

NN31545.0767

NOTA 767

december 1973

Instituut voor Cultuurtechniek en Waterhuishouding
Wageningen

BIBLIOTHEEK DE HAAFF

Droevendaalsesteeg 3a
Postbus 241
6700 AE Wageningen

ZWAARTEKRACHTSMODELLEN VOOR DE SPREIDING VAN
WOONGEBIEDEN, AFGELEID UIT RITDISTRIBUTIES
VAN WOON-WERKVERKEER

ir. Th. Michels

**BIBLIOTHEEK
STARINGGEBOUW**

Nota's van het Instituut zijn in principe interne communicatiemiddelen, dus geen officiële publikaties.

Hun inhoud varieert sterk en kan zowel betrekking hebben op een eenvoudige weergave van cijferreeksen, als op een concluderende discussie van onderzoeksresultaten. In de meeste gevallen zullen de conclusies echter van voorlopige aard zijn omdat het onderzoek nog niet is afgesloten.

Bepaalde nota's komen niet voor verspreiding buiten het Instituut in aanmerking.

1780/262



INHOUD

	Blz.
SAMENVATTING	1
1. INLEIDING	3
2. STRUCTUUR EN CALIBRATIE VAN HET ZWAARTE- KRACHTSMODEL	4
2.1. Zonering van het studiegebied	4
2.2. Formulering van het model	4
2.3. Calibratie van het model	11
2.4. Analyse van parameterwaarden	17
3. ENKELE TOEPASSINGEN VAN CALIBRATIE EN PARAMETER ANALYSE	18
3.1. Werkforensen op de gemeente Utrecht in 1960	18
3.2. Werkforensen op 14 centra in de Randstad in 1960	21
3.3. Beroepsbevolking van de drie Randstadprovincies in 1960	28
4. CONCLUSIES	35
LITERATUUR	36

SAMENVATTING

In de ruimtelijke ordening bestaat het verlangen, facetsgewijze interactiemodellen (bijv. wonen-werken, wonen-verzorging, wonen-recreatie), zoals die zijn ontwikkeld in de verkeertheorie, toe te passen als macroplanologische spreidingsmodellen voor diverse vormen van grondgebruik in grote gebieden voor lange termijnprognoses.

In deze nota wordt daarom een analyse verricht van de structuur van zwaartekrachtsmodellen, gecalibreerd op ritdistributies van woon-werkverkeer, teneinde na te gaan of de diverse modelparameters zich lenen voor voorspelling van grondgebruik.

Na formulering van het zwaartekrachtsmodel met zijn nevenvoorwaarden wordt een interpretatie van de erin voorkomende parameters gegeven. Daarbij wordt geconcludeerd, dat met het model geen lange termijnprognose van grondgebruik kan worden gegeven zonder de beschikking te hebben over in minstens twee verschillende jaren waargenomen ritdistributies. Door de parameters te correleren aan specifieke gebiedsvariabelen kan uit een tijdreeks een extrapolatie ervan naar de toekomst worden afgeleid.

Uiteengezet wordt op welke wijze het zwaartekrachtsmodel kan worden gecalibreerd op een ritdistributiematrix door schatting van de parameters met behulp van de 'maximum likelihood'-methode in een iteratieve rekenprocedure met een computer.

Aan de hand van een voorbeeld met één werkzone (Utrecht) wordt geïllustreerd dat de woonplaatskeuze in de omringende woonzones niet kan worden beschreven zonder daarbij de interactie met andere werkzones te beschouwen, en dat met de aldus gevonden parameters de bevolkingsontwikkeling in de woonzones niet kan worden voorspeld.

Op grond van een toepassing op de ritdistributie tussen 14 werkzones en 245 woonzones in de Randstad wordt geconcludeerd, dat de uit deze steekproef berekende parameterwaarden nog onvoldoende betrouwbaar zijn om te worden gecorreleerd met verklarende variabelen.

Met een calibratie op een vierkante ritmatrix voor 84 zones in de drie Randstadprovincies wordt aangetoond, dat de spreiding in de attractieparameters per zone bovendien storend wordt beïnvloed indien in het model slechts één afstandsfunctie voor het hele studiegebied wordt gehanteerd. De conclusie luidt dat het wellicht zinvol is, voor de werkzones individuele afstandsfuncties aan te passen aan de betreffende ritdistributies, en de parameters hierin te relateren aan specifieke variabelen van de ter plaatse werkende beroepsbevolking.

1. INLEIDING

Voor de beschrijving van de ruimtelijke relaties van diverse grondgebonden maatschappelijke activiteiten met behulp van wiskundige interactiemodellen wordt langzamerhand vrij algemeen de voorkeur gegeven aan het zwaartekrachtmodel. Operationele toepassingen ervan zijn tot nu toe vrijwel voorbehouden gebleven aan de planning van verkeersvoorzieningen in zeer specifieke situaties en op betrekkelijk korte termijn, bijvoorbeeld 5 à 10 jaar (o. a. GOUDAPPEL, 1970). Algemene toepasbaarheid van gevonden modelparameters in andere gebieden alsmede lange termijn-voorspellingen (10 à 20 jaar) worden met name in de weg gestaan door een beperkt inzicht in de specifieke parameterwaarden (zoals voor ritproductie, ritattractie en afstandsgedrag), en daardoor in de ontwikkeling ervan in de tijd.

In de ruimtelijke ordening bestaat grote behoefte aan integratie van facetsgewijze interactiemodellen (wonen-werken, wonen-verzorging, wonen-recreatie) in macroplanologische spreidingsmodellen voor het grondgebruik, toegepast op grote gebieden over lange perioden (BONE, 1971; BUCHANAN c. s., 1973). Hierbij dient men bekend te zijn met de differentiatie van parameterwaarden naar plaats en tijd.

Het in deze nota beschreven onderzoek beoogt: 1. het inzicht te vergroten in de mathematische structuur van het zwaartekrachtmodel door analyse van resultaten van een aantal calibraties; 2. het zwaartekrachtmodel te verfijnen door onderzoek naar verklarende variabelen voor de diverse relevante parameters. Eén en ander geschiedt aan de hand van toepassingen op woon-werkrelaties, hetgeen niet wegneemt, dat de ontwikkelde benaderingswijze evenzeer op andere facetten (verzorging, openluchtrecreatie) toepasbaar is.

2. STRUCTUUR EN CALIBRATIE VAN HET ZWAARTEKRACHTS-MODEL

2.1. Zonering van het studiegebied

Een interactiemodel beschrijft de distributie van relaties tussen verschillende vormen van grondgebruik. Deze relaties roepen verplaatsingen op van personen en goederen. Die verplaatsingen zijn te lokaliseren met behulp van hun begin- en eindpunten, oftewel de plaatsen van herkomst en bestemming. Bij toepassing van een distributiemodel worden daarom in een studiegebied onderscheiden een verzameling van N_I herkomstzones, $I = \{i : \text{herkomstzones}\}$ en een verzameling van N_J bestemmingszones, $J = \{j : \text{bestemmingszones}\}$. De verzamelingen I en J kunnen gelijk zijn: $I = J$, maar dit is niet noodzakelijk, hun doorsnede kan zelfs leeg zijn: $I \cap J = \emptyset$. Omvang en aantal van de onderscheiden zones hangt af van beschikbaarheid en detaillering van in te voeren gegevens, maar vooral van het doel dat met de modelanalyse wordt beoogd, met andere woorden, welke eisen men stelt aan de vorm waarin de resultaten worden gepresenteerd. Overigens dient vaak om praktische redenen het aantal zones te worden beperkt, bijvoorbeeld wegens de capaciteit van de gebruikte computer of die van verwerking van gegevens met de hand.

2.2. Formulering van het model

De interactie tussen een herkomstzone i ($i \in I$) en een bestemmingszone j ($j \in J$) wordt door het zwaartekrachtmodel in zijn algemene gedaante als volgt gedefinieerd:

$$T_{ij} = O_i A_j f(C_{ij}) \quad \text{voor alle } i \in I \text{ en alle } j \in J \quad (1)$$

Hierin is het aantal ritten T_{ij} tussen i en j een functie van een maat voor de ritproduktie O_i in zone i , van een maat voor de ritattractie A_j in zone j en van het gegeneraliseerde relatieve afstandsbezwaar C_{ij} tussen i en j ($f(C_{ij})$ neemt af naarmate C_{ij} toeneemt).

Indien nu het hierboven geformuleerde gravitatiemodel wordt toegepast op de distributie van werknemers over woonzones bij een

waargenomen aantal werknemers E_i , waarvan de rit in zone i start, dan moet (1) voldoen aan de eis, dat

$$\sum_{j \in J} T_{ij} = E_i \quad \text{voor alle } i \in I \quad (2)$$

zodat volgens (1)

$$O_i \sum_{j \in J} A_j f(C_{ij}) = E_i \quad \text{of} \quad O_i = \frac{E_i}{\sum_{j \in J} A_j f(C_{ij})} = a_i E_i$$

Hierdoor wordt (1)

$$T_{ij} = \frac{E_i A_j f(C_{ij})}{\sum_{j \in J} A_j f(C_{ij})} = a_i E_i A_j f(C_{ij}) \quad \text{voor alle } i \in I \text{ en alle } j \in J \quad (3)$$

Volgens (3) wordt dus het totaal aantal waargenomen ritten E_i van werknemers werkzaam in een zone i , over de woonzones j verdeeld volgens de relatieve produkten van attractie en afstandsfunctie $A_j \cdot f(C_{ij})$, en wel relatief, doordat de factor a_i zijnde de inverse som van deze produkten als distributiedeler fungeert.

Omgekeerd redenerend kan men de eis stellen, dat bij R_j waargenomen ritten van beroepsbevolking naar een woonzone, (1) voldoet aan

$$\sum_{i \in I} T_{ij} = R_j \quad \text{voor alle } j \in J \quad (4)$$

Substitutie van (1) in (4) geeft dan

$$A_j \sum_{i \in I} O_i f(C_{ij}) = R_j \quad \text{of} \quad A_j = \frac{R_j}{\sum_{i \in I} O_i f(C_{ij})} = b_j R_j$$

Hierdoor wordt (1)

$$T_{ij} = \frac{R_j O_i f(C_{ij})}{\sum_{i \in I} O_i f(C_{ij})} = b_j R_j O_i f(C_{ij}) \quad \text{voor alle } i \in I \text{ en } j \in J \quad (5)$$

In (5) verloopt voor elke woonzone j de verdeling van de gegeven beroepsbevolking R_j over de werkzones i volgens de relatieve producten $O_i \cdot f(C_{ij})$, waarbij de factor b_j als distributiedeler fungeert.

Indien tegelijkertijd aan beide randvoorwaarden (2) en (4) wordt voldaan, krijgt (1) de vorm

$$T_{ij} = \frac{E_i R_j f(C_{ij})}{\sum_{j \in J} A_j f(C_{ij}) \sum_{i \in I} O_i f(C_{ij})} = a_i E_i b_j R_j f(C_{ij})$$

voor alle $i \in I$ en alle $j \in J$ (6)

De distributiedelers a_i en b_j , in de literatuur ook wel competitie- of evenwichtsfactoren genoemd, zijn in (6) wederzijds afhankelijk en geven het zwaartekrachtsmodel een gecompliceerde structuur:

$$a_i = \frac{O_i}{E_i} = \frac{1}{\sum_{j \in J} b_j R_j f(C_{ij})} \quad \text{voor elke } i \in I \quad (7)$$

en

$$b_j = \frac{A_j}{R_j} = \frac{1}{\sum_{i \in I} a_i E_i f(C_{ij})} \quad \text{voor elke } j \in J \quad (8)$$

De factor a_i blijkt de inverse van de gesommeerde, met b_j gewogen 'nabijheden' van beroepsbevolking ($R_j \times$ afstandsfunctiewaarden). Omgekeerd is b_j de inverse som van de met a_i gewogen 'nabijheden' van arbeidsplaatsen ($E_i \times$ afstandsfunctiewaarden). Dit betekent, dat a_i en b_j elkaar wederzijds beïnvloeden. WILSON (1970, pag. 19 e.v.) wijst erop, dat de complexiteit van deze beïnvloeding in bepaalde gevallen slechts betrekkelijk is; indien namelijk slechts één E_i verandert, stel E_1 wordt E'_1 , dan wordt volgens (6)

$$T'_{ij} = a_1 E'_1 b_j R_j f(C_{1j}) \quad \text{voor elke } j \in J \quad (9)$$

zodat de stromen T_{1j} van zone 1 naar iedere j veranderen met eenzelfde factor als E_1 , indien we aannemen dat a_1 niet verandert.

Door de wijziging van E_1 ondergaan nu volgens (8) alle b_j 's een wijziging, welke evenwel gering is, doordat in de noemer de term $a_1 E_1 f(C_{1j})$ slechts één element vormt van een aantal N_1 . Vervolgens veroorzaken volgens (7) de geringe veranderingen van de b_j 's een nóg geringere wijziging van alle a_i 's. Het blijkt dus niet geheel correct, a_1 onveranderlijk te veronderstellen, maar aangezien de verandering zeer gering is, nemen we toch aan, dat alle stroompjes T_{1j} bij benadering met dezelfde factor als E_1 veranderen.

De functie van de b_j 's is kennelijk, alle stromen T_{ij} iets te wijzigen als compensatie voor de verandering van $E_1 = \sum_{j \in J} T_{1j}$ in zone 1. De b_j 's kunnen dus worden gezien als competitiefactoren, die de stromen T_{ij} beïnvloeden, afhankelijk van de 'aantrekkingskrachten' van de zones i . Het is nu ook begrijpelijk dat b_j het omgekeerde is van de gesommeerde 'nabijheid' (eng.: accessibility) van arbeidsplaatsen in alle zones i , ten opzichte van een woonzone j .

Omgekeerd kan analoog hieraan worden geconcludeerd, dat de factoren a_i competitietermen zijn, waarmee T_{ij} beïnvloed wordt door de 'aantrekkingskrachten' van woonzones j . De a_i 's zijn de inversen van de gesommeerde 'nabijheden' van beroepsbevolking in alle zones j ten opzichte van een werkzone i .

Blijkbaar is dus het effect van één veranderende E_i of R_j met model (6) nog te overzien. Indien echter veel of alle E_i 's veranderen, veranderen volgens (8) ook alle b_j 's en daarmee volgens (7) alle a_i 's aanzienlijk. Omgekeerd, als alle R_j 's veranderen, ondergaan alle a_i 's en daarmee alle b_j 's aanzienlijke wijzigingen. De waarden a_i en b_j kunnen in dat geval alleen door calibratie (zie par. 2.3) uit de nieuwe situatie worden bepaald in samenhang met nieuwe functiewaarden van $f(C_{ij})$.

Dit betekent, dat met (6) het gevolg van een wijziging van één enkele E_i voor de R_j 's in het gebied op zeer korte termijn (bijv. 1 jaar) redelijk betrouwbaar kan worden voorspeld, doordat dan alle b_j -waarden onveranderd mogen worden gesteld. Als voorbeeld valt te noemen de vestiging van een groot bedrijf in een zone i , (E_i wordt E_i') bij gegeven totale beroepsbevolking $\sum_{j \in J} R_j$, waardoor de stromen T_{ij} zich herverdelen, en het aantal vacatures in alle werkzones van het gebied iets toeneemt (c.q. de werkeloosheid afneemt).

Evenzo kan met (6) het effect van één zich wijzigende R_j op de diverse R_j 's worden benaderd door alle a_i 's en alle b_j 's onveranderlijk te stellen. Bijvoorbeeld kan worden genoemd de openstelling van een nieuw recreatiegebied in zone j (recreatiecapaciteit R_j wordt R_j') bij gegeven aantallen recreanten E_i per herkomstzone i , zodat $\sum_{i \in I} E_i$ gelijk blijft, de stromen T_{ij} zich wijzigen, en de overcapaciteit in alle zones j van het gebied iets toeneemt, (c. q. de recreatiedruk er iets afneemt).

Wanneer echter een prognose een langere periode (bijv. 10 jaar) beslaat, zullen in het algemeen de overige factoren eveneens veranderingen ondergaan (werkgelegenheid verandert in vele werkzones, beroepsbevolking verandert in alle woonzones, het aantal recreanten verandert in elke herkomstzone, enz.). De veranderingen in alle factoren a_i en b_j kunnen daartoe op geen andere wijze worden geschat dan door extrapolaties op grond van de ontwikkeling ervan in de tijd. Indien geen tijdreeksen van a_i en b_j kunnen worden berekend, kan met (6) geen betrouwbare prognose op langere termijn worden verricht, maar is het model alleen interessant ter beschrijving van een bestaande interactie in een gebied en van het effect van een incidentele verandering in één zone daarvan.

Naar verkiezing kan (6) dus worden gebruikt om R_j te voorspellen op grond van wijzigingen in een E_i , waartoe de betreffende $O_i (= a_i E_i)$ met eenzelfde factor wijzigt als de betreffende E_i , tevens kan men E_i voorspellen door een $A_j (= b_j R_j)$ te wijzigen overeenkomstig de betreffende R_j . In het eerste geval wordt elke b_j bekend verondersteld, in het tweede geval elke a_i .

Welke van beide procedures nu voor woon-werkverkeer wordt gevolgd hangt af van de werkelijke interactie tussen arbeidsplaatsen en woonplaatsen; gedeeltelijk kiezen werknemers hun woonplaatsen bij gegeven situering van hun arbeidsplaatsen, een ander deel zal een arbeidsplaats kiezen afhankelijk van de gegeven situering van de woonplaats.

Empirisch is evenwel gebleken dat de eerstgenoemde situatie zich veel frequenter voordoet dan de tweede (o. a. ALBRECHT, 1972). De hiërarchie van het grondgebruik in de benadering van LOWRY (1964) alsmede die in de hiervan afgeleide modellen

(GOLDNER, 1971) berust zelfs op de aanname dat de spreiding van woongebieden afhankelijk is van de spreiding van werkgelegenheid, onder uitsluiting van het omgekeerde. Ook in deze nota wordt een model van de gedaante (3) onderworpen aan (2) en (4) ten behoeve van analyse van de factoren $A_j = b_j R_j$.

De functie $f(C_{ij})$ kan worden beschouwd als een instrument waardoor (3) berust op een niet-continue kansverdeling van ritlengten. Behalve voor zeer lage, in regionale studies minder relevante waarden van C_{ij} ($C_{ij} < 1$ à 2 km) blijkt empirisch namelijk steeds, dat de relatieve ritlengtefrequentie negatief-exponentieel afneemt, naarmate de ritlengte groter is (fig. 1).

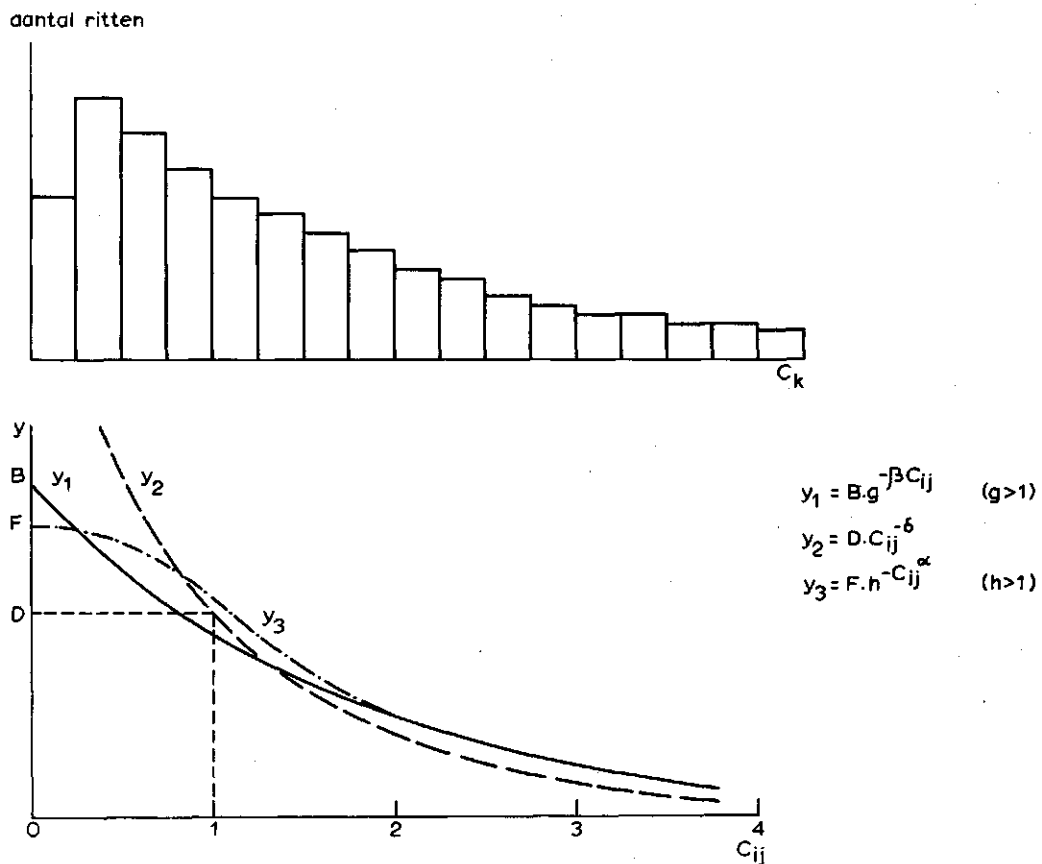


Fig. 1. Theoretische frequentieverdeling van ritlengten en enkele eenvoudige functies ter beschrijving daarvan

Indien de ritlengte als stochastische variabele \underline{C} wordt gedefinieerd voor $0 \leq \underline{C} < \infty$ voor hele waarden van \underline{C} ($= 0, 1, 2, \dots, \infty$, bijvoorbeeld klassebreedten van 1 km) en in een gebied komt

een verzameling K van verschillende ritlengten voor, dan kan de kans, dat \underline{C} in een ritlengteklasse C_k valt, bijvoorbeeld worden benaderd door

$$P(\underline{C} \in C_k) = \frac{e^{-\beta C_k}}{\sum_{k \in K} e^{-\beta C_k}} \quad (10)$$

Nemen we nu in (3) dezelfde kansverdeling aan voor de distributie van T_{ij} als functie van C_{ij} , dan kunnen we invoeren

$$f(C_{ij}) = e^{-\beta C_{ij}} \quad (11)$$

Daardoor zou, indien alle A_j gelijk waren, voor elke i de verdeling van E_i in stroompjes T_{ij} verlopen volgens

$$\frac{T_{ij}}{E_i} = P(\underline{C} \in C_{ij}) = \frac{e^{-\beta C_{ij}}}{\sum_{j \in J} e^{-\beta C_{ij}}} \quad \text{voor alle } j \in J \quad (12)$$

Aangezien de distributie echter mede afhankelijk is gesteld van een factor A_j per woonzone, wordt het kansmodel uitgebreid tot

$$\frac{T_{ij}}{\sum_{j \in J} T_{ij}} = \frac{T_{ij}}{E_i} = P(\underline{T}_i \in T_{ij}) = \frac{A_j e^{-\beta C_{ij}}}{\sum_{j \in J} A_j e^{-\beta C_{ij}}} \quad \text{voor alle } j \in J \quad (13)$$

De kans dat, van alle in het gebied gemaakte ritten, een rit van i naar j voert hangt bovendien per werkzone van O_i af:

$$P(\underline{T} \in T_{ij}) = \frac{T_{ij}}{\sum_{i \in I} \sum_{j \in J} T_{ij}} = \frac{T_{ij}}{\sum_{i \in I} E_i} = \frac{O_i A_j e^{-\beta C_{ij}}}{\sum_{i \in I} \sum_{j \in J} O_i A_j e^{-\beta C_{ij}}}$$

voor alle $i \in I$ en alle $j \in J$ (14)

Door invoering van (11) krijgt (3) nu dus de vorm:

$$T_{ij} = \frac{E_i A_j e^{-\beta C_{ij}}}{\sum_{j \in J} A_j e^{-\beta C_{ij}}} = O_i A_j e^{-\beta C_{ij}} \quad \text{voor alle } i \in I \quad (15)$$

en alle $j \in J$

Voor C_{ij} kan worden genomen de afstand d_{ij} tussen i en j , hemelsbreed of over de weg, de reistijd t_{ij} , of een combinatie van beide, bijvoorbeeld van de vorm $C_{ij} = p d_{ij} + q t_{ij}$, waarin de wegingsfactoren p en q de waarderingsverhouding aangeven van resp. afstand (= voertuig- of transportkosten) en reistijd (= tijd-kosten), en als zodanig normatief moeten worden vastgesteld. Toepassing van zo'n afstand-tijd functie is alleen mogelijk als waarnemingen voor t_{ij} beschikbaar zijn. Daar t_{ij} afhankelijk is van congestie, dus van de specifieke uurintensiteiten op ieder wegvak, zijn metingen aan t_{ij} vrij tijdrovend en wordt in de praktijk doorgaans volstaan met invoeren van de afstand als reisbezwaar.

2.3. Calibratie van het model

Calibratie van een model houdt in, dat de diverse parameters in het model zó worden gekozen dat daarmee een bepaald proces of systeem wordt afgebeeld op een wijze, die aan vooraf gestelde eisen van nauwkeurigheid voldoet. Daartoe is een set van bij elkaar behorende gegevens nodig, waargenomen aan alle in het model voorkomende variabelen. Uit deze waarnemingen worden de parameters dan berekend. Met behulp van statistische grootheden kan vervolgens worden bepaald in welke mate de modelwaarden aansluiten bij de gemeten waarden, en hieruit zijn conclusies te trekken over de statistische betrouwbaarheid van de modeluitkomsten in het algemeen.

In het model (15) zijn de te bepalen parameters: β , de A_j 's en de O_i 's. Hiertoe zijn de waarnemingen benodigd betreffende R_j , E_i , C_{ij} en T_{ij} . BATTY (1970) heeft aangetoond, dat het ten behoeve van een betrouwbare aanpassing van de parameters noodzakelijk is, te beschikken over een volledige matrix met waarnemingen n_{ij} voor T_{ij} , dus de ritten van alle $i \in I$ naar alle $j \in J$, waaruit dan tevens

als randtotalen per zone de waarden van de werkgelegenheid

$E_i = \sum_{j \in J} n_{ij}$ (voor elke i) en de beroepsbevolking $R_j = \sum_{i \in I} n_{ij}$ (voor elke j), voorzover in de waarnemingenmatrix, volgen.

Men kan niet volstaan met alleen de spreiding van arbeidsplaatsen en beroepsbevolking (dus C_{ij} , E_i en R_j) te meten, omdat de gebruikelijke statistische grootheden die een maat zijn voor de aanpassing, niet voldoende gevoelig voor parametervariatie blijken om vast te stellen of de eindoplossing in het aanpassingsproces bereikt is. Indien alleen de randtotalen E_i en R_j bekend zijn, dan kunnen uiteenlopende waarden van n_{ij} daarin zó tegen elkaar zijn weggeïmmiddeld, dat weliswaar een groot aantal oplossingen voor O_i , A_j en β bij deze gegevens past, maar deze sluiten nog niet alle aan bij de individuele waarden n_{ij} .

Bij de calibratie van model (15) bedraagt het aantal op te lossen onbekenden $N_I + N_J + 1$. De ritmatrix (T_{ij}) levert $N_I \cdot N_J$ vergelijkingen (15), waaraan deze onbekenden moeten worden aangepast. Bij de oplossing worden vergelijkingen gebruikt in de partiële afgeleiden van T_{ij} naar de diverse parameters. Daar (15) een niet-lineair verband tussen de parameters geeft bevat iedere partiële afgeleide ook de overige, nog onbekende, parameters. Er ontstaat dus een simultaan stelsel niet-lineaire vergelijkingen dat niet rechtstreeks oplosbaar is. De parameters moeten dan ook worden benaderd met beginschattingen, welke door herberekening uit elkaar itererend worden verbeterd totdat de verlangde statistische aansluiting is bereikt.

Stel nu dat we de beschikking hebben over waarnemingen van woon-werkritten in een studiegebied, waarvan n_{ij} ritten plaatsvinden tussen i en j , dus $\sum_{i \in I} \sum_{j \in J} n_{ij} = n$.

Volgens HYMAN (1969) moeten nu de $N_I + N_J + 1$ schattingen voor resp. O_i , A_j en β zo gekozen worden, dat

1. de rijtotalen van de matrix (T_{ij}) gelijk zijn aan die van de waarnemingenmatrix (n_{ij}) :

$$\sum_{j \in J} T_{ij} = \sum_{j \in J} n_{ij} = \sum_{j \in J} O_i A_j e^{-\beta C_{ij}} \quad \text{voor alle } i \in I$$

hieruit volgt:

$$O_i = \sum_{j \in J} n_{ij} \left[\sum_{j \in J} A_j e^{-\beta C_{ij}} \right]^{-1} \quad \text{voor alle } i \in I \quad (16)$$

2. de kolomtotalen van de matrix (T_{ij}) gelijk zijn aan die van de waarnemingenmatrix (n_{ij}) :

$$\sum_{i \in I} T_{ij} = \sum_{i \in I} n_{ij} = \sum_{i \in I} O_i A_j e^{-\beta C_{ij}} \quad \text{voor alle } j \in J$$

hieruit volgt:

$$A_j = \sum_{i \in I} n_{ij} \left[\sum_{i \in I} O_i e^{-\beta C_{ij}} \right]^{-1} \quad \text{voor alle } j \in J \quad (17)$$

3. het gemiddelde reisbezwaar (in afstand, tijd of kosten) in het model gelijk is aan dat in de waarnemingen:

$$\sum_{i \in I} \sum_{j \in J} C_{ij} T_{ij} = \sum_{i \in I} \sum_{j \in J} C_{ij} n_{ij} = \sum_{i \in I} \sum_{j \in J} C_{ij} O_i A_j e^{-\beta C_{ij}} \quad (18)$$

(16), (17) en (18) leveren aldus $N_I + N_J + 1$ vergelijkingen waaruit evenveel onbekenden kunnen worden opgelost.

Langs een theoretisch fraaiere weg zijn deze vergelijkingen ook af te leiden met behulp van de schattingstheorie (EVANS, 1971): Als we de distributie van n waargenomen ritten met model (15) willen beschrijven, dan kunnen we volgens (14) de kans dat een rit van i naar j voert, formuleren als

$$P(\underline{T} \in T_{ij}) = \frac{1}{n} O_i A_j e^{-\beta C_{ij}} \quad \text{voor alle } i \in I \text{ en alle } j \in J \quad (19)$$

Voor de gezamenlijke kans $P \left\{ \bigwedge_{i,j} T_{ij} = n_{ij} \right\}$ dat de modelwaarden T_{ij} overeenkomen met de waarnemingen n_{ij} voor alle $i \in I$ en alle $j \in J$ tegelijk, geldt dat deze afhangt van alle individuele kansen $P(\underline{T} \in T_{ij})$ en als zodanig een multinomiale verdeling bezit. Indien we uit m verschillende steekproeven over m verschillende waarnemingsmatrices (n_{ij}) beschikken, dan fungeert het produkt van deze m gezamenlijke kansen als aannemelijkheidsfunctie ('likelihood').

Aangezien we uit de Volkstelling 1960 voor het woon-werkverkeer slechts één matrix (n_{ij}) kunnen samenstellen ($m=1$) is de aanmelijheidsfunctie in dit geval:

$$\begin{aligned}
 L &= \prod_m P \left\{ \bigwedge_{i,j} T_{ij} = n_{ij} \right\} = \prod \binom{n}{n_{11} \ n_{12} \ \dots \ n_{N_I N_J}} \prod_{i \in I} \prod_{j \in J} P(T \in T_{ij})^{n_{ij}} \\
 &= \binom{n}{n_{11} \ n_{12} \ \dots \ n_{N_I N_J}} \left(\frac{1}{n}\right)^n \prod_{i \in I} \prod_{j \in J} (O_i A_j e^{-\beta C_{ij}})^{n_{ij}} \\
 &= \text{constante} \cdot \prod_{i \in I} \prod_{j \in J} (O_i A_j e^{-\beta C_{ij}})^{n_{ij}}
 \end{aligned}$$

voor alle $i \in I$ en alle $j \in J$ (20)

De parameters moeten zó worden geschat, dat L maximaal is ('maximum likelihood estimation'). Aangezien $\log L$ een continue stijgende functie van L is, mogen we daartoe ook $\log L$ maximaliseren:

$$\log L = \text{constante} \cdot \sum_{i \in I} \sum_{j \in J} n_{ij} (\log O_i + \log A_j - \beta C_{ij}) \quad (21)$$

Hierbij moet worden voldaan aan de voorwaarde

$$\sum_{i \in I} \sum_{j \in J} T_{ij} = \sum_{i \in I} \sum_{j \in J} O_i A_j e^{-\beta C_{ij}} = n \quad (22)$$

Met behulp van de multiplicatorenmethode van Lagrange voeren we de voorwaarde (22) in (21) in en maximaliseren:

$$F = \sum_{i \in I} \sum_{j \in J} n_{ij} (\log O_i + \log A_j - \beta C_{ij}) - \lambda \left(\sum_{i \in I} \sum_{j \in J} O_i A_j e^{-\beta C_{ij}} - n \right) \quad (23)$$

in $N_I + N_J + 1$ richtingen, t.w. O_i voor alle $i \in I$, A_j voor alle $j \in J$ en β .

$$\frac{\delta F}{\delta O_i} = 0 \text{ geeft } \lambda \sum_{j \in J} O_i A_j e^{-\beta C_{ij}} = \sum_{j \in J} n_{ij} \text{ voor alle } i \in I \quad (24)$$

Sommering van (24) voor $i \in I$ geeft, na substitutie van (22):

$$\lambda \cdot n = \sum_{i \in I} \sum_{j \in J} n_{ij} \quad , \quad \text{zodat} \quad \lambda = 1$$

Hiermee wordt (24):

$$O_i \sum_{j \in J} A_j e^{-\beta C_{ij}} = \sum_{j \in J} n_{ij} \quad \text{voor alle } i \in I \quad (25)$$

Analoog hieraan volgt uit $\frac{\delta F}{\delta A_j} = 0$:

$$A_j \sum_{i \in I} O_i e^{-\beta C_{ij}} = \sum_{i \in I} n_{ij} \quad \text{voor alle } j \in J \quad (26)$$

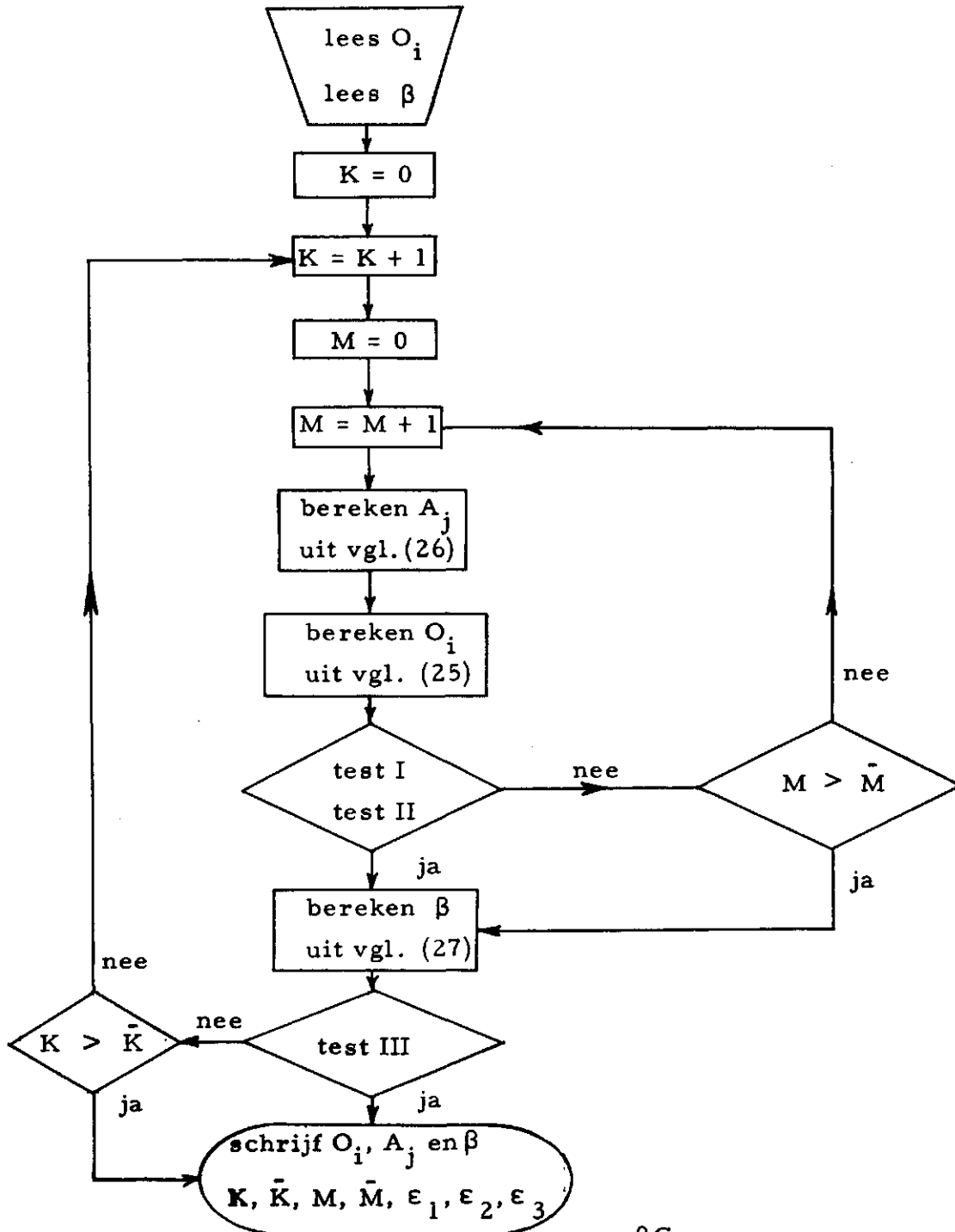
en uit $\frac{\delta F}{\delta \beta} = 0$ volgt:

$$\sum_{i \in I} \sum_{j \in J} C_{ij} O_i A_j e^{-\beta C_{ij}} = \sum_{i \in I} \sum_{j \in J} C_{ij} n_{ij} \quad (27)$$

De vergelijkingen (25), (26) en (27) zijn dezelfde als resp. (16), (17) en (18).

Zoals hiervoor beschreven, worden met behulp van deze vergelijkingen de parameterschattingen in een iteratief rekenproces gecorrigeerd, totdat ze gezamenlijk zó aansluiten bij de gegevens, dat aan gestelde eisen is voldaan. Een en ander is schematisch weergegeven in fig. 2.

Het aantal correctielussen dat men de computer achter elkaar laat uitvoeren zonder tussentijds printen van de waarden O_i , A_j en β , kan op grond van ervaring en inzicht worden gekozen met behulp van de tellers K en M , waarvan de drempelwaarden \bar{K} en \bar{M} , een beveiliging vormen om te voorkomen, dat de berekening eindeloos lang wordt voortgezet, indien aan de tests I, II of III niet wordt voldaan, doordat de aanpassing divergeert of te langzaam convergeert. De itererende berekening wordt stopgezet wanneer aan de tests I, II en III is voldaan of als $K > \bar{K}$.



$$\text{test I : } \sum_{i \in I} \left(\sum_{j \in J} n_{ij} - \sum_{j \in J} O_i A_j e^{-\beta C_{ij}} \right)^2 < \epsilon_1$$

$$\text{test II : } \sum_{j \in J} \left(\sum_{i \in I} n_{ij} - \sum_{i \in I} O_i A_j e^{-\beta C_{ij}} \right)^2 < \epsilon_2$$

$$\text{test III: } \beta_K - \beta_{K-1} < \epsilon_3$$

Fig. 2. Blokschema van de calibratie van model (15) met een computer

2.4. Analyse van parameterwaarden

De aldus aangepaste parameterwaarden gelden steeds voor de specifieke situatie waarin de waarnemingen aan de variabelen zijn verricht. Teneinde te komen tot een algemene geldigheid, onverschillig plaats en tijd, zal het model moeten worden verfijnd door aan de in diverse situaties berekende parameterwaarden een interpretatie te geven, met andere woorden door te onderzoeken of de parameters een functioneel verband vertonen met variabelen, welke in iedere specifieke situatie meetbaar zijn, danwel met een zekere betrouwbaarheid kunnen worden voorspeld.

Zoals in par. 2.3 geformuleerd, bevat het model (15) voor iedere zone $j \in J$ een factor A_j . Deze kan wellicht worden gezien als een functie van variabelen (u, v, w, \dots) van iedere specifieke zone:

$$A_j = f(u_j, v_j, w_j, \dots) \quad (28)$$

Met name het verloop van A_j in de tijd kan mogelijk op deze wijze worden beschreven.

Een dergelijke beschouwing geldt voor de parameter β in de afstandsfunctie, zij het dat deze in (15) een algemene geldigheid bezit voor het gehele studiegebied, en dan ook slechts afhankelijk kan worden gesteld van over het studiegebied geaggregeerde variabelen van de beide groepen zones, I en J:

$$\beta = f(x_I, y_I, z_I, \dots, x_J, y_J, z_J, \dots) \quad (29)$$

Essentieel bij deze modelverfijningen blijft uiteraard wel, dat men weet, welke betekenis aan de berekende parameters moet worden toegekend, met andere woorden dat ze zó zijn uitgerekend dat een gekozen interpretatie ervan is toegestaan.

In hoofdstuk 3 zal worden gepoogd aan deze interpretatie gestalte te geven, aan de hand van een chronologische beschrijving van de tot nu toe uitgevoerde berekeningen.

3. ENKELE TOEPASSINGEN VAN CALIBRATIE EN PARAMETER-ANALYSE

3.1. Werkforensen op de gemeente Utrecht in 1960

Dit eerste voorbeeld heeft betrekking op de relaties van werkforensen tussen één werkzone ($N_I = 1$), t.w. de gemeente Utrecht, en de woongemeenten binnen een cirkel met straal ± 30 km hemelsbreed ($= \pm 35$ km over de weg) rond de stad Utrecht ($N_J = 98$). Omdat hier sprake is van slechts één herkomstzone neemt de zwaarte-krachtsformule (15) een triviale vorm aan; in dit geval is namelijk

$$\frac{E_i}{\sum_{j \in J} A_j e^{-\beta C_{ij}}} = O_i = \text{constant} \quad \text{voor elke relatie } i - j \quad (30)$$

Als attractiefactor voor iedere woongemeente kunnen we nu $O_i A_j$ vervangen door de omvang P_j van de beroepsbevolking (= alle inwoners met een beroep):

$$T_j = P_j e^{-\beta C_j} \quad \text{voor elke } j \in J \quad (31)$$

Met (31) wordt aldus beweerd, dat T_j/P_j , dat wil zeggen de fractie van de beroepsbevolking uit woonzone j , welke als forens in Utrecht werkt, negatief-exponentieel afneemt als functie van het afstandsbezwaar C_j . Voor het theoretische geval dat $C_j = 0$ betekent dit, dat de ter plaatse wonende beroepsbevolking in haar geheel in Utrecht werkzaam is (namelijk $T_j/P_j = 1$). In fig. 3 is de waargenomen n_j/P_j grafisch uitgezet tegen C_j , waarbij n_j en P_j zijn overgenomen uit volkstellingsgegevens voor 1960 (CBS, 1962), en C_j de kortste weg-afstand in km is, ontleend aan wegenkaarten (ANWB, 1963).

De calibratie van (31) is in dit geval vereenvoudigd tot een eenvoudige regressieberekening van de parameter β . Het resultaat hiervan toont fig. 3: $\beta = 0,1672$ ($r_{n,T}^2 = 0,96$).

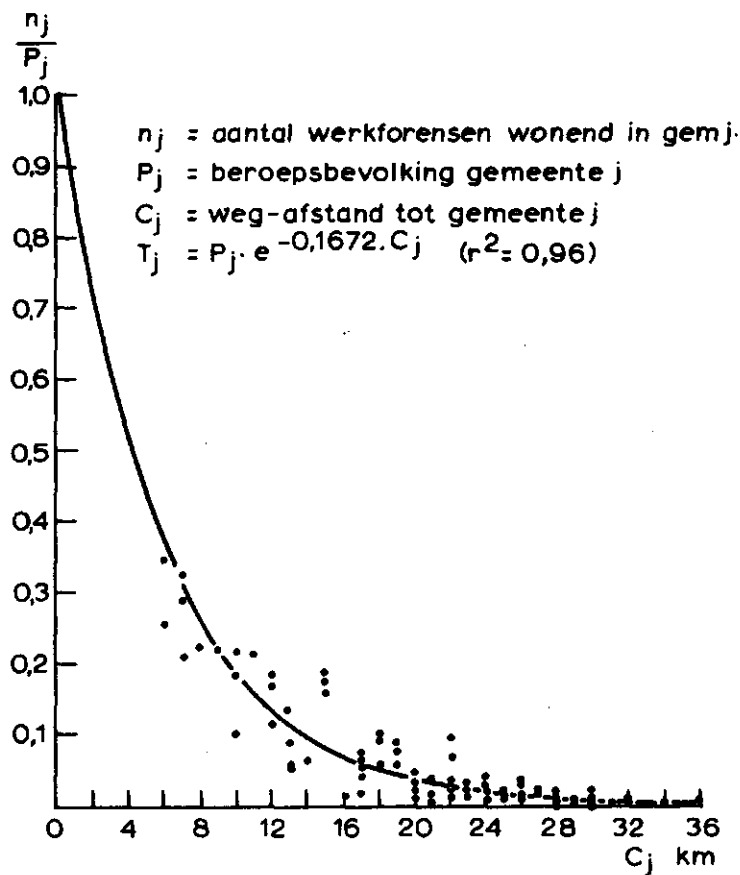


Fig. 3. Gravitiemodel werkforensen gemeente Utrecht, 1960.
 Gemeten fracties in Utrecht werkzame beroepsbevolking n_j/P_j
 en de daaraan aangepaste functie T_j/P_j

Voortbouwend op de benadering van MIDDELKOOP (1970) zou men vervolgens het quotiënt van n_j en T_j voor elke j kunnen beschouwen als een 'relatieve aantrekkelijkheidscoëfficiënt', welke aangeeft, in welke mate een gemeente relatief meer in Utrecht werkzaam onder haar beroepsbevolking telt dan een andere. Fig. 4 is evenwel voldoende illustratief om in te zien, dat de term 'aantrekkelijkheid' in dit geval uitermate misleidend is, de verhouding tussen gemeten en berekende aantallen vertoont een duidelijke geografische concentratie van hoge resp. lage waarden, hetgeen te verwachten viel. Het laat zich namelijk verstaan, dat per woongemeente het relatieve aantal in Utrecht werkzame personen wordt

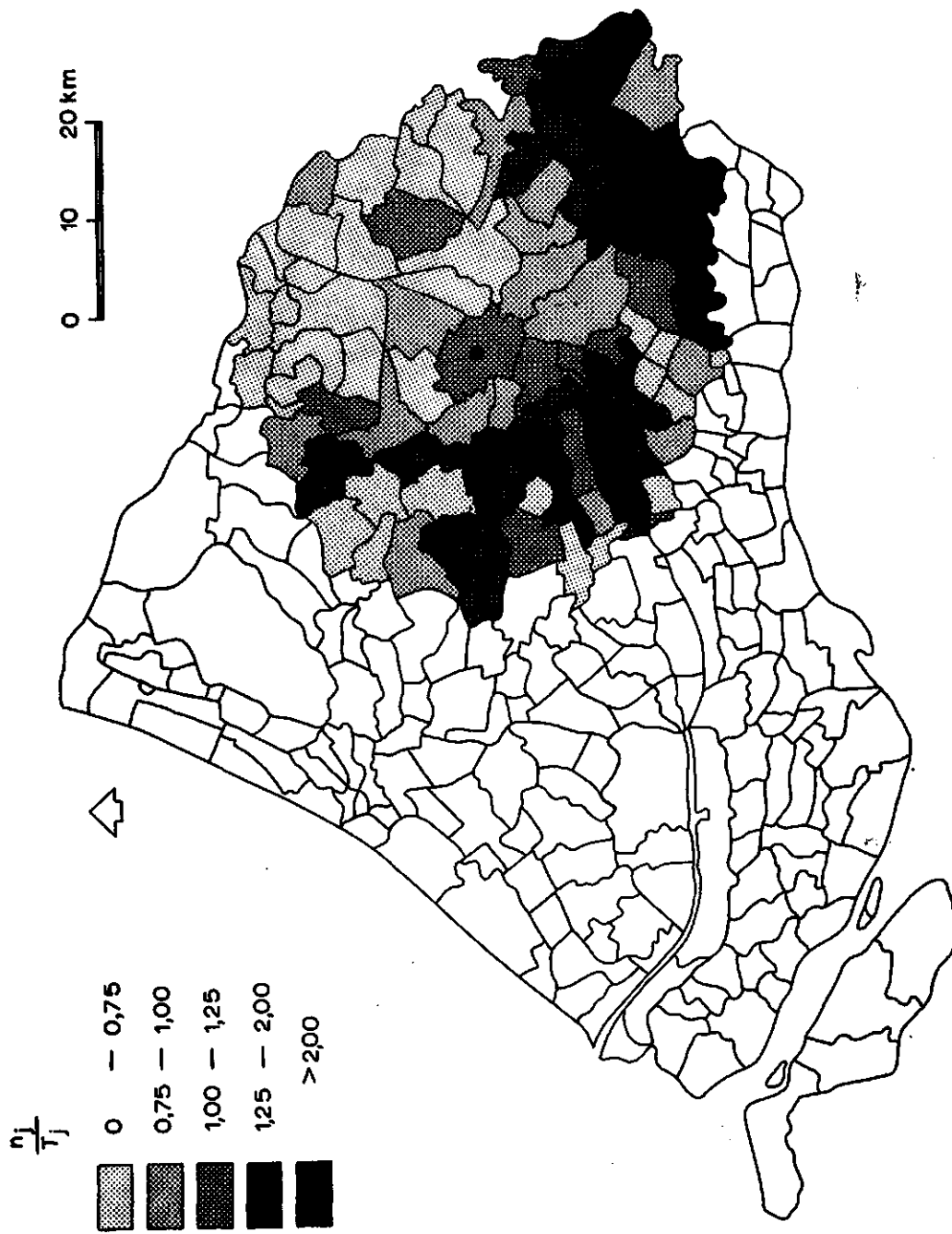


Fig. 4. Gravitiemodel werkforensen gemeente Utrecht, 1960. Verhouding van gemeten en berekende fracties in Utrecht werkzame beroepsbevolking per gemeente

beïnvloed door de spreiding van (= omvang van en afstand tot) alternatieve werkgelegenheid ('job opportunity') in en rond de betreffende gemeente. Het kaartbeeld spreekt in dit opzicht voor zichzelf: lage waarden in en rond Hilversum, Amersfoort en Zeist, hoge waarden in de rivierstreek ten zuiden en het weidegebied ten westen van Utrecht. Terzijde zij opgemerkt, dat bovendien een rol kan spelen de betere bereikbaarheid van Utrecht vanuit zuidelijke en westelijke richting via (spoor-)wegen van hogere kwaliteit, hetgeen tot uitdrukking zou komen indien voor C_j reistijden in plaats van weg-km konden worden ingevoerd.

Het aldus aangepast model is alleen interessant ter beschrijving van de spreiding van een aantal in Utrecht werkzame personen over de omliggende gemeenten, maar het geeft geen enkele informatie omtrent de spreiding van woongebieden, of een daarbij bestaande relatieve voorkeur voor bepaalde gemeenten. De steekproefgrootte, t.w. $\sum_{j \in J} n_j / \sum_{j \in J} P_j$ voor de beschouwde gemeenten bedraagt 0,0525. Aan voorspellingen van de bevolkingsontwikkeling in bijvoorbeeld de Lopikerwaardgemeenten (MIDDELKOOP, 1970) op grond van dit soort parameters kan dan ook slechts een zeer betrekkelijke waarde worden toegekend.

Dit voorbeeld maakt duidelijk, dat het voor onderzoek naar het woonplaatskeuzegedrag noodzakelijk is, de arbeidsplaatsen van meerdere zones tegelijk in de beschouwing te betrekken, dat wil zeggen niet alleen $N_J > 1$, maar ook $N_I > 1$. Dat betekent ook, dat het distributiemodel (15) werkelijk inhoud krijgt, doordat de term

$$O_i = E_i \left[\sum_{j \in J} A_j e^{-\beta C_{ij}} \right]^{-1} \quad \text{voor elke } i \in I \text{ een andere waarde aanneemt.}$$

3.2. Werkforensen op 14 centra in de Randstad in 1960

Uit de in paragraaf 3.1 beschreven berekening werd duidelijk dat de woon-werkrelaties van diverse gemeenten met Utrecht worden beïnvloed door de aanwezigheid van alternatieve werkgelegenheid in de omgeving. De berekening gaf aldus geen juist beeld van de spreiding van woongebieden in relatie met de spreiding van arbeids-

plaatsen, doordat het onderzochte deel van de totale beroepsbevolking zeer gering was (ca. 5%) en bovendien een specifieke groep vormde (de in Utrecht werkzame personen).

Om dit bezwaar te ondervangen is het nodig een groter gedeelte van de beroepsbevolking, waarvan bovendien de arbeidsplaatsen over meerdere centra gespreid zijn, te beschouwen. Daartoe is een studiegebied gekozen, bestaande uit de provincie Noordholland bezuiden het Noordzeekanaal, de provincies Zuidholland en Utrecht alsmede de Gelderse gemeenten benoorden de Waal en ten westen van Kesteren. Het gebied omvat 245 gemeenten, welke als woonzones zijn beschouwd ($N_J = 245$), terwijl in en op de rand van het gebied 14 duidelijke centra van werkgelegenheid zijn gekozen als werkzones ($N_I = 14$), welke bestaan uit één of meer gemeenten, te weten:

1. Amsterdam
2. Zaandam, Koog a.d. Zaan, Wormerveer, Krommenie
3. Beverwijk, Velsen
4. Haarlem, Heemstede, Bloemendaal
5. Haarlemmermeer
6. Leiden
7. Den Haag, Rijswijk, Voorburg
8. Delft
9. Rotterdam, Schiedam, Vlaardingen
10. Dordrecht, Zwijndrecht
11. Utrecht
12. Zeist
13. Amersfoort
14. Hilversum

Aan de woon-werkrelaties in 1960 van deze 14 zones met de 245 genoemde woonzones is model (15) aangepast op de in paragraaf 2.3 beschreven wijze (zie ook MICHELS, 1972). Bron voor de aantallen relaties, n_{ij} , $E_i = \sum_{j \in J} n_{ij}$ en $R_j = \sum_{i \in I} n_{ij}$, was weer de Volkstelling 1960 (CBS, 1962), terwijl voor C_{ij} de hemelsbrede afstanden tussen i en j in km werden ingevoerd, gemeten op wegenkaarten (ANWB, 1963). De steekproefgrootte ten opzichte van de totale beroepsbevolking, $\sum_{j \in J} R_j / \sum_{j \in J} P_j$ bedraagt 0,7418. De berekening

resulteerde in een parameter $\beta = 0,229966$, en parameters A_j zoals ze in fig. 5 in kaart zijn gebracht. De absolute waarden van A_j hebben op zichzelf weinig betekenis; deze worden mede bepaald door de dimensies waarin de diverse variabelen zijn ingevoerd. Alleen de onderlinge verhouding van de A_j -waarden geeft een indruk van de relatieve attracties, zoals die volgen uit de waarnemingen n_{ij} . Alle waarden A_j mogen dus met eenzelfde willekeurige positieve factor worden vermenigvuldigd, dit brengt in formule (15) geen wijziging in de functiewaarden T_{ij} .

Fig. 6 geeft de geografische spreiding van de waarden b_j , verkregen door de aldus gecalibreerde A_j -waarden te delen door $R_j = \sum_{i \in I} n_{ij}$. Het kaartbeeld illustreert de uit (8) afgeleide interpretatie van b_j als omgekeerde van de gesommeerde nabijheid van werkgelegenheid ('job opportunity'). De waarden van b_j zijn laag in en rond de grote centra van de Randstad, b_j is relatief hoog in de ten opzichte van de werkgelegenheid meer excentrisch gelegen zones van het studiegebied.

In dit verband dient nu echter wel te worden bedacht, dat de onderlinge verhouding van de b_j 's, bij calibratie op een niet-vierkante ritmatrix ($I \neq J$, steekproefgrootte = 0,7418), afhankelijk is van de in het studiegebied gekozen werkcentra i . Zo zijn bijv. de b_j -waarden in het meest zuidoostelijke deel van het studiegebied beduidend hoger dan die rond Dordrecht en Zwijndrecht. Dit wordt gedeeltelijk veroorzaakt door het feit, dat in het eerste gebied geen werkcentrum i is beschouwd, in het tweede wel. Wat bij het Utrechtse voorbeeld (paragraaf 3.1) zeer evident was, speelt ook hier nog steeds een rol: de niet in de berekening betrokken werkgelegenheid veroorzaakt te hoge waarden van b_j in de nabijgelegen zones. De enige mogelijkheid om deze oneerlijke concurrentie tussen de woonzones uit te sluiten is, het model te calibreren op een vierkante matrix (n_{ij}) met waarnemingen aan alle werkzones in het studiegebied ($I = J$, en dus $N_I = N_J$).

Bovendien bestaat geen zekerheid omtrent de invloed van het gebruik van hemelsbreed gemeten afstanden op de waarden b_j . In het algemeen mag worden aangenomen, dat wegafstanden gemiddeld globaal 1,1 à 1,2 x de hemelsbrede afstanden bedragen. Bij toepas-

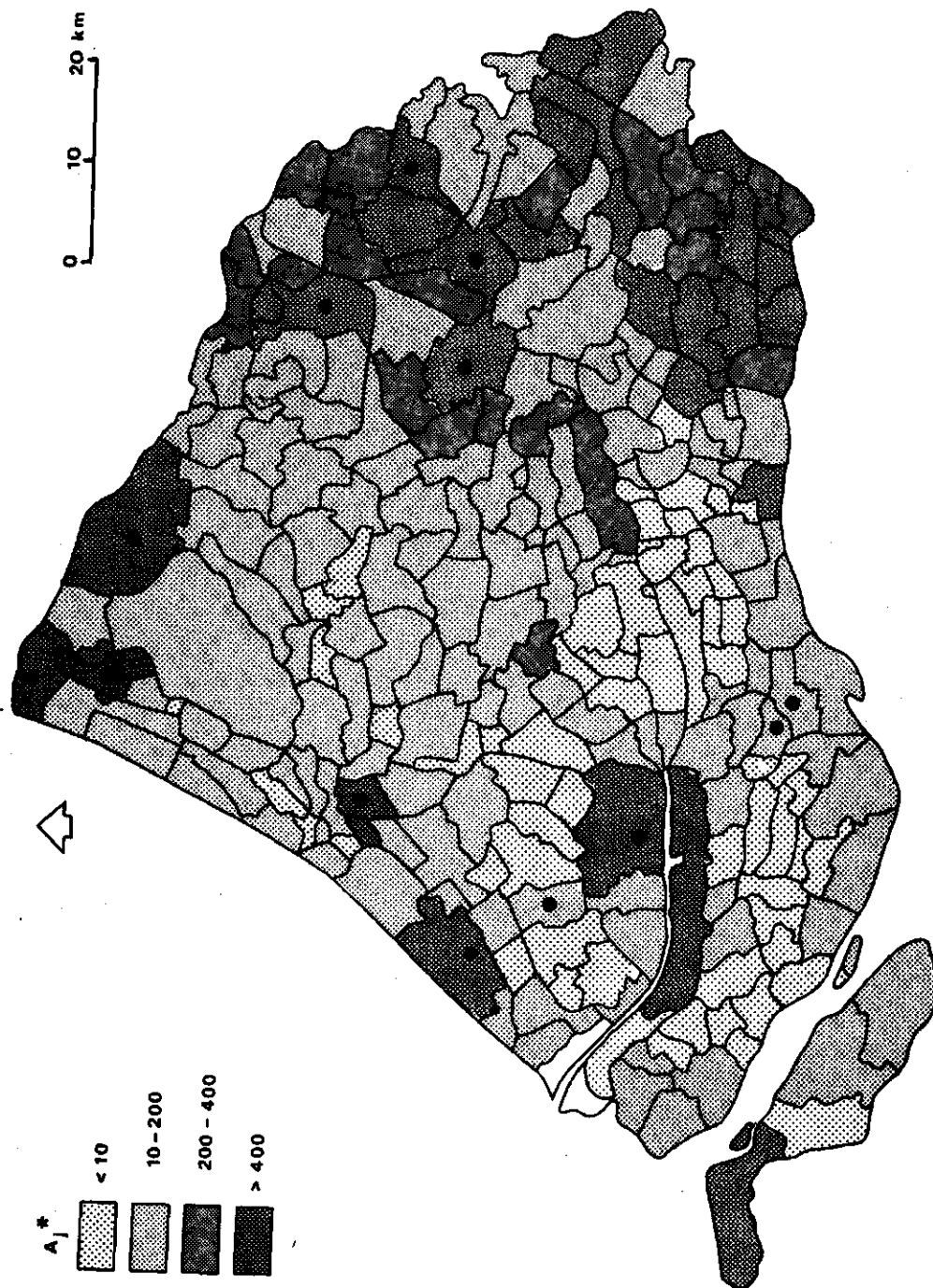


Fig. 5. Attractiefactoren A_j uit het zwaartekrachtmodel (15), aangepast aan de distributie van woonwerkkriten uit 14 grote centra van werkgelegenheid in 1960. (*Voor de werkelijke waarden vermenigvuldige men alle klassegrenzen met 133,6206)

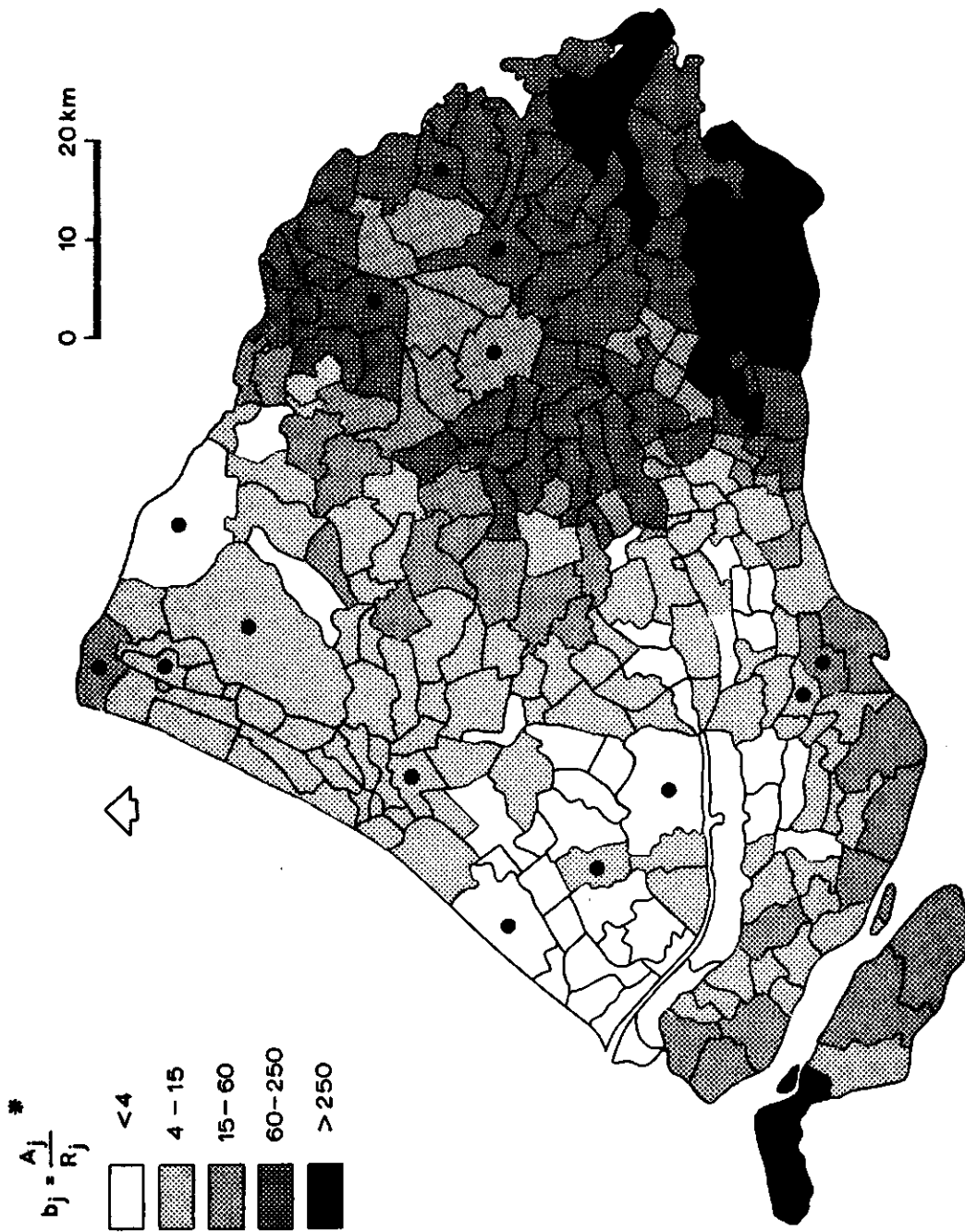


Fig. 6. Competitiefactoren b_j in model (15), aangepast aan de distributie van woon-werkritten op 14 grote centra van werkgelegenheid in 1960. (*Voor de werkelijke waarden vermenigvuldige men alle klassegrenzen met 0,1336206)

sing van hemelsbrede afstanden gaat men dan ook impliciet uit van een voor alle relaties gelijke verhoudingsfactor. Voor de door spoor- en autowegen relatief goed ontsloten gemeenten, bijvoorbeeld wederom in de zuidoosthoek van het gebied (Beesd, Culemborg, Geldermalsen, Waardenburg e. a.) vormt de hemelsbrede afstand tot de werkcentra wellicht een te hoge schatting van C_{ij} ten opzichte van die van andere gemeenten, waardoor de b_j in zulke gemeenten relatief te hoog wordt.

Om dit gevaar zo veel mogelijk te vermijden is het belangrijk, dat het werkelijke reisbezwaar zo goed mogelijk wordt benaderd. In ieder geval verdient het aanbeveling, wegafstanden in te voeren. In paragraaf 3.1 is evenwel reeds aangestipt, dat ook in dat geval nog verschillen in ontsluitingskwaliteit (in de vorm van reistijden) de parameters A_j en b_j kunnen beïnvloeden.

Beide overwegingen zijn mede aanleiding geworden tot de in paragraaf 3.3 beschreven calibratie van (15) op een vierkante waarnemingsmatrix van alle woon-werkrelaties in het studiegebied met invoering van wegafstanden voor C_{ij} .

In verband met het bovenstaande behoeft het geen verwondering te wekken, dat in dit stadium een poging, de A_j -waarden te relateren aan andere specifieke variabelen per woonzone, nog weinig bevredigende resultaten opleverde. Fig. 7 illustreert, dat van de aangepaste A_j -waarden per woongemeente een significante correlatie werd geconstateerd met de landoppervlakte indien de gemeenten werden onderscheiden naar urbanisatiegraad. Met behulp van lineaire regressie is het verband afgeleid tussen de logaritmen van A_j en L_j , hetgeen resulteert in: $A_j = 10^t \cdot L_j^s$ met t , s en correlatiecoëfficiënt r als gegeven in de figuur. Aangezien de urbanisatiegraad de gemeenten in klassen A, B en C globaal onderscheidt naar hun landelijke resp. stedelijke karakter (CBS, 1964), mag een dergelijk verband redelijkerwijze worden verwacht; vergelijking van fig. 5 en 6 maakt duidelijk, dat in alle drie klassen de b_j -waarden een grote spreiding vertonen, en dat de R_j -waarden een overheersende rol spelen in de variantie in A_j . Hiermee zij verklaard dat de A_j 's in de C-gemeenten (relatief hoge R_j 's met grote spreiding) gemiddeld een iets hogerniveau hebben met grotere spreiding, dan

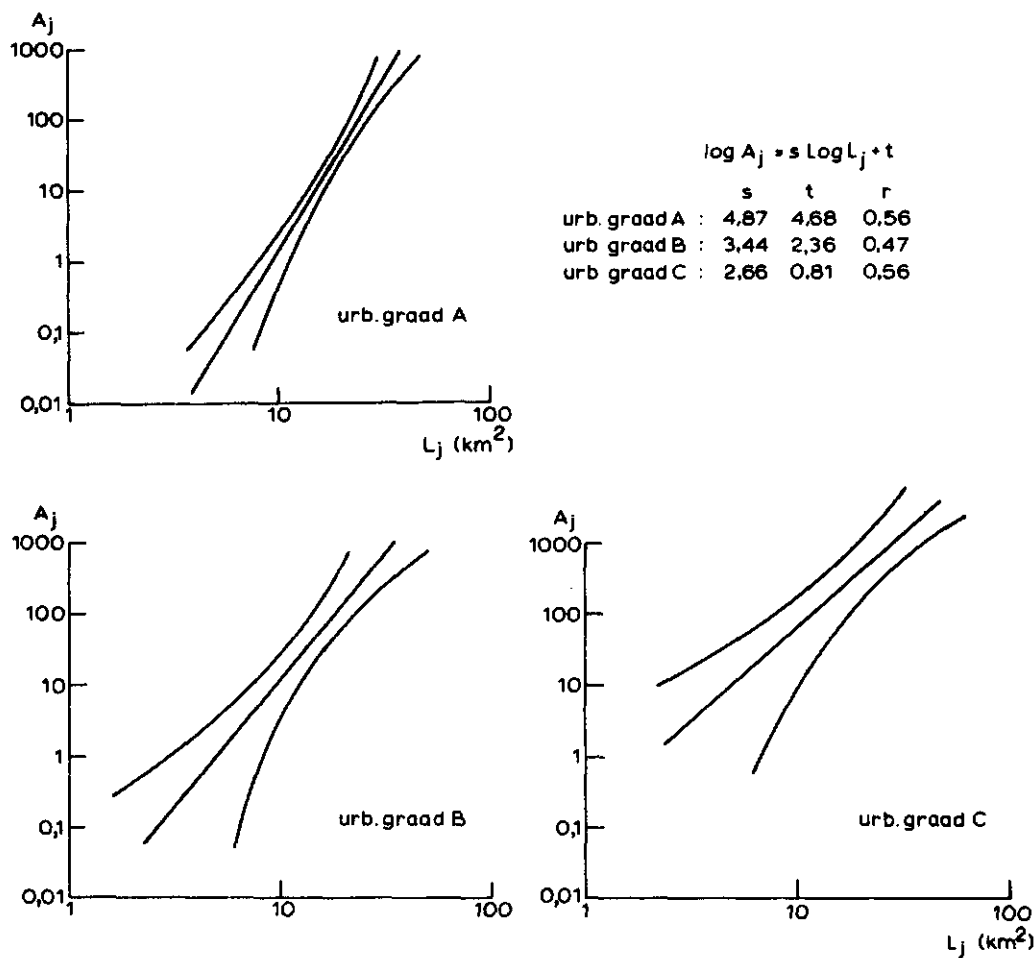


Fig. 7. Gravitiemodel woon-werkverkeer 14 centra, 1960.
 Verband tussen A_j en de landoppervlakte L_j per gemeente
 naar urbanisatiegraad, met 95% betrouwbaarheidsintervallen

de A-gemeenten (relatief lage R_j 's met geringere spreiding). Het is evenwel duidelijk, dat correlaties als deze onvoldoende betrouwbaar zijn om als basis te dienen voor bruikbare prognoses van A_j . Voordat tot correlatieberekeningen kan worden overgegaan, dienen we ons ervan te vergewissen dat de calibratie van de A_j -waarden zó realistisch is uitgevoerd, dat deze een interpretatie van b_j volgens (8) rechtvaardigt.

3.3. Beroepsbevolking van de drie Randstadprovincies in 1960

Op grond van de in paragraaf 3.2 opgedane ervaringen zijn vervolgens de woon-werkbewegingen van de totale beroepsbevolking met een vast werkadres tussen alle onderscheiden zones onderling in een calibratie ingevoerd. Daartoe zijn als studiegebied gekozen de provincies Noordholland, Zuidholland en Utrecht, waarvan de gemeenten zijn samengevoegd tot 84 zones. Als herkomst- of bestemmingszones van de relaties die de grens van het studiegebied overschrijden, zijn de overige acht provincies als onderscheiden zones toegevoegd, zodat de calibratie is uitgevoerd op grond van een vierkante (92 x 92) matrix van waargenomen ritaantallen n_{ij} ($I = J$, $N_I = N_J = 92$).

De waargenomen aantallen ritten n_{ij} tussen alle zones zijn berekend uit de Volkstellingsgegevens per gemeente (CBS, 1962). De weerstanden C_{ij} tussen de zones zijn als wegafstanden in km langs de kortste route ingevoerd, welke zijn ontleend aan wegenkaarten (NOB, 1966). Bij het kiezen van deze kortste routes zijn vertragingen door niet-vaste oeververbindingen als fictieve wegafstanden benaderd, en wel het veer Hellevoetsluis-Middelharnis als 50 km, het veer Den Helder- Texel als 20 km en alle overige in Nederland varende veren als 10 km. Dit betekent dus dat, waar mogelijk, wordt omgereden indien de alternatieve omrijdroute korter is dan de kortste route + de fictieve 'veerafstand'.

Zoals gezegd, bevat de berekening nu alle personen met een vast werkadres, die in het studiegebied wonen of werken, waardoor de steekproefgrootte ten opzichte van de totale beroepsbevolking wordt: $\sum_{j \in J} R_j / \sum_{j \in J} P_j = 0,9828$. De calibratie van (15), wederom uitgevoerd zoals in paragraaf 2.3 beschreven, leverde als resultaat:

$\beta = 0,2577$ en waarden A_j zoals in kaart gebracht in fig. 8.

Evenals in paragraaf 3.2 is weer de factor b_j per woonzone berekend door A_j te delen door $R_j = \sum_{i \in I} n_{ij}$, welke in dit geval in het algemeen nagenoeg overeenkomt met de totale beroepsbevolking P_j (fig. 9). Globaal vertoont de figuur eenzelfde geografische spreiding van de b_j -waarden als het kaartbeeld in fig. 6; lage b_j 's in en rond de steden, hoge waarden in de excentrische zones.

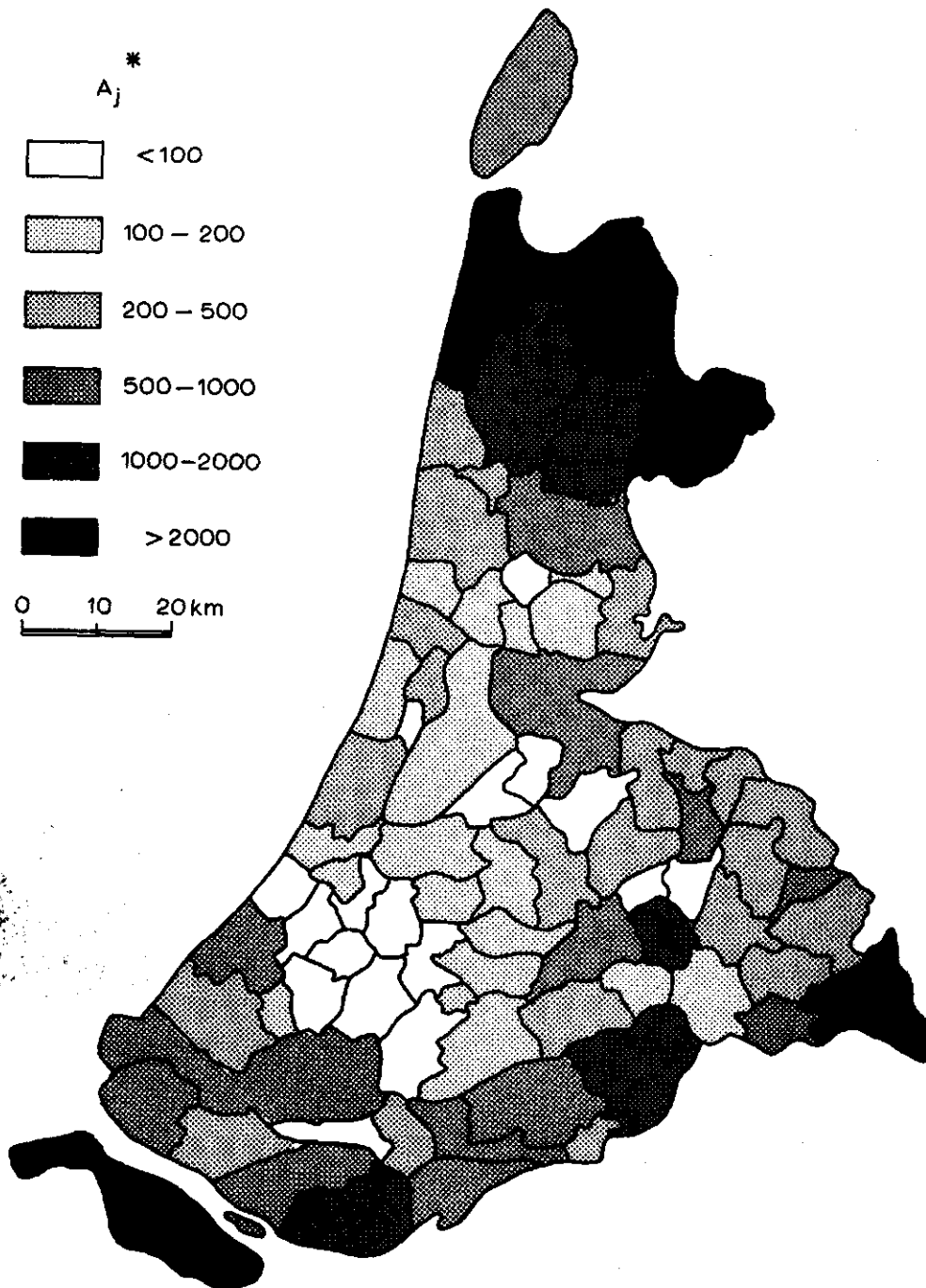


Fig. 8. Attractiefactoren A_j uit model (15), gecalibreerd op de distributie van woon-werkritten tussen 84 zones in 1960. (*Voor de werkelijke waarden vermenigvuldige men alle klassegrenzen met 198.6301)

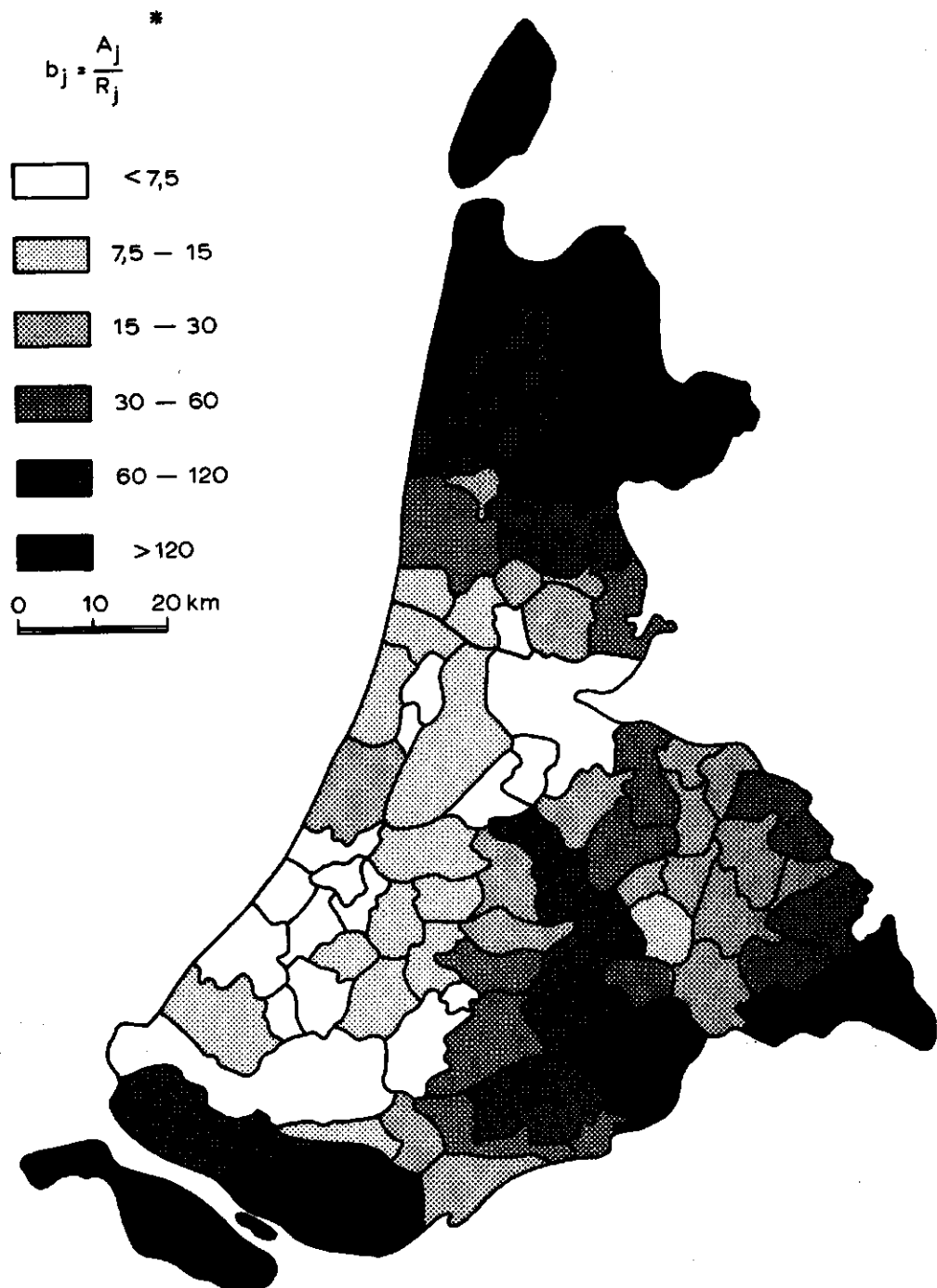


Fig. 9. Competitiefactoren b_j in model (15), gecalibreerd op de distributie van woon-werkritten tussen 84 zones in 1960. (*Voor de werkelijke waarden vermenigvuldigde men alle klassegrenzen met 0,1986301)

Het feit dat alle in het studiegebied werkzaam in de berekening van A_j zijn betrokken, scheidt nu de gelegenheid, om het model te toetsen op de randtotalen $\sum_{i \in I} T_{ij}$ voor elke woonzone j in het studiegebied, die ermee worden voorspeld voor 1971. Uit de Volkstelling 1971 (CBS, 1973) zijn namelijk weliswaar nog niet de woon-werkbewegingen n_{ij} bekend, maar wel reeds de randtotalen van de ritmatrix, $R_j = \sum_{i \in I} n_{ij}$ per woonzone en $E_i = \sum_{j \in J} n_{ij}$ per werkzone. De vergelijking van $\sum_{i \in I} T_{ij}$ met $\sum_{i \in I} n_{ij}$ wordt geïllustreerd door fig. 10. De bovenste figuur toont de sterke correlatie tussen berekende en gemeten randtotalen voor 1960, dus voor de gegevens waarop het model werd gecalibreerd. Kennelijk is het model voldoende flexibel om een goede aansluiting te geven bij de waarnemingen. Met de voor 1960 aangepaste parameters A_j en β is vervolgens uit (15) een ritmatrix berekend voor 1971, met als variabelen de werkgelegenheid E_i per zone en wegaftstanden C_{ij} in 1971, waarbij de waarden a_i en b_j van 1960 onveranderd werden gesteld. De aldus verkregen uitkomsten $\sum_{i \in I} T_{ij}$ staan in de onderste figuur uitgezet tegen de waarnemingen $\sum_{i \in I} n_{ij}$. Behalve de spreiding rond de lijn $y = x$, welke voor de geringe periode van 11 jaar reeds vrij aanzienlijk blijkt, valt verder op, dat de puntenzwerm systematisch van de lijn afwijkt; met name voor de zones waarin de grotere steden liggen zijn de waarnemingen kleiner dan de modeluitkomsten, terwijl de beroepsbevolking in de overige zones in het algemeen sterker blijkt te zijn gegroeid dan dat het model voorspelde. Deze afwijking illustreert, dat het, met het oog op voorspellingen op langere termijn, vooral indien ze de grondslag vormen voor andere prognoses (BUCHANAN c. s., 1973), noodzakelijk is, de modelparameters beter te kunnen beschrijven dan op grond van één steekproef in de tijd mogelijk is. Model (15) zal pas voorspellende waarde kunnen krijgen nadat minstens één soortgelijke set van parameters is berekend op grond van een matrix (n_{ij}) voor een ander tijdstip. Wellicht is het dan mogelijk, de veranderingen in de waarden A_j (danwel b_j) per zone te relateren aan specifieke variabelen van die zones. De eerstvolgende, volledige matrix (n_{ij}) na 1960 voor het onderhavige gebied is waargenomen in 1971. De volkstellingsgegevens van dat jaar betreffende het intergemeentelijk woon-werkverkeer

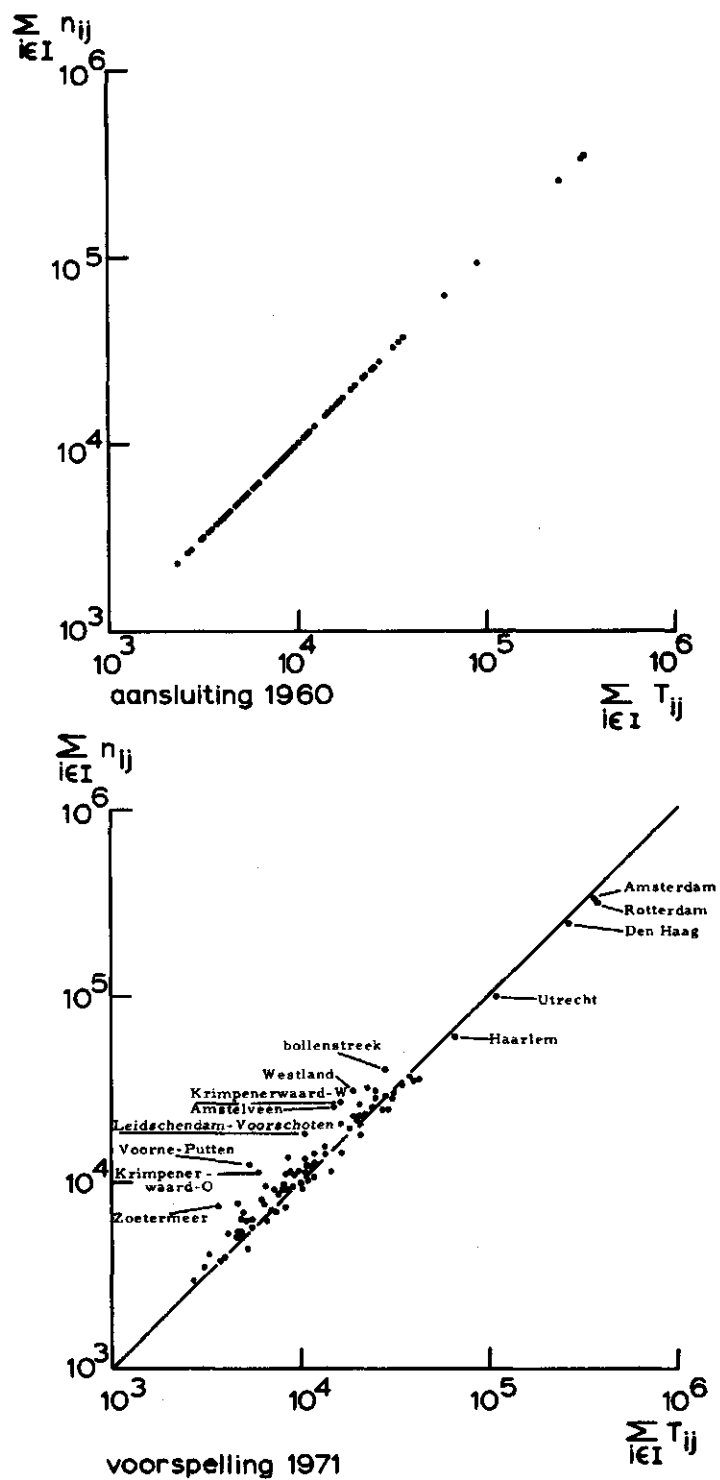


Fig. 10. Gravitiemodel woon-werkverkeer tussen 84 zones in 1960. Aansluiting van (15), gecalibreerd op 1960, en de toetsing van $\sum_{i \in I} T_{ij}$ voor 1971, berekend uit de E_i -waarden van 1971, en de parameters van 1960

zullen evenwel niet eerder dan eind 1974 volledig gepubliceerd kunnen zijn. Niettemin dient hierop te worden gewacht, aangezien ze onmisbaar zijn voor het schatten van de parameters ten behoeve van prognoses.

Tenslotte blijkt het gerechtvaardigd, zeer kritisch te staan tegenover het gebruik van een afstandsfunctie (11) met slechts één voor alle zones geldende, parameter β . Uit partiële regressieberekeningen op ritlengtedistributies van individuele zones volgt namelijk een aanzienlijke spreiding in de β -waarden. Ter illustratie toont tabel 1 enkele van deze waarden, berekend voor rijen en kolommen uit de (n_{ij}) -matrix van het in deze paragraaf behandelde voorbeeld.

Tabel 1. Enkele waarden van de afstandsparemeter β , aangepast aan ritlengteverdelingen van individuele zones

Werkzone	β_i	Woonzone	β_j
Leiden	3,03	Amsterdam	2,85
Utrecht	2,35	Zaanstreek	1,05
Amsterdam	2,08	Beverwijk/Heemskerk	0,85
Delft	1,15	De Streek	0,65
Haarlem	1,07		
Rotterdam	0,90		
Zaanstreek	0,85		
Den Haag	0,82		
Purmerend	0,39		
Haarlemmermeer	0,08		

Dit houdt in, dat, indien slechts één β voor alle zones wordt gehanteerd, de aangepaste A_j -waarden op storende wijze worden beïnvloed. Doordat model (15) deze spreiding niet in de parameter β kan opvangen, nemen de diverse parameters A_j deze taak over en wijken daardoor af van hun werkelijke waarde, naarmate de werkelijke vorm van de ritlengteverdeling afwijkt van de aangepaste functie (fig. 11).

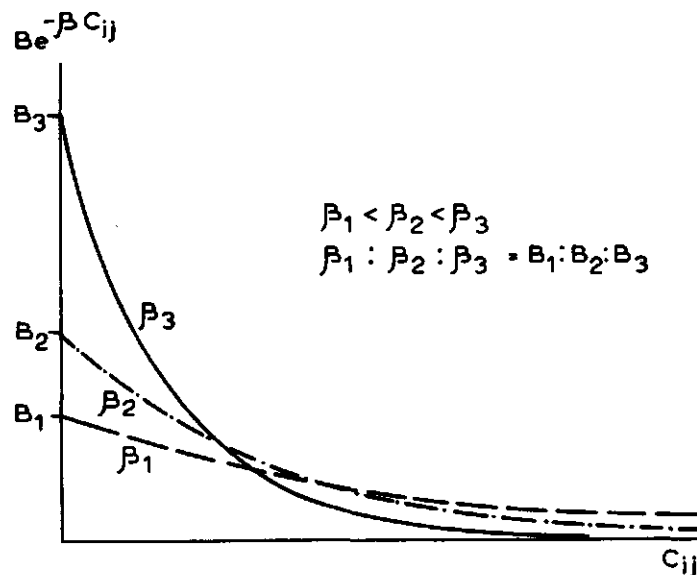


Fig. 11. Voorbeelden van verschillende ritlengteverdelingen met eenzelfde rittentotaal

Deze extra spreiding in A_j als gevolg van de mathematische structuur van (15) is misleidend, indien men A_j poogt te relateren aan bepaalde variabelen per woonzone. Het is daarom dan ook nodig, de flexibiliteit van het model te vergroten, door voor bijvoorbeeld iedere werkzone een specifieke afstandsfunctie aan te passen:

$$f(C_{ij}) = e^{-\beta_i C_{ij}} \quad \text{voor elke } i \in I \quad (32)$$

Aannemelijk lijkt, dat het afstandsgedrag per werkzone afhangt van diverse socio-economische eigenschappen van de in die zone werkzame personen. Hierbij valt te denken aan factoren als het gemiddeld inkomensniveau, sociale beroepsgroep, bedrijfscategorie, gezinsfase, opleidingsniveau, enz. Het is ten behoeve van prognoses dan ook wellicht mogelijk, de parameter β_i te schrijven als functie van dergelijke variabelen.

Dit laatste zou erop neerkomen dat, aan de vraagzijde, de druk van werknemers op woongebieden kan worden gekarakteriseerd door

het afstandsgedrag per werkzone, beschreven door de parameters β_i , terwijl aan de aanbodzijde de verdeling van deze druk wordt bepaald door de attractiewerking van woonzones, beschreven door de parameters A_j . Het onderzoek wordt voortgezet aan de hand van deze benaderingswijze.

4. CONCLUSIES

1. Zoals onder meer door fig. 10 wordt geïllustreerd, mag op grond van de behandelde toepassingen worden geconcludeerd dat op zichzelf een zwaartekrachtsmodel van de vorm (6), onderworpen aan de randvoorwaarden (2) en (4), goed bruikbaar is voor de beschrijving van bestaande (dus meetbare) interactiesystemen, zoals dat van de woon-werkrelaties in de Randstad.
2. Ter voorspelling van de gevolgen van één of enkele veranderingen van de grondgebruiksvariabelen per zone (arbeidsplaatsen, c. q. beroepsbevolking) kan met een op de bestaande situatie gecalibreerd model het nieuwe interactiepatroon nog redelijk worden benaderd door de parameters a_i en b_j van alle overige zones ongewijzigd aan te nemen. Dit laatste is aanvaardbaar indien de wijzigingen in het grondgebruik gering zijn ten opzichte van de totale hoeveelheid in het studiegebied (paragraaf 2.2).
3. Bij voorspellingen op langere termijn (bijv. meer dan 10 jaar) zal in het algemeen het gehele te beschrijven interactiesysteem zó veranderen, dat alle parameters a_i en b_j aanzienlijk kunnen veranderen, waarbij deze elkaar op gecompliceerde wijze beïnvloeden. De enige manier om voor zulke prognoses de parameters te schatten is extrapolatie op grond van een reeks berekeningen ervan in de tijd. Daartoe kan wellicht naar correlatie worden gezocht tussen de individuele veranderingen van de parameters enerzijds en specifieke variabelen per zone anderzijds. Nadere analyse van deze veranderingen in het kader van de hier beschreven toepassingen voor 1960 zal worden verricht na beschikbaar komen van de volkstellingsgegevens betreffende het woon-werkverkeer in 1971.

4. Ten behoeve van de in conclusie 3 genoemde correlatieberekeningen dienen de parameters evenwel zó te worden gecalibreerd, dat ze zo zuiver mogelijk voldoen aan de in paragraaf 2.2 ontwikkelde interpretatie. Dit betekent dat de onderlinge verhouding tussen de berekende parameterwaarden niet mag worden beïnvloed door het onjuist trekken van een steekproef uit de ritdistributie, of door de mathematische structuur van het gecalibreerde model. In verband hiermee wordt in hoofdstuk 3 duidelijk, dat daartoe alle onderscheiden zones in een studiegebied zowel als werkzone als als woonzone moeten worden beschouwd, d.w.z. dat een vierkante herkomst- en bestemmingstabel wordt gebruikt. Verder blijkt in paragraaf 3.3 dat de parameters nog aanzienlijk worden beïnvloed, indien de afstandsfunctie in (15) slechts één parameter β bevat, welke voor het hele gebied geldt.
5. In verband met conclusie 4 verdient het aanbeveling, in het vervolg op dit onderzoek na te gaan of het zinvol is, per werkzone een afstandsfunctie (32) aan te passen aan de betreffende ritlengteverdeling. Vervolgens kan dan worden gepoogd veranderingen in de aldus berekende attractieparameters A_j te beschrijven als functie van ruimtelijke en maatschappelijke variabelen per woonzone, (zoals woonkosten, voorzieningenniveau, woonmilieubeleving, urbanisatiegraad, recreatiemogelijkheden, sociografische structuur, e.d.), en bovendien de afstandsparementers β_i als functie van socio-economische variabelen van de ter plaatse werkende beroepsbevolking (zoals inkomen, sociale beroepsgroep, gezinsfase, opleidingsniveau, e.d.).

LITERATUUR

- ALBRECHT, G. 1972. Soziologie der geographischen Mobilität. Ferdinand Enke Verlag, Stuttgart.
- ANWB KON. NED. TOERISTENBOND. 1963. Autokaart van Nederland, schaal 1 : 200 000, 3 delen. 's-Gravenhage.
- BATTY, M. 1970. Some problems of calibrating the Lowry model. Environment and Planning 2: 95-114.

- BONE, R. 1971. The Tyne-Wear Model. Proceedings of the Seminar on Urban Growth Models. P. T. R. C., London.
- BUCHANAN AND PARTNERS, GRONTMIJ N. V., ADVIESBUREAU ARNHEM B. V. 1973. Midden Randstad, Part I Final Report. London, Arnhem.
- CENTRAAL BUREAU VOOR DE STATISTIEK. 1962. 13e Algemene Volkstelling, 31 mei 1960. (Gemeentelijke gelichtdrukte tabellen). 's-Gravenhage.
- 1964. Typologie van de Nederlandse gemeenten naar urbanisatiegraad, 31 mei 1960. Zeist.
- 1973. 14e Algemene Volkstelling, 28 februari 1971. Voorlopige uitkomsten per gemeente. 's-Gravenhage.
- EVANS, A. W. 1971. The calibration of trip distribution models with exponential or similar cost functions. Transportation Research 5: 15-38.
- GOLDNER, W. 1971. The Lowry model heritage. Journal of the American Institute of Planners XXXVII-2: 100-110.
- GOUDAPPEL, H. M. 1970. Verkeers- en vervoersstudies. ANWB, Serie Verkeerskunde en Verkeerstechniek no. 9. 's-Gravenhage.
- HAMERSLAG, R. 1973. Automatische calibratie van distributiefuncties. Verkeerstechniek 11: 571-575.
- HYMAN, G. M. 1969. The calibration of trip distribution models. Environment and Planning 1: 105-112.
- LOWRY, I. S. 1964. A Model of Metropolis. Rand Corporation R. M. - 4035 - RC, Santa Monica.
- MICHELS, Th. 1972. Een distributiemodel van het forensisme in West-Nederland. Nota ICW 652, Wageningen.
- MIDDELKOOP, L. J. 1970. Forensisme en buiten wonen. Nota ICW 561, Wageningen.
- NATIONALE ORGANISATIE VOOR HET BEROEPSGOEDERENVERVOER WEGTRANSPORT (N. O. B.). 1966. Afstandenwijzer Nederland.

STOL, Ph. Th. 1973. Het gebruik van de negatief-exponentiële kansverdeling voor afvoergegevens en voor het bepalen van de maatgevende afvoer. Deel I: Theoretische achtergronden. Nota ICW 724, Wageningen.

WILSON, A.G. 1970. Entropy in urban and regional modelling. Pion Ltd., London.