fasemodel heeft dat verder geëxpliciteerd. Voor elk van die fasen kan op verschillende aspecten de ontwikkeling in de tijd in variabelen op geaggregeerd niveau worden vastgelegd. Het aantal onderzoeken dat dergelijke variabelen in één model met elkaar in verband brengt is echter zeer beperkt gebleven. Het blijkt nu mogelijk dat wel te doen. Een dergelijke poging kan dus ook in Nederland ondernomen worden, wanneer het materiaal daarvoor beschikbaar gekregen kan worden.

Dit hoeft niet noodzakelijkerwijs met dezelfde analysetechnieken te gebeuren die hier beschreven zijn. Het GLIM-pakket bijvoorbeeld zou dergelijke exploratieve analyses ook mogelijk maken, zonder dat de statistische toetsing een overmaat van aandacht vraagt. De modellen die Gaudry gebruikt, hebben het voordeel dat adequaat rekening gehouden wordt met een aantal problemen die inherent zijn aan de analyse van tijdafhankelijke gegevens. Andere pakketten bieden die mogelijkheden waarschijnlijk ook, zij het niet altijd op dezelfde wijze. Zowel in Duitsland als in Frankrijk (en wellicht ook in België) wordt een 'nationale variant' van het model van Gaudry voorbereid. De eerste indruk is dat daarmee een forse hoeveelheid werk gemoeid is. Dit heeft zowel te maken met het verzamelen van extra gegevens, als met het 'vertalen' van het model van Gaudry naar de nationale situatie. Het gebruik van andere analysetechnieken die dezelfde conceptuele mogelijkheden bieden kan daarom een goed alternatief zijn. Zo'n model zou minimaal in staat moeten zijn een aantal niveaus te onderscheiden, zoals dat in het DRAG-model gebeurt.

7.3. Gegevens

Welke gegevens wenselijk zijn hangt in de eerste plaats af van het aantal niveaus dat onderscheiden wordt. De vraag is of in de Nederlandse situatie (veel fietskilometers) de vervoerwijzekeuze niet een extra niveau zou behoren te zijn, ten opzichte van de niveaus die in het DRAG-model onderscheiden zijn (vervoerbehoefte, hoeveelheid verkeer, ongevallen).

Dit wordt versterkt door het feit dat het weer in een aantal onderzoeken naast de expositie een relatief belangrijke variabele blijkt te zijn. Het weer heeft in Nederland waarschijnlijk een grote invloed op de keuze voor de fiets of een ander vervoermiddel.

Te voren dient ook de vraag gesteld te worden voor welk doel de eventuele kennis uit een dergelijke analyse gebruikt zou kunnen worden. Wanneer het de primaire doelstelling is om prognoses of scenarioberekeningen te ondersteunen moet wel bedacht worden dat tot nu toe veel van de gevonden relevante variabelen (zoals bijv. het weer) zich niet lenen voor prognoses, laat staan voor enigerlei ingreep van de overheid.

Ten slotte is er de vraag naar de beschikbaarheid van gegevens. Er is in Nederland een veelheid aan statistisch materiaal beschikbaar. Ten eerste de in het verkeersveiligheidsonderzoek veel gebruikt tijdreeksen die direct met verkeer en verkeersveiligheid te maken hebben (mobiliteit, ongevallen en slachtoffers). Daarnaast echter bestaan er ook veel tijdreeksen (met zowel jaar- als maandgegevens) die informatie geven over aspecten van andere niveaus van het verkeerssysteem. Deze gegevens zijn voor een gedeelte reeds bij de SWOV beschikbaar, veel andere gegevens kunnen bij het CBS worden betrokken. Ook

uit andere bronnen (bijv. KNMI) zijn relevante gegevens te krijgen. Ten aanzien van brandstofverbruik, wagenpark, inwonertal, weglengte, alcoholconsumptie, economische toestand, enz. enz. zijn gegevens beschikbaar. De lengte en de frequentie van de reeksen zal echter verschillen. Er zijn echter voldoende gegevens om op de verschillende niveaus een aantal aspecten in een variabele vast te leggen.

7.4. Het voorstel concreet

Er zijn mogelijkheden voor exploratief onderzoek naar de vraag of de relaties tussen de aspecten van het verkeers- en vervoersysteem (op verschillende niveaus van het systeem) ook in één model te vatten zijn. De in dit rapport besproken onderzoeken geven aan dat er methoden zijn om dergelijk exploratief onderzoek uit te voeren.

De ervaringen wijzen echter ook uit dat het opzetten van een volledig uitgewerkt model veel tijd en menskracht vergt. Het is ook mogelijk op kleinere schaal te beginnen, met direct beschikbaar gegevensmateriaal, en met direct toepasbare statistische technieken.

Het voorstel voor de volgende stap kan dan in de volgende punten samengevat worden.

Onderzoek starten

Het is goed mogelijk exploratief onderzoek naar de relaties tussen tijdreeksen op verschillende niveaus van het verkeerssysteem te starten. Als niveaus kunnen daarbij bijvoorbeeld onderscheiden worden:

- bevolking;
- economie;
- klimaat en weer;
- beschikbare infrastructuur;
- regelgeving;
- hoeveelheid verkeer;
- hoeveelheid ongevallen;
- hoeveelheid slachtoffers.

De precieze keuze van de te onderscheiden niveaus zal onder meer moeten afhangen van de prioriteiten die beleidsmatig gesteld worden: op welke van de verschillende beïnvloedingsmogelijkheden van de overheid moet de aandacht gericht worden?

Het onderzoek moet gericht zijn op het formuleren van een model. De eerste stap naar zo'n model is het aangeven (bijvoorbeeld in een 'stroommodel') hoe elk van de hiervoor genoemde niveaus de andere niveaus beïnvloedt.

De gegevens

Voor dergelijk onderzoek is het nodig een aantal tijdreeksen op verschillende niveaus te selecteren, die elk een aspect op dat niveau 'meten'.

Voor veel van de hiervoor genoemde niveaus zijn tijdreeksen (meestal ook in de vorm van maandelijks gegevens) beschikbaar. Voor elk van de niveaus kan een aantal mogelijkheden genoemd worden:

- bevolking: totaal aantal, aantal jongeren, aantal bejaarden;
- economie: beroepsbevolking, aantal werklozen, prijsindexcijfer;
- klimaat en weer: aantal millimeters neerslag, aantal uren zon, gemiddelde windsnelheid;
- beschikbare infrastructuur: weglengte (evt. afzonderlijk autosnelweg, overig buiten de bebouwde kom, binnen de bebouwde kom);
- regelgeving: van kracht zijn gordelwet, alcoholwet, APK, bromfietshelm, enz.;
- hoeveelheid verkeer (personenkilometers naar vervoerwijze);

- hoeveelheid ongevallen (naar vervoerwijze);
- hoeveelheid slachtoffers (naar ernst).

Dit overzicht is niet uitputtend, maar geeft aan welke mogelijkheden beschikbaar zijn.

In een onderzoekopzet dient allereerst nagegaan worden wat de kwaliteit van de gegevens is en in hoeverre nog aanpassingen nodig zijn. Daarbij moet ook duidelijk worden in hoeverre daar nog (aanvullende) stappen gezet moeten worden (kosten, wellicht privacy-maatregelen).

De meeste gegevens zijn echter zonder aanvullende kosten bij de SWOV beschikbaar.

Het model schatten

Na deze twee acties is enerzijds bepaald welke niveaus onderscheiden worden en hoe deze niveaus mogelijk samenhangen, anderzijds zijn de tijdreeksen geselecteerd die dat moeten operationaliseren.

Nu zullen de relaties tussen de verschillende aspecten die op de niveaus onderscheiden worden ingevuld moeten worden: de parameters van het model moeten geschat worden. Daarvoor kunnen bestaande bij de SWOV direct toepasbare technieken ingezet worden. Mogelijkheden daarvoor zijn bijvoorbeeld het toepassen van een GLIM-analyse of pad-analyses. De SAS-procedure CALIS biedt daarvoor mogelijkheden, met bijvoorbeeld LISREL-analyses (lineair structural relations) en auto-regressieve modellen.

Zo mogelijk moeten interpretaties aan de gevonden relaties gegeven kunnen worden, zodanig dat dit kan leiden tot toetsbare hypothesen. Het onderzoek kan echter niet tegelijkertijd gericht zijn op het toetsen van die hypothesen.

Uitvoering

Deze eerste stap moet tot twee soorten resultaten leiden.

In de eerste plaats moeten er inhoudelijke conclusies getrokken kunnen worden over de vraag in hoeverre met behulp van dergelijke geaggregeerde gegevens meer inzicht in de relaties binnen het verkeers- en vervoersysteem kan worden verkregen. De gevonden relaties moeten ook geïnterpreteerd worden.

In de tweede plaats moeten ook eventuele praktische en methodologische problemen bij de uitvoering van vervolgstappen aan de orde komen. Deze moeten gericht zijn op de toepasbaarheid van de gebruikte analysetechnieken en op de beschikbaarheid en de bruikbaarheid van de gegevens. Wanneer voor een volgende stap andere analysetechnieken zouden moeten worden gebruikt moet aangegeven worden aan welke aanvullende voorwaarden die dan zouden moeten voldoen.

Ook moet aangegeven worden op welke punten voor een volgende stap meer of betere gegevens verkregen zouden moeten worden. Daarbij moet zo goed mogelijk inzicht gegeven worden in de omvang van de daartoe te verrichten werkzaamheden.

Bij de uitwerking van deze activiteiten kan gebruik gemaakt worden van ervaringen die intussen elders opgedaan worden. In internationaal verband (FERSI) wordt een aantal analyse-activiteiten gecombineerd (o.a. workshop 'Models for Traffic and Traffic Safety', met bijdragen uit Duitsland, Groot-Brittannië, de Scandinavische landen, Frankrijk en Nederland).

Literatuur

- Bos, J. (1993). *Prognosemodel mobiliteit en verkeersveiligheid; deelstudie 1: Risicoprognoses*. R-93-64. SWOV, Leidschendam.
- Box & Cox (1964). *Journal of the Royal Statistical Society*, 211–243.
- Bijleveld, F. (1990). *Dynamische processen*. SWOV, Leidschendam.
- Bijleveld, F. (1993). *Point processes in traffic safety analysis*. Master's thesis.
- Durkheim, E., (1951). *Suicide*. Free Press.
- Elvik, R. (1992). *Explaining randomness in accident counts: Three mechanisms generating the Poisson-distribution*. Institute of Transport Economics TØI.
- Fridstrøm, L. (1991). *In favor of aggregate econometric accident models*. Paper presented at the 6th International Conference on Travel Behaviour, Quebec, May 22-24, 1991.
- Fridstrøm, L.; e.a. (1992). *The contribution of exposure, weather, daylight, and randomness to the variation in accident counts: A four-country analysis*. Paper presented at the 6th World Conference on Transport Research, Lyon, June 29 - July 3, 1992.
- Gaudry, M. (1984). *DRAG, un modèle de la demande routière, des accident et de leur gravité, appliqué au Québec de 1956 à 1982*. nr. 359. Centre de Recherche sur les Transports, Université de Montréal.
- Gaudry, M. (1989). *Responsibility for accidents; relevant results selected from the DRAG model*. *Canadian Business Law Journal*, 16, 1: 21–33.
- Gaudry, M. (1990). *An outline of the DRAG comprehensive road safety model*. Transportation in Canada.
- Partyka, S. (1991). *Simple models of fatality trends revisited seven years later*. *Accident Analysis & Prevention*, 23: 423–430.
- Poppe, F. (1986). *De verkeersonveiligheid in Noord-Brabant IX; Onderzoek kruispunten tweede fase*. SWOV, Leidschendam.
- Robinson, W.S. (1950). *Ecological correlations and the behaviour of individuals*. *American Sociological Review*, 15: 351–357.
- Vorko-Jović, A. & Jović, F. (1992). *Macro model prediction of elderly people's injury and death in road traffic accidents in Croatia*. *Accident Analyses & Prevention*, 24, 6: 667–672.

Bijlage A. Enkele samenvattingen

A.1. Oudere verkeersslachtoffers in Kroatië

Vorko-Jović & Jović [1992] gaan uit van het jaarlijkse aantal verkeersdoden en -gewonden onder ouderen in de periode 1971–1986. Dit wordt in verband gebracht met het aantal inwoners, het aantal geregistreerde motorvoertuigen, de brandstofconsumptie en de weglengte. In de analyse worden alleen samengestelde variabelen gebruikt als brandstofverkoop per weglengte (benadering van verkeersintensiteit) en aantal motorvoertuigen per inwoner (de motorisatie). Ook de jaarlijkse verandering in die variabelen wordt in het model geïntroduceerd. De analyse wordt uitgevoerd als een rangorde-regressie, waarbij dus niet de absolute waarden van onafhankelijke variabelen worden gemodelleerd, maar het rangorde-nummer dat voor die variabele in elk jaar geldt. In feite is dit een transformatie die voorafgaande aan de analyse op het materiaal wordt uitgevoerd. Waarom gekozen is voor deze transformatie wordt niet duidelijk. De selectie van het model gaat volgens “essentially a trial and error procedure”, en wordt verder niet toegelicht.

Het model dat het ‘best’ voorspelt, verklaart het aantal verkeersslachtoffers onder ouderen uit de verkeersintensiteit minus de verandering in de verkeersintensiteit (dit lijkt de verkeersintensiteit van het vorig jaar te moeten zijn, maar daar wordt niet op ingegaan), een constante gedeeld door de motorisatie en een constante maal de verandering in het vorig jaar in de motorisatie.

Geconstateerd wordt dat het model goed voorspelt. Er wordt niet ingegaan op mogelijke interpretaties van de gevonden relaties.

A.2. Pleidooi voor econometrische modellen

Fridstrøm [1991] houdt een algemeen pleidooi voor het toepassen van (geaggregeerde) econometrische modellen in de verkeersveiligheidsanalyse. Deze zijn immers bedoeld om te pogen ‘oorzakelijke’ verbanden te leggen bij gegevens (vnl. tijdreeksen) die vrijwel uitsluitend op niet-experimentele wijze verkregen worden. In veel verkeersveiligheidsanalyses wordt met hetzelfde probleem geworsteld. Hoewel hij erkent dat het gevaar bestaat dat statistische verbanden op individueel niveau geïnterpreteerd worden, stelt hij dat vooralsnog het omgekeerde vaker lijkt te gebeuren. Bij wijze van voorbeeld worden met behulp van het GLIM-pakket ‘quasi-likelihood’ schattingen uitgevoerd op maandelijkse gegevens. Verschillende verkeersonveiligheidsmaten worden in verband gebracht met gegevens over expositie, weer, daglicht, weglengte, registratie, voertuiginspectie, politietoezicht, gordelgebruik, rijervaring en alcoholverkoop. Statistisch gezien zijn de resultaten bevredigend. Expositie blijkt veruit de belangrijkste verklarende variabele.

Fridstrøm concludeert dat gegeven de eigenschappen van onveiligheidsgegevens (weinig voorkomende toevalsgebeurtenissen volgens een Poissonproces) econometrische modellen uitstekend toegepast kunnen worden. De resultaten zullen wel altijd tegen gedisaggregeerde analyses getoetst moeten blijven. De kans op goede resultaten wordt het grootst wanneer men gegevens die de werking van een specifieke maatregel of omstandigheid meten getoetst kunnen worden aan dat gedeelte van de populatie dat voor die maatregel gevoelig is. Als bijvoorbeeld gevonden wordt dat het totale aantal ongevallen stijgt met het percentage zwangere vrouwen onder de bevolking [Gaudry, 1989], dan zou volgens Fridstrøm een relatief veel sterker effect gevonden moeten worden onder ongevallen waarbij vrouwen tussen de 15 en 45 jaar betrokken zijn, en geen effect bij andere ongevallen.

A.3. De vier Noordse landen

Onder auspiciën van de ‘Noordse commissie voor verkeersveiligheidsonderzoek’ hebben de onderzoeksinstituten uit Noorwegen, Zweden, Finland en Denemarken een gezamenlijk onderzoek uitgevoerd naar de relatie tussen het maandelijks aantal geregistreerde ongevallen en een aantal achtergrondvariabelen, zoals expositie, weers- en lichtomstandigheden, en registratiesysteem. De methode van analyseren was er op gericht een scheiding aan te brengen tussen de ‘ruis’ in de gegevens die bij zulke relatief zelden voorkomende gebeurtenissen als ongevallen relatief groot is, en de systematische variatie die geweten kan worden aan verschillen in omstandigheden. Het waargenomen aantal ongevallen in een gebied r gedurende een tijdvak t is dan $y(r, t)$, de realisatie van een Poisson-proces met verwachting $\lambda(r, t)$. Deze verwachting wordt op zijn beurt getrokken gedacht uit een Γ -verdeling met parameter ξ . Het waargenomen aantal ongevallen krijgt dan een negatief-binomiale verdeling. Wanneer de verwachting gelijk aan $\mu(r, t)$ gesteld wordt, dan geldt voor de spreiding

$$\sigma^2(r, t) = \mu(r, t) (1 + \theta \mu(r, t))$$

met $\theta = 1/\xi$. Het is duidelijk dat wanneer $\theta = 0$, er weer een simpele Poisson-verdeling ontstaat.

Wanneer in een model blijkt dat $\theta > 0$, is de hoeveelheid onverklaarde variantie groter dan de normale witte ruis die in een perfect gespecificeerd model zou moeten worden verwacht. De conclusie die dan getrokken wordt is dat niet alle systematische variatie verklaard is.

Om de effecten van de ‘verklarende’ variabelen te schatten wordt nu gesteld dat

$$\mu(r, t) = e^{\sum_j x_j(r, t) \beta_j}$$

zodat de β_j 's geschat kunnen worden. Dit is het gegeneraliseerde Poisson regressiemodel, zoals vaak toegepast bij GLIM-analyses.

De auteurs zijn tevreden over de prestaties van het model. Hoewel het aantal te analyseren ongevallen soms klein was, ‘degenereerde’ de oplossing niet snel. Overigens was ook hier expositie de belangrijkste verklarende variabele, verantwoordelijk voor ca. 50% van de variatie. Daarnaast zijn weers- en lichtomstandigheden belangrijke factoren, maar de auteurs constateren dat de richtingen van het effect soms tegen de intuïtie ingaan. Men veronderstelt dat hieraan correlaties met andere variabelen, niet in het model, ten grondslag liggen.

A.4. De verantwoording van het DRAG-model

Het DRAG-model wordt uitgebreid beschreven door de ontwerper van het model in [Gaudry, 1984]. Omdat deze rapportage moeilijk verkrijgbaar is en bovendien uitsluitend in het Frans beschikbaar is, worden hier enkele passages uit die rapportage kort samengevat. Deze paragrafen zijn dus niet bedoeld als een kritische beschouwing. Overigens is volgens een telefonische mededeling van de auteur een Engelstalige versie in voorbereiding. Die zal ook enkele nieuwe analyses bevatten, terwijl het statistische gedeelte gewijzigd zal zijn. Dit zou vooral betrekking hebben op de statistische toetsen in verband met de Box-Cox transformaties. De significantie-grenzen zouden aldus scherper getrokken kunnen worden.

A.4.1. Inleiding

Het acronym DRAG staat voor *la Demande Routière, les Accidents et leur Gravité*, in het Engels (zoals de auteur zelf aangeeft, wat gewrongen) vertaald als *Demand for Road use, Accidents and their Gravity*. In de inleidende paragrafen wordt uiteengezet welke splitsing in verschillende niveaus achter deze drie

elementen schuil gaat.

Externe factoren leiden tot een zekere hoeveelheid weggebruik. Daaruit, en uit (gedeeltelijk) andere externe factoren volgt een ongevalsrisico. Ook de ernst van de ongevallen hangt af van (de omvang van het) weggebruik en weer een aantal externe factoren. Schematisch:

DR — external factors

A — DR, external factors

G — DR, external factors

Hieruit kan het aantal slachtoffers dan bepaald worden: weggebruik maal ongevalsrisico maal ongevallenernst.

Vervolgens wordt aangegeven hoe nutsfuncties (zoals die in de economie en de econometrie gebruikelijk zijn) gebruikt kunnen worden om de effecten van allerlei veranderingen in omstandigheden op de keuzen van individuen in beeld te brengen. Voorbeelden worden gegeven van veranderende omstandigheden en daarmee samenhangende keuzen die op elk van de niveaus weggebruik, ongevalsrisico en ongevallenernst invloed hebben.

Vervolgens wordt (in par. 3.1) in een aantal stappen de te modelleren variabelen beschreven:

- de benzine- en dieselverkoop;
- het brandstofverbruik voor verkeersdoeleinden;
- het aantal UMS-ongevallen, het aantal letselongevallen en het aantal dodelijke ongevallen;
- de morbiditeit (gewonden per letselongeval) en mortaliteit (doden per letselongeval);
- het aantal gewonden en het aantal doden.

Naast afhankelijke variabelen worden er zes categorieën onafhankelijke variabelen onderscheiden. Deze worden (in par. 3.2) kort aangeduid:

- prijzen van brandstof, openbaar vervoer, auto-onderhoud, enz.;
- motorisatie, kwantitatief en kwalitatief (gewicht, aanwezigheid gordels, enz.);
- omstandigheden, onderscheiden in drie soorten:
 - wetten, regels en politietoezicht;
 - voorzieningenniveaus van de onderscheiden wijzen van vervoer;
 - infrastructuur en weer;
- de ‘gebruikers’, in vier soorten variabelen beschreven:
 - algemeen (omvang, rijbewijsbezit);
 - leeftijd
 - geslacht
 - alertheid
- economische activiteiten
- diversen:
 - ‘administratieve’ beslissingen en procedures (o.a. registratie van ongevallen);
 - aggregatie en schaal (o.a. aantal dagen per maand);
 - seizoensinvloeden en constanten (o.a. dummy-variabelen voor de invloed van de wereld-Expo van 1967 en de Olympische Spelen van 1976);

A.4.2. De econometrische beschrijving

Vervolgens wordt (in par. 3.4) overgegaan op de modelmatige beschrijving, waarbij uitgegaan wordt van econometrische modellen. Er wordt een conse-

quent onderscheid gemaakt tussen het 'vaste' (systematische), en het 'toeval-
lige' (de fout) van het model.

Het vaste gedeelte kan dan als volgt worden beschreven (voor elk tijdvak t van
de N waarnemingen):

$$y_t^{(\lambda_y)} = \sum_{k=1}^K \beta_k X_{k_t}^{(\lambda_x)} + u_t$$

waarbij het model voor de fout u_t verder niet gespecificeerd wordt.

De notatie $y^{(\lambda)}$ staat voor de Box-Cox [1964] transformatie:

$$y^{(\lambda)} = \max_{\lambda} -\frac{2}{N} \ln \left(\frac{1}{N} \sum_{j=1}^N (x_j^{(\lambda)} - \bar{x}^{(\lambda)})^2 \right) + (\lambda - 1) \sum_{j=1}^N \ln x$$

waar

$$\bar{x}^{(\lambda)} = \frac{1}{N} \sum_{j=1}^N x_j^{(\lambda)}$$

en

$$x^{(\lambda)} = \begin{cases} \frac{x^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln x & \text{if } \lambda = 0 \end{cases}$$

Deze transformatie is bruikbaar wanneer men de vorm van de relatie tussen
variabelen nog niet kent. Voor $\lambda_x = \lambda_y = 1$ levert dit bijvoorbeeld een gewone
lineaire regressie, en voor $\lambda_x = \lambda_y = 0$ een 'gewone' logaritmische regressie.

Het model voor de fouten is:

$$u_t = \left[e^{\delta_0 + \sum_{m=1}^M \delta_m Z_{m_t}^{(\lambda_z m)}} \right]^{-\frac{1}{2}} v_t$$

$$v_t = \sum_{l=1}^r \rho_l v_{t-l} + w_t$$

De volgende twee doelen worden nagestreefd.

(1) De spreiding van de fout moet constant zijn. De Box-Cox transformatie
beïnvloedt de spreiding in de fout. Om een fout te krijgen waarvan de spreiding
constant is (homoscedasticiteit) hebben we een model nodig dat het mogelijk
maakt de heteroscedasticiteit te controleren en niet de vorm van het model
en spreidingfout door elkaar haalt: twee instrumenten voor twee doelen. Deze
strategie is al eerder algemener geformuleerd. Merk op dat de eerste vergelij-
king het voordeel heeft dat de klassieke heteroscedasticiteit als bijzonder geval
meegenomen is.

(2) De verkregen foutenresiduen dienen niet in de tijd gecorreleerd zijn ('white
noise'). In werkelijkheid wordt de tweede foutenvergelijking gebruikt om, met
een hogere orde autoregressief proces, een benadering te maken van een proces
dat zowel een autoregressief schema bevat als een gedeelte met een doorschui-
vend gemiddelde. De ervaring leert dat dit werkt (bij voldoende parameters
zijn de modellen equivalent).

De gedachte achter dit econometrische model is het balanceren tussen het werk
aan het vaste gedeelte en het werk aan het modelleren van de fout. Bij het vaste
gedeelte laten we de gegevens de functionele vorm bepalen, bij het fouten model
zoeken we de systematische of in een model te vatten informatie. Deze taken
worden gelijktijdig volbracht door het maximaliseren van de log-likelihood
van de $N - r$ waarnemingen. Onder de voorwaarde dat we veronderstellen
dat de w_t een normale verdeling hebben, en met weglaten van de eerste r
waarnemingen wordt de log-likelihood als volgt geschreven:

$$\mathcal{L} = -\frac{N-r}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sum_{t=l+r}^N \ln f(Z_t)$$

$$- \frac{1}{2} \sigma^2 \sum_{t=l+r}^N \left[\frac{u_t}{\sqrt{f(Z_t)}} - \sum_{l=1}^r \rho_l \frac{u_{t-l}}{\sqrt{f(Z_t)}} \right]^2$$

$$+ (\lambda_y - 1) \sum_{t=l+r}^N \ln y_t$$

waarbij $f(\mathcal{Z}_t)$ het gedeelte tussen vierkante haken uit de vergelijking voor u_t voorstelt en σ^2 de spreiding van w_t is die men terug vindt binnen de haken in het kwadraat.

Het algoritme wordt verder bescheven bij Liem et.al.

Dit evenwicht tussen de beschrijving van het model $y_t^{(\lambda)}$ (waar men de variabelen gebruikt waarmee men de afhankelijke variabele ‘begrijpt’; en de twee foutenvergelijkingen waar men de restfout u_t probeert te beschrijven (in de vorm van) w_t , (‘white noise’ met constante spreiding), combineert twee soorten van modellering.

In het eerste perspectief stelt men een regressiemodel op en verwaarloost men het werk aan het fouten model; in het uiterste geval veronderstelt men simpelweg dat u_t homoscedastisch is en niet lijdt onder autocorrelatie. Bovendien komt het vaak voor dat men een lineair model schat ($\lambda_y = \lambda_x = 1$) en een log-lineair ($\lambda_y = \lambda_x = 0$) en de resultaten als de ‘beste’ publiceert.

In het tweede perspectief interesseert men zich niet voor de verklarende variabelen, men probeert slechts de afhankelijke variabele y_t te reproduceren uit zijn voorgaande waarden en uit de voorgaande waarden van de fout w_t ; men noemt deze soort van analyse ‘tijdreeksanalyse’ of Box-Jenkins-analyse. De auteur gebruikt deze nu om de structuur vast te stellen.

Sinds enige tijd corrigeert men enigszins de uitschieters van deze vorm van ‘curve fitting’ door één of twee verklarende variabelen toe te voegen, vaak binaire dummies, die structurele veranderingen of ‘trapsgewijze’ fenomenen beschrijven. Deze verandering van Box-Jenkins (interventie-analyse of Box-Tiao) kan ook gebruikt worden om effecten te onderscheiden die geleidelijk veranderen.

Zie verder Wiorkowski & Heckard.

Gaudry stelt dat zijn model, in vergelijking met andere modellen die tot dan toe in de literatuur gevonden kunnen worden, zich op vijf punten onderscheidt:

- het geïntegreerd onderzoeken van de verschillende niveaus (na deze eerste onderscheiden te hebben);
- het gebruik van verschillende samengestelde functies, zowel voor brandstof als voor ongevallen en slachtoffers;
- het gebruik van maandelijkse tijdreeksen met een veelheid aan informatie;
- een poging om informatie te verkrijgen over zowel het aantal als de ernst van de ongevallen, wat iets anders is dan de informatie die men verkrijgt als men zich slechts interesseert voor het aantal slachtoffers;
- het gebruik van een soepele functionele vorm (Box-Cox) met een gelijktijdige correctie voor zowel autocorrelaties van hogere orde als een zeer algemene vorm van heteroscedasticiteit.

A.4.3. Interpretatie van de resultaten

Onderscheid wordt gemaakt tussen *calibratie* van de parameters en het *schatten* ervan.

Bij de calibratie vraagt men zich af of de doelfunctie (de kwadratensom) gevoelig is voor veranderingen in bijv. het aantal parameters, het toevoegen van een variabele, een verandering in de functionele vorm enz.

Bij het schatten gaat het om de vraag of de variabele in kwestie werkelijk ‘random’ is, en men gebruikt statistische theorieën om te bezien of de waarschijnlijkheid in significante mate afhangt van dergelijke veranderingen. Maar uiteindelijk gaat het om de robuustheid. Daar zijn echter slechts pragmatische methoden voor.

Als maat voor de significantie wordt steeds Student’s t gegeven, bij de resultaten

van een regressie analyses als waarde tussen haakjes, bij de elasticiteiten met de aanduidingen *, **, *** en /// (t is dan resp. 0-1, 1-2, 2-3 of ≥ 3).

Omdat de factoren in een regressie analyse weinig zeggen over de grootte van het effect van een variabele of de afhankelijke variabele (dat hangt bijv. af van de meeteenheden) worden ook de elasticiteiten gegeven. Omdat er sprake is van verschillende niveaus kan deze niet altijd direct berekend worden. De volgende formule wordt gebruikt:

$$\epsilon_{y, X_k} = \frac{\delta \mathbf{E}(y)}{\delta X_k} \frac{\bar{X}_k}{\mathbf{E}(y)} \rightarrow \frac{\delta y}{\delta X_k} \frac{\bar{X}_k}{\bar{y}}$$

(waarbij $\mathbf{E}(y)$ de mathematische verwachting voor y aanduidt).

Wanneer de verklarende variabele niet gedurende de hele periode 'aktief' is wordt een correctie toegepast. Voor een binaire variabele (waar geen partiële differentiaal vergelijking bestaat) wordt eenvoudig het relatieve verschil (met- zonder) berekend. In de tabellen wordt met een enkele onderstreping aangeduid wanneer de eerste correctie is toegepast, met een dubbele onderstreping wordt de tweede variatie bedoeld.

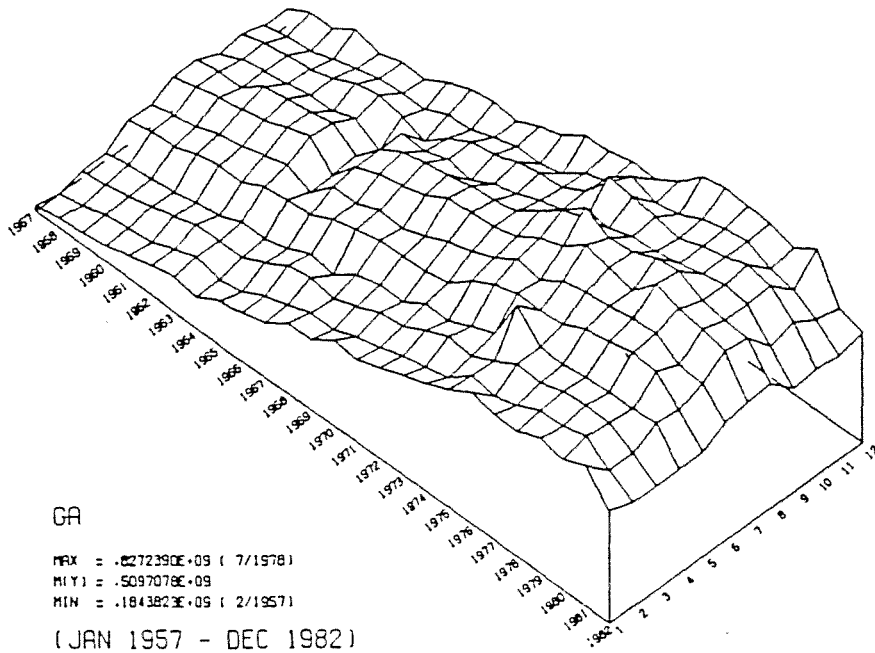
A.4.4. Enkele resultaten

Gaudry [1990] geeft nog een samenvatting van enkele interessante resultaten. Enkele punten daaruit.

- Hij constateert dat een aantal verklarende variabelen op de verschillen- de onderscheiden niveaus (weggebruik, ongevallenkans, letselkans) in verschillende richtingen werken. Was het onderscheid in niveaus niet gemaakt dan zouden een aantal effecten of niet gevonden zijn, of mogelijk verkeerd geïnterpreteerd zijn.
- Een stijging van de onderhoudskosten leidt tot minder, en minder ernstige ongevallen.
- Er is een grotere ongevallenkans maar een geringere ongevallenernst gekoppeld aan zwangerschap, vooral in de eerste twee maanden. Ook in het 'hoofdrapport' wordt veel aandacht hieraan, en aan een mogelijke verklaring, gegeven. De auteur formuleert de hypothese dat een verhoging van de verhouding tussen oestrogene hormonen en progesteron die niet gecompenseerd wordt door androgene hormonen de rijvaardigheid beïnvloeden.
- Een relatief grote alcoholconsumptie is gerelateerd aan verminderd aantal verkeersdoden. Verondersteld wordt dat een relatief groot gedeelte van die toename 'gebruikt' wordt door bestuurders die per hoofd nog weinig alcohol blijven gebruiken, door compenserend gedrag van deze 'kleine drinkers' neemt het totale risico af.

A.4.5. Enkele bladzijden uit het rapport

In dit rapport wordt (vnl. bij wijze van voorbeeld) verwezen naar enkele figuren uit het rapport van Gaudry dat hier kort werd samengevat. Op de volgende bladzijden volgen kopieën van de relevante bladzijden uit dat rapport.

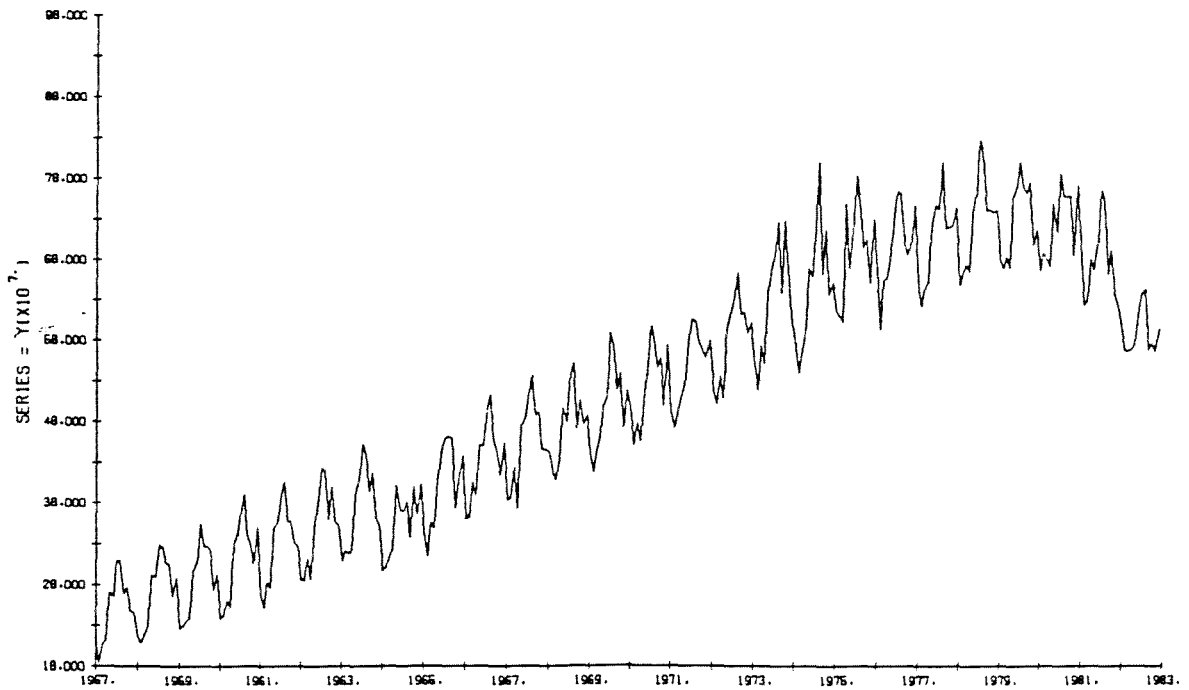


GA

MAX = .8272390E+09 (7/1978)
MIN = .5097078E+09
MIN = .1843823E+09 (2/1957)

(JAN 1957 - DEC 1982)

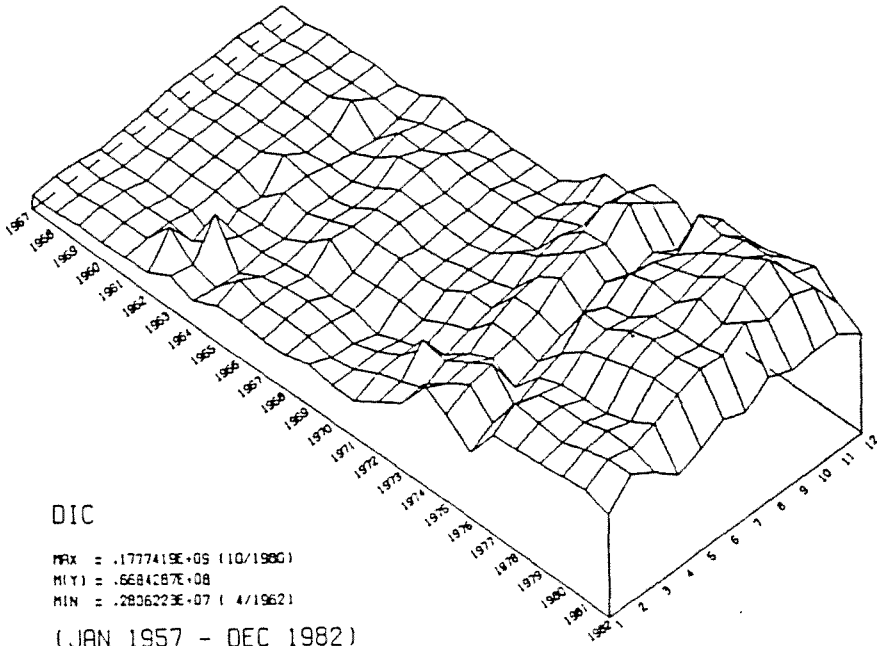
VENTES BRUTES D'ESSENCE AU QUEBEC (LITRES)



GA

VENTES BRUTES D'ESSENCE AU QUEBEC (LITRES)

(JAN 1957 - DEC 1982)

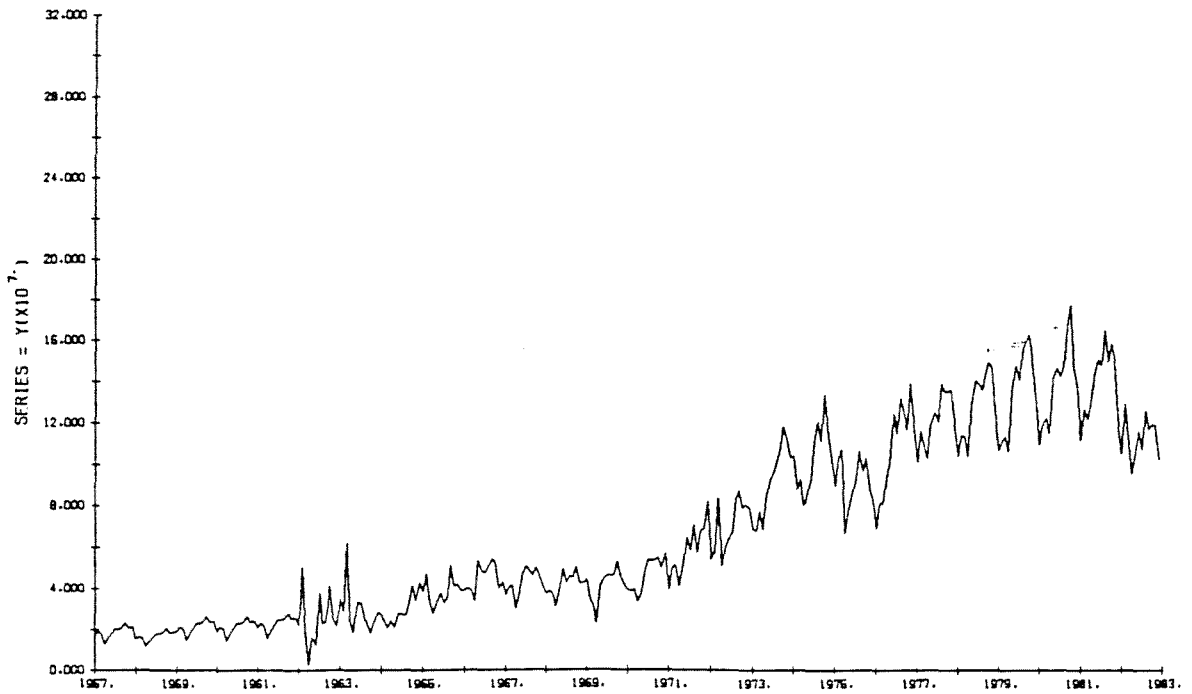


DIC

MAX = .1777419E+05 (10/1980)
MAY = .6684287E+08
MIN = .2806223E+07 (4/1962)

(JAN 1957 - DEC 1982)

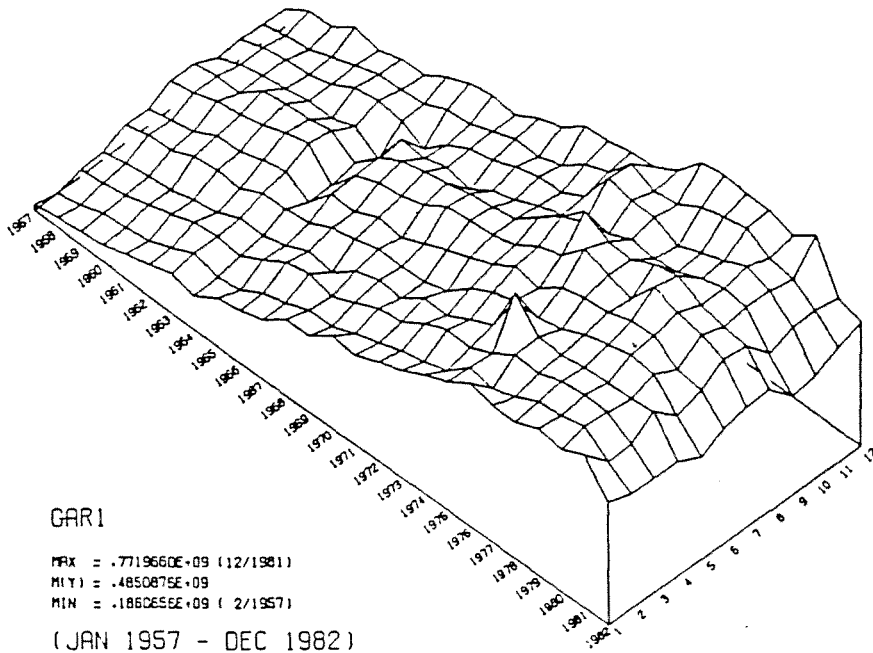
VENTES BRUTES RECORRIGES DE DIESEL AU QUEBEC (LITRES)



DIC

VENTES BRUTES RECORRIGES DE DIESEL AU QUEBEC (LITRES)

(JAN 1957 - DEC 1982)



GAR1

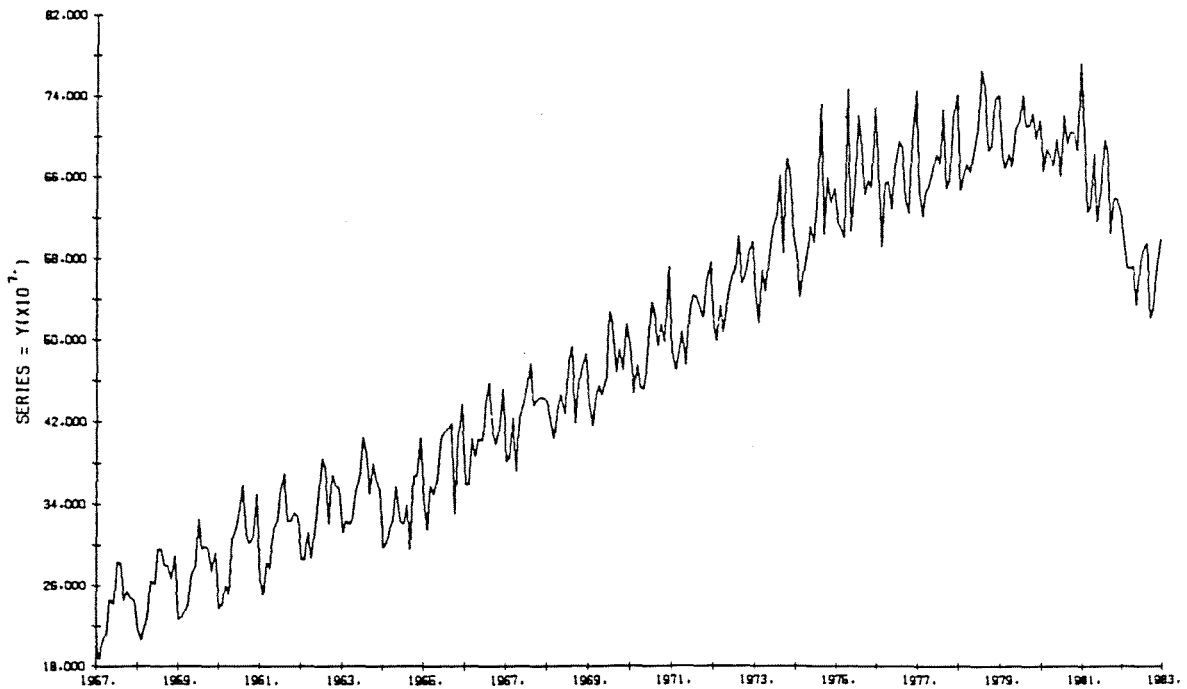
MAX = .7719660E+09 (12/1981)

MIY = .4850875E+09

MIN = .1860552E+09 (2/1957)

(JAN 1957 - DEC 1982)

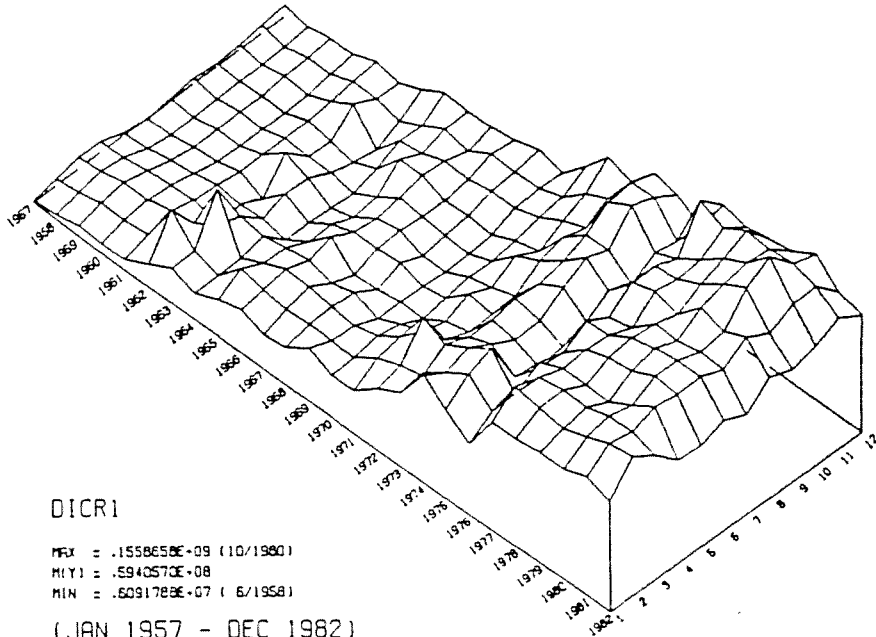
VENTES BRUTES D'ESSENCE ROUT. (MESURE 1) AU QUEBEC (LI)



GAR1

VENTES BRUTES D'ESSENCE ROUT. (MESURE 1) AU QUEBEC (LI)

(JAN 1957 - DEC 1982)

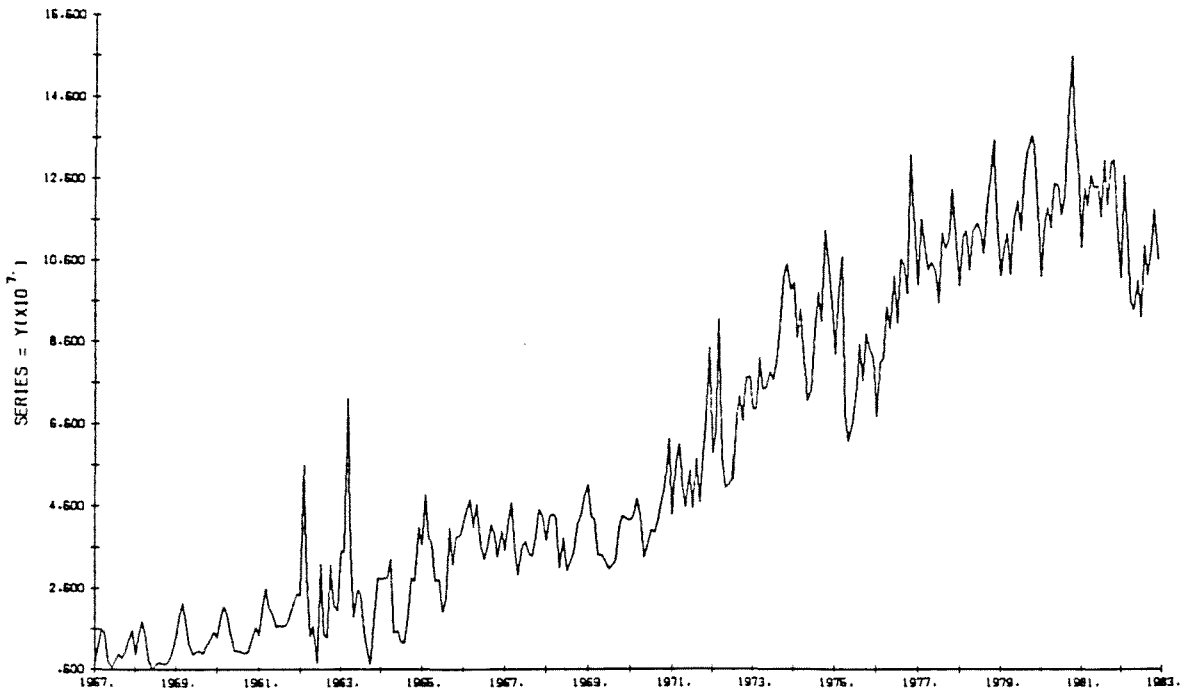


DICR1

MAX = .1558650E+09 (10/1980)
MIN = .6091783E+07 (6/1958)

(JAN 1957 - DEC 1982)

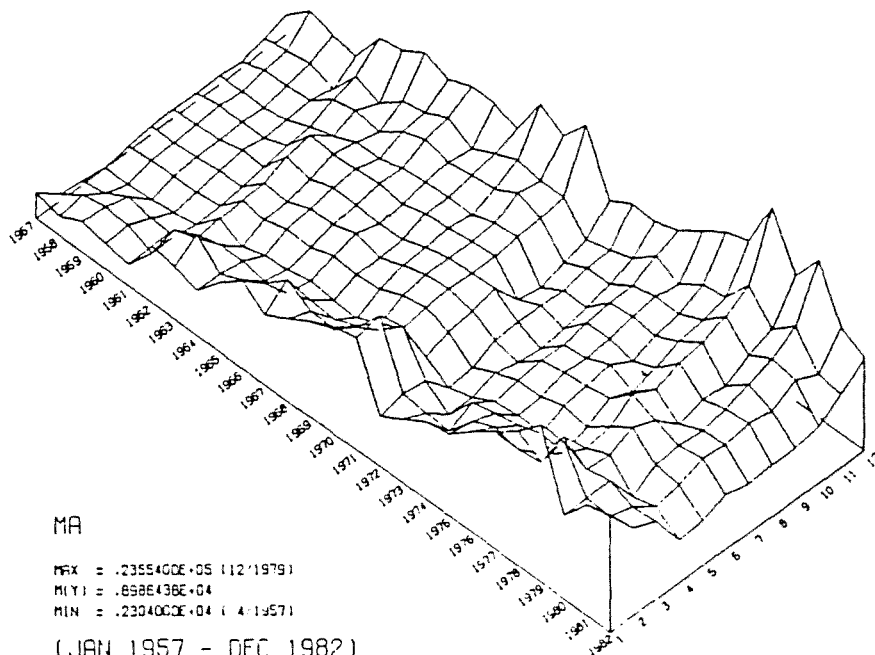
VENTES BRUTES RECOR. DE DIESEL ROUT. (M.1) AU QUEBEC (LI)



DICR1

VENTES BRUTES RECOR. DE DIESEL ROUT. (M.1) AU QUEBEC (LI)

(JAN 1957 - DEC 1982)



MA

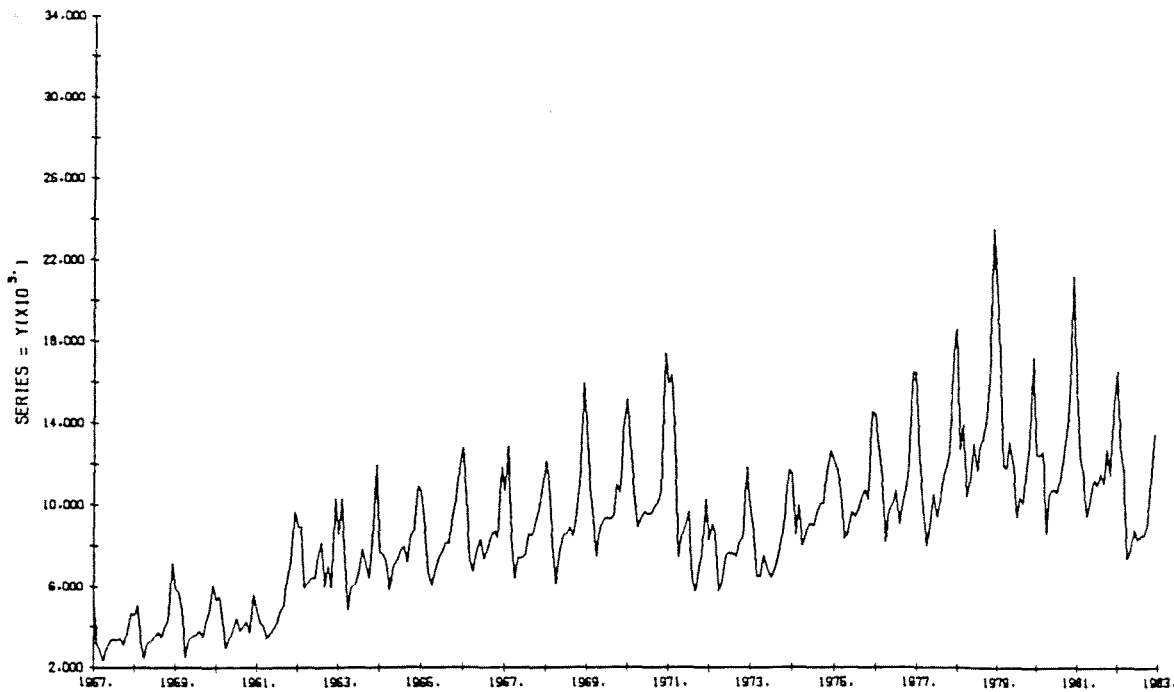
MAX = .2355400E+05 (12/1979)

MEAN = .658E+38E+04

MIN = .2304000E+04 (4/1957)

(JAN 1957 - DEC 1982)

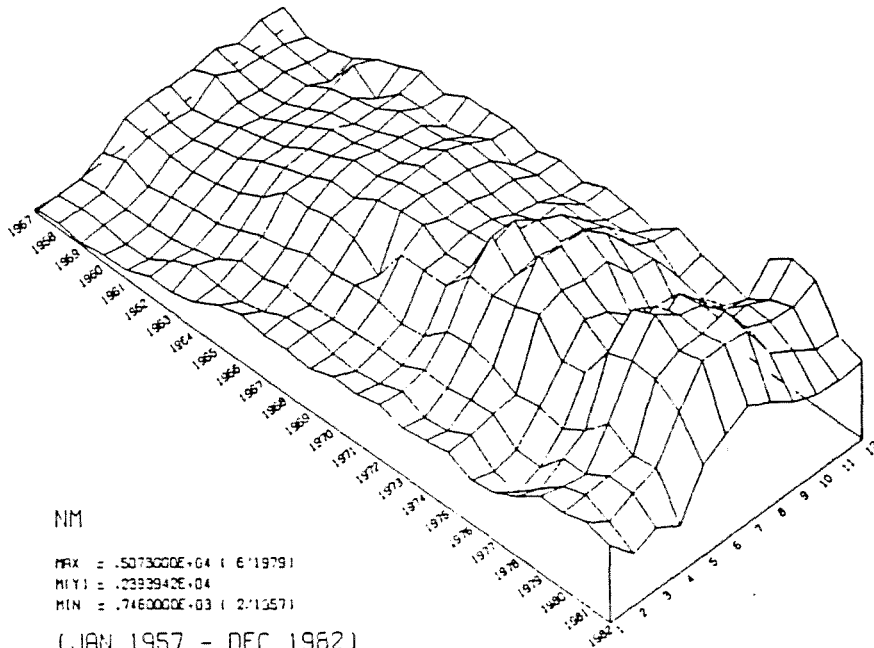
NOMBRE D'ACCIDENTS AVEC DOMMAGES MATERIELS AU QUEBEC



MA

NOMBRE D'ACCIDENTS AVEC DOMMAGES MATERIELS AU QUEBEC

(JAN 1957 - DEC 1982)



NM

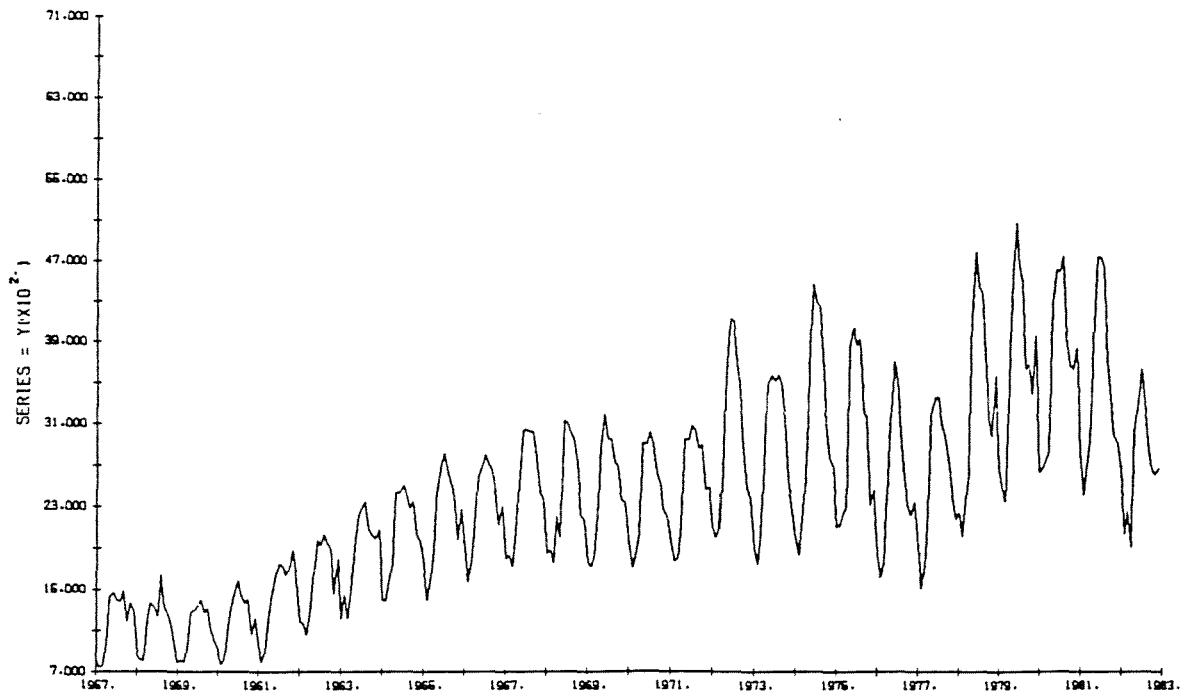
MAX = .5073000E+04 (6:1979)

MIYI = .2333942E+04

MIN = .7460000E+03 (2:1957)

(JAN 1957 - DEC 1982)

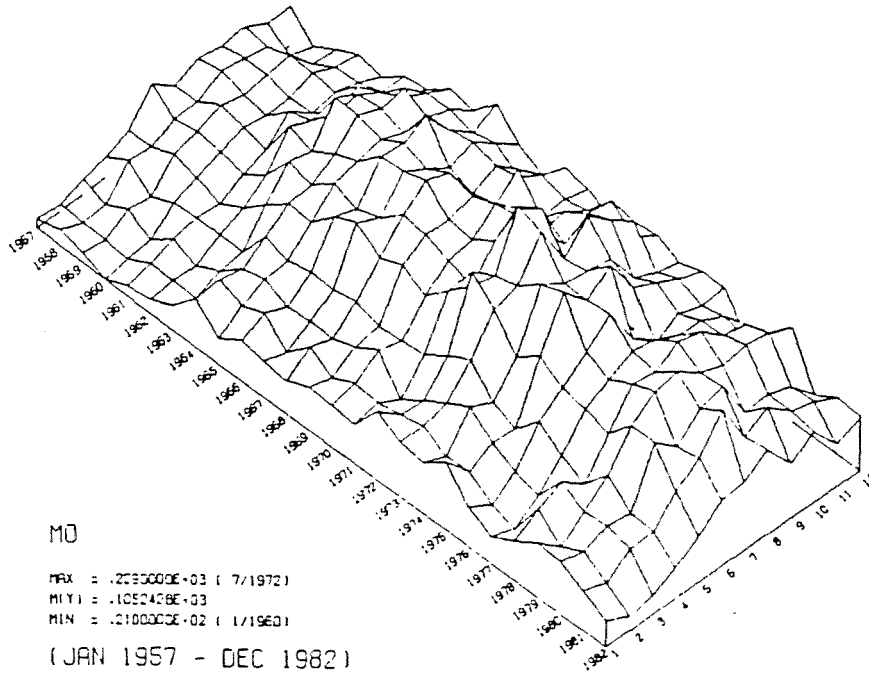
NOMBRE D'ACCIDENTS NON MORTELS (AVEC BLESSES) AU QUEBEC



NM

NOMBRE D'ACCIDENTS NON MORTELS (AVEC BLESSES) AU QUEBEC

(JAN 1957 - DEC 1982)



MO

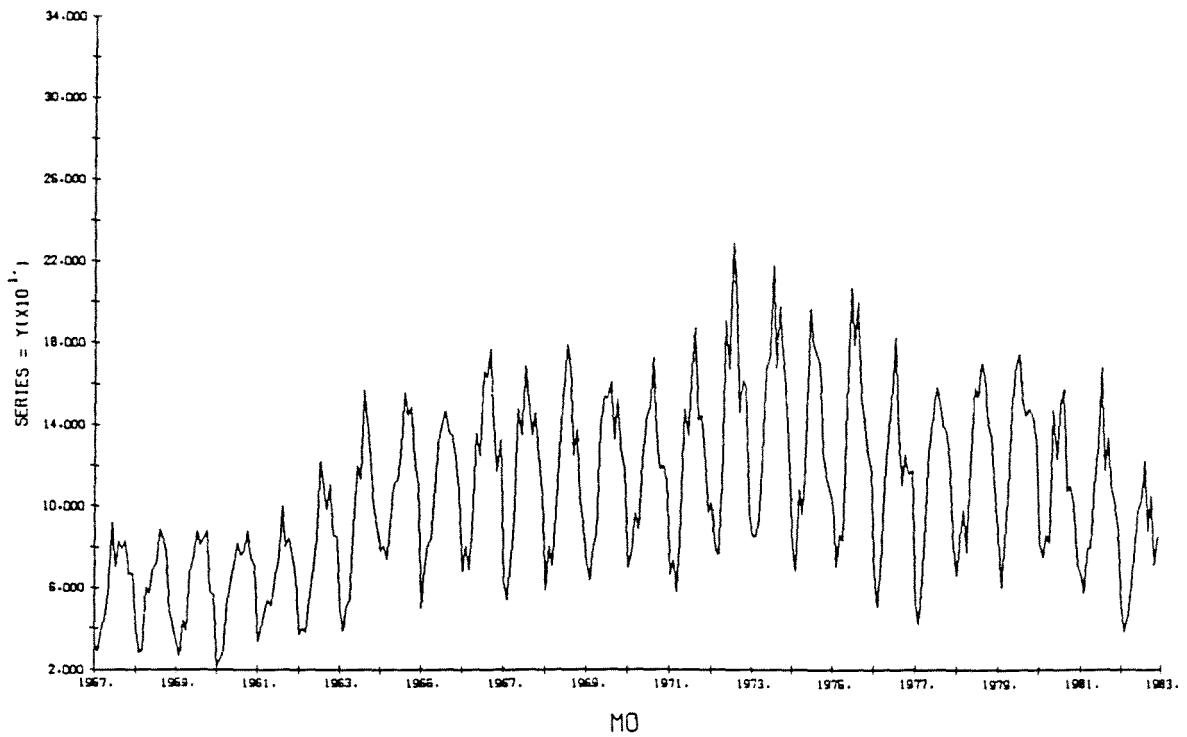
MAX = .2733000E+03 (7/1972)

MIY = .1052400E+03

MIN = .2100000E+02 (1/1960)

(JAN 1957 - DEC 1982)

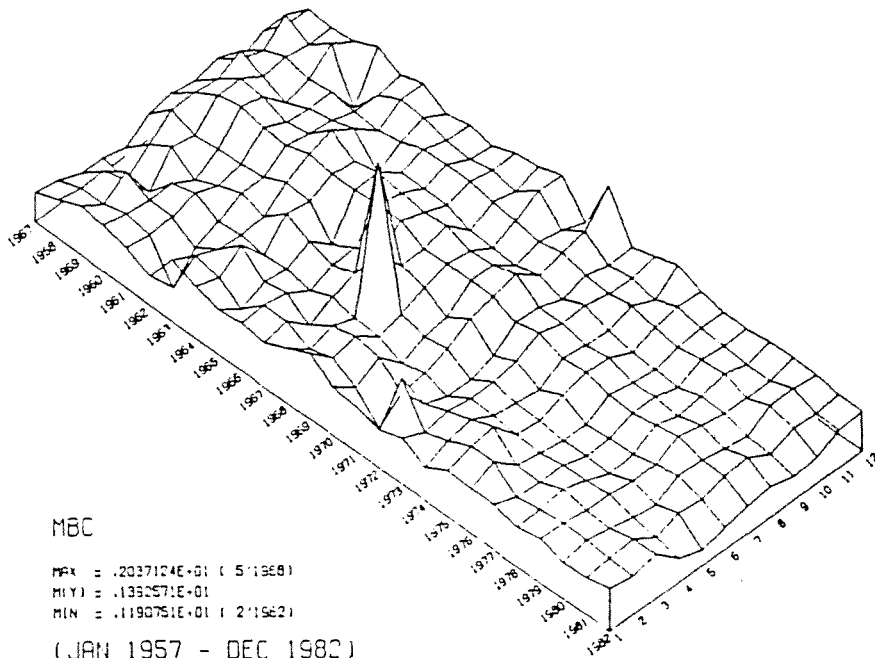
NOMBRE D'ACCIDENTS MORTELS (AVEC MORTS) AU QUEBEC



MO

NOMBRE D'ACCIDENTS MORTELS (AVEC MORTS) AU QUEBEC

(JAN 1957 - DEC 1982)

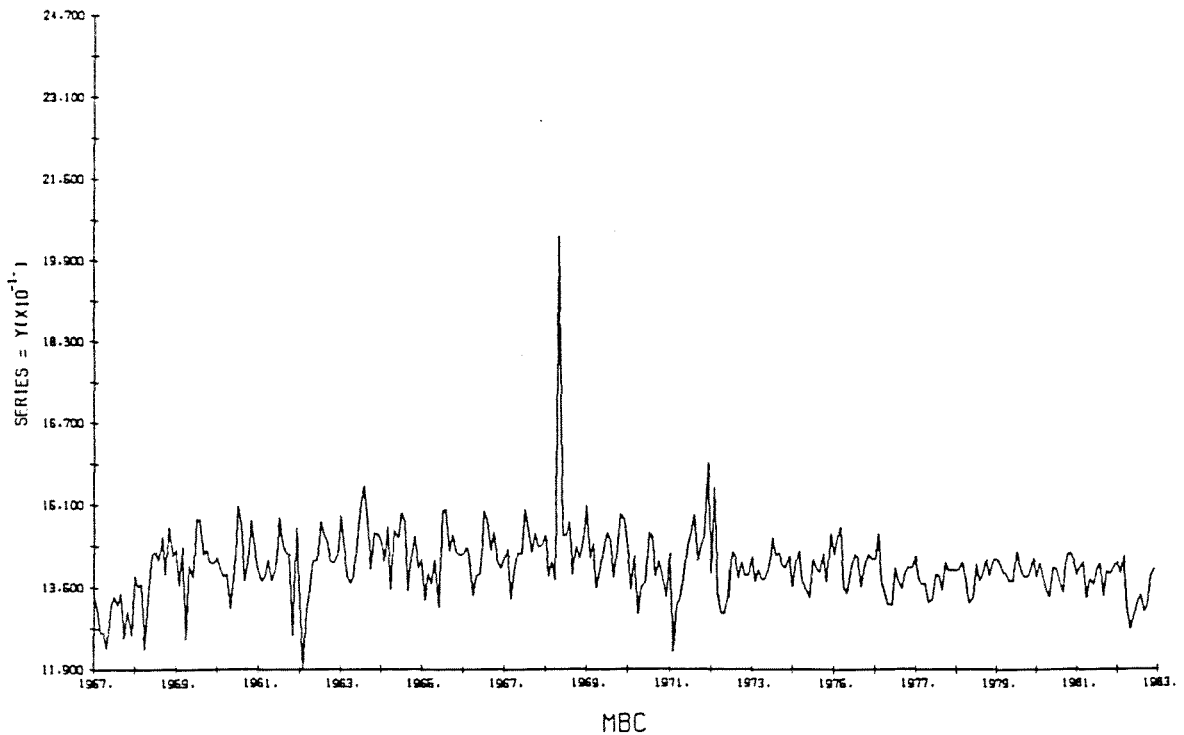


MBC

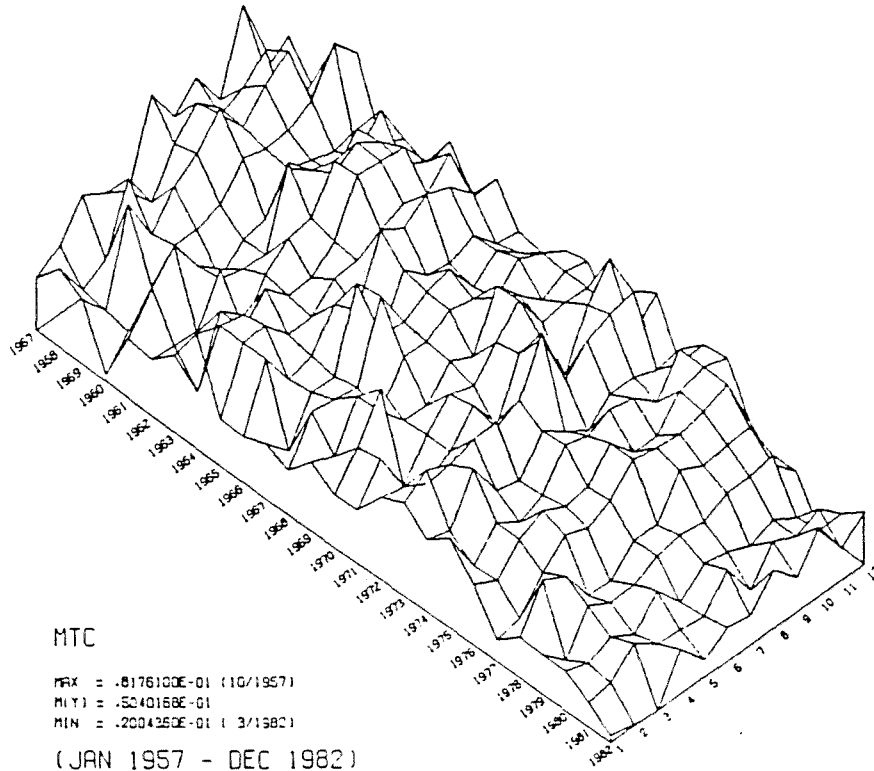
MAX = .2037104E+01 (5'1968)
MIN = .1190751E+01 (2'1962)

(JAN 1957 - DEC 1982)

MORBIDITE DES ACC. CORPORELS (BLESSES/ACC. CORPORELS)



MORBIDITE DES ACC. CORPORELS (BLESSES/ACC. CORPORELS)
(JAN 1957 - DEC 1982)

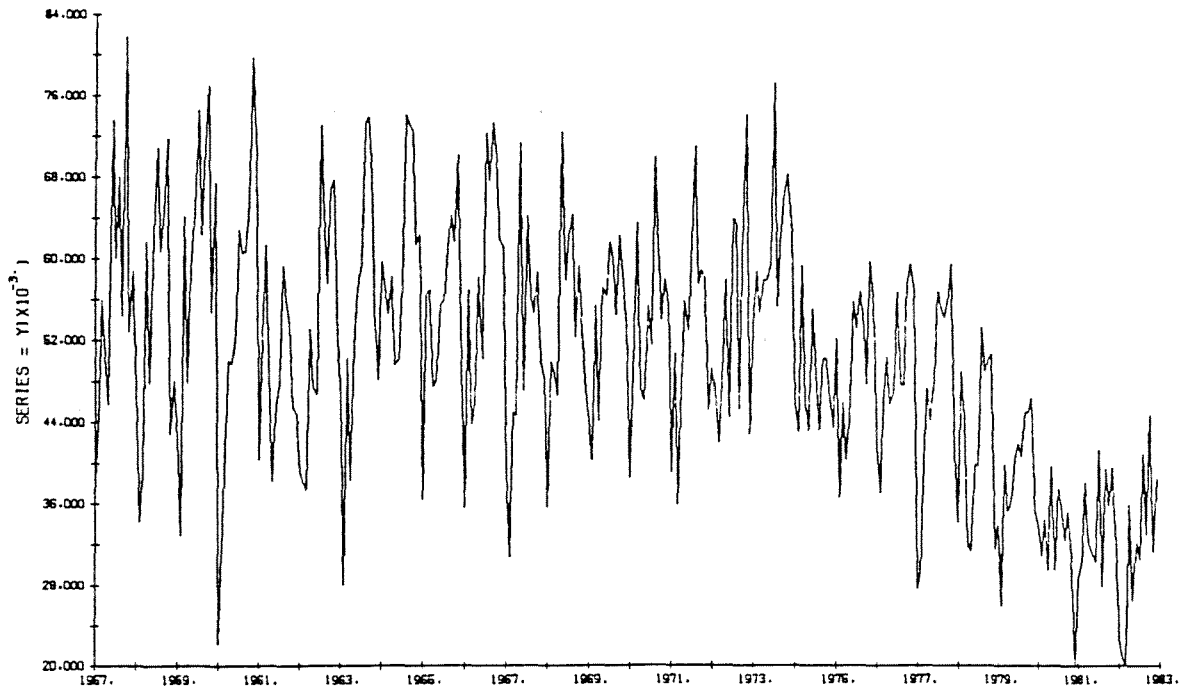


MTC

MAX = .0176100E-01 (110/1957)
MIN = .5040160E-01
MIN = .2004250E-01 (3/1982)

(JAN 1957 - DEC 1982)

MORTALITE DES ACC. CORPORELS (TUES/ACC. CORPORELS)



MTC

MORTALITE DES ACC. CORPORELS (TUES/ACC. CORPORELS)

(JAN 1957 - DEC 1982)

TABLEAU 9.9 : Élasticités directes et indirectes,
Y-S = consommateurs - sexe

A. Modèle de référence

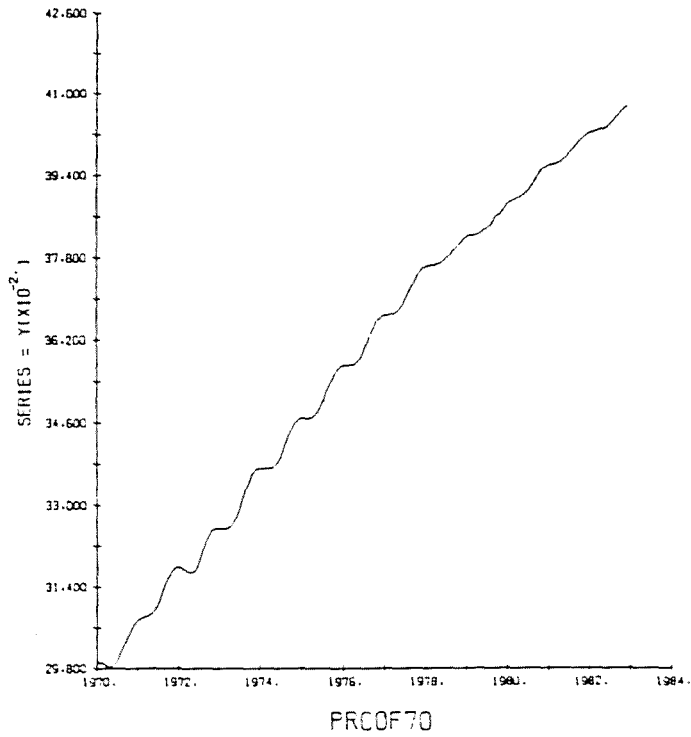
Y-S = CONSOMMATEURS - SEXE		GARI	DICRI	MA	NH	MO	COR	ACC	MBC	MTC	HT	DE	VI
PROB.COND.SOIT FEMME ENCEINTE	PRFEVM	-.198		.211	-.175	-.211	-.177	.127	.057	-.289	-.119	-.466	-.131
		..		-.172	-.095	-.116	-.096	-.156	-.013	-.149	-.110	-.245	-.114

B. Variante : désagrégation du nombre de femmes enceintes

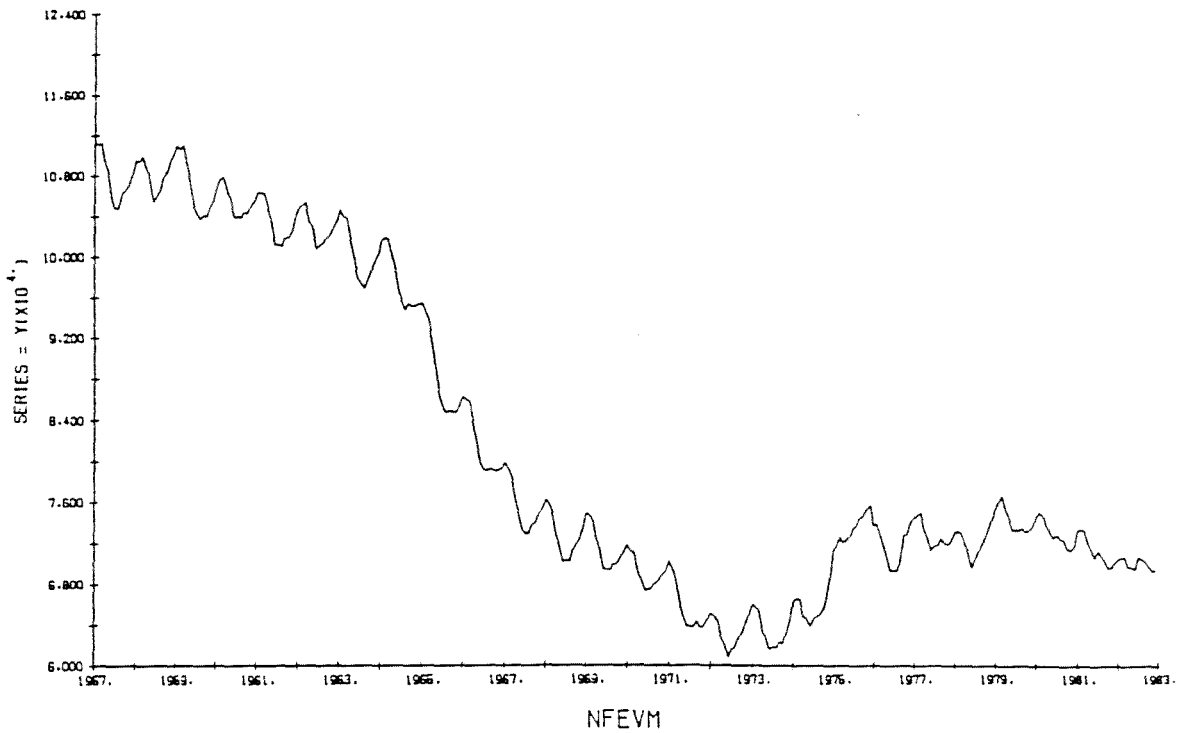
Y-S = CONSOMMATEURS - SEXE		GARI	DICRI	MA	NH	MO	COR	ACC	MBC	MTC	HT	DE	VI
PROB.COND. SOIT FEM.ENC. 1 MOIS	PRFE1	-.024		.229	-.037	.036	-.034	.172	.099	-.120	.064	-.184	.057
		.		-.019	-.011	-.017	-.012	-.017	-.001	-.017	-.013	-.028	-.013
PROB.COND. SOIT FEM.ENC. 2 MOIS	PRFE2	-.015		.139	.303	1.010	.333	.181	-.024	.314	.309	.647	.320
		.		-.012	-.007	-.010	-.037	-.011	-.001	-.010	-.008	-.018	-.008
PROB.COND. SOIT FEM.ENC. 3 MOIS	PRFE3	-.070		-.142	-.150	-.533	-.165	-.147	-.017	.097	-.182	-.068	-.178
		.		-.055	-.034	-.050	-.035	-.052	-.003	-.050	-.037	-.085	-.039
PROB.COND. SOIT FEM.ENC. 4 MOIS	PRFE4	.026		.062	-.084	-.628	-.106	.026	.027	-.042	-.079	-.149	-.082
		.		.021	-.013	.015	.013	.019	.021	.019	.014	.032	.015
PROB.COND. SOIT FEM.ENC. 5 MOIS	PRFE5	.028		.258	.052	-.115	.045	.212	.041	-.078	.086	-.034	.082
		.		.023	.014	.020	.014	.021	.001	.020	.015	.034	.016
PROB.COND. SOIT FEM.ENC. 6 MOIS	PRFE6	-.077		.003	.032	-.134	.025	.008	.001	-.317	.025	-.292	.015
		..		-.062	-.037	-.055	-.038	-.057	-.003	-.055	-.041	-.093	-.043
PROB.COND. SOIT FEM.ENC. 7 MOIS	PRFE7	-.111		-.231	-.115	.027	-.109	-.205	-.032	-.225	-.141	-.335	-.148
		..		-.069	-.054	-.079	-.055	-.082	-.004	-.079	-.059	-.134	-.062
PROB.COND. SOIT FEM.ENC. 8 MOIS	PRFE8	.022		-.044	-.005	.090	-.001	-.035	-.008	.202	-.010	.201	-.003
		.		.017	.010	.015	.011	.016	.001	.015	.012	.026	.012
PROB.COND. SOIT FEM.ENC. 9 MOIS	PRFE9	.020		-.067	-.133	-.076	-.129	-.080	-.031	-.179	-.158	-.307	-.164
		.		.016	.010	.014	.010	.015	.001	.014	.011	.024	.011

C. Variante : ajout de la proportion de conducteurs de sexe féminin

Y-S = CONSOMMATEURS - SEXE		GARI	DICRI	MA	NH	MO	COR	ACC	MBC	MTC	HT	DE	VI
PROPORTION CONDUCT.SEXE FEMIN.70-82	PRCOF70	.145		1.220	1.310	.300	1.268	1.230	-.069	-.717	1.199	.561	1.176
		.		.108	.048	.080	.049	.095	.010	.131	.059	.180	.064
PROB.COND.SOIT FEMME ENCEINTE	PRFEVM	-.196		.201	-.176	-.229	-.178	.119	.061	-.242	-.118	-.420	-.128
		..		-.146	-.064	-.108	-.066	-.129	-.014	-.178	-.080	-.244	-.086



PROPORTION DES CONDUCTEURS FEMININS 1970-82 AU QUEBEC
(JAN 1970 - DEC 1982)



NOMBRE DE FEMMES ENCEINTES AU QUEBEC
(JAN 1957 - DEC 1982)