

waterloopkundig laboratorium  
delft hydraulics laboratory

multikriteria evaluatie: een verkenning

deel III

technische aspecten:  
toetsing van methoden aan de praktijk

onder redactie van: H. A. Pennekamp

R 999

mei 1985



H13 - 3

toegepast onderzoek  
waterstaat

Roggg-13-C

---

multikriteria evaluatie: een verkenning

deel III

technische aspecten:  
toetsing van methoden aan de praktijk

onder redactie van: H. A. Pennekamp

---

R 999

mei 1985

**low**

H13 - 3

toegepast onderzoek  
waterstaat



## VOORWOORD

Het voor u liggende rapport vormt een weergave van een oriënterende studie naar de mogelijkheden multikriteria evaluatie toe te passen bij de planvorming in het waterbeheer. Deze studie is uitgevoerd door het Waterloopkundig Laboratorium in het kader van het Toegepast Onderzoek Waterstaat, projekt H (systeembenadering in het waterbeheer).

Het rapport bestaat uit drie delen:

- deel I - Toepassingsmogelijkheden van multikriteria evaluatiemethoden  
in het waterbeheer
- deel II - Beslissen bij meervoudige doelstellingen:  
begrippen en methoden
- deel III - Technische aspecten: toetsing van methoden aan de praktijk.

In deel I wordt de relevantie van multikriteria evaluatiemethoden (MKE-methoden) voor in het waterbeheer voorkomende evaluatieprocedures behandeld. en wordt verslag gedaan van gesprekken die gevoerd zijn met een aantal mensen die ervaring hebben opgedaan met de toepassing van MKE-methoden. In deel II wordt nader ingegaan op begrippen die een rol spelen in evaluatieprocedures en op de achtergrond en de kenmerken van methoden die daarbij gebruikt worden. In deel III wordt een aantal van de meest gebruikte MKE-methoden getoetst aan een praktijksituatie en worden de technische aspecten behandeld.

De samenvattende konklusies ten aanzien van de toepassing van MKE-methoden in het algemeen en in het waterbeheer in het bijzonder zijn weergegeven in hoofdstuk 2 van deel I

## INHOUD DEEL III

### LIJST VAN TABELLEN

### LIJST VAN FIGUREN

	blz.
1 <u>Inleiding</u> .....	1
2 <u>Rekenvoorbeeld</u> .....	4
2.1   Algemeen.....	4
2.2   Keuzemotivering.....	5
2.3   Probleemstelling.....	5
2.4   Alternatieven.....	6
2.5   Criteria.....	7
2.6   Effektenmatrices.....	10
2.7   Gewichtenmatrix.....	10
3 <u>MKE-methoden</u> .....	15
3.1   Overzicht.....	15
3.2   Goals achievement matrix (GAM).....	17
3.3   Verwachtingswaardemethoden.....	19
3.4   Permutatiemethoden.....	22
3.5   Eigenwaardemethoden.....	26
3.6   Konkordantiemethoden.....	28
3.7   Meerdimensionale schalingsmethoden.....	31
3.8   MKE-methoden en soort gegevens.....	33
4 <u>Standaardisatie</u> .....	35
4.1   Algemeen.....	35
4.2   Methoden.....	37
4.3   Externe extremen.....	45
4.4   Vergelijking.....	48
5 <u>Gewichten</u> .....	53
5.1   Inleiding.....	53
5.2   Preferentiemethoden.....	54
5.3   Impliciete gewichten.....	56



INHOUD DEEL III (vervolg)

	blz.
5.4 Interne weging.....	59
5.5 WATERLAB.....	61
6 <u>Toepassing</u> .....	71
6.1 Inleiding.....	71
6.2 Invloed van de keuze van standaardisatie- en evaluatiemethode.....	71
6.3 Vergelijking.....	75
6.4 Invloed van gewichten.....	78
6.5 Invloed van onzekerheden.....	84
7 <u>Samenvatting</u> .....	90
7.1 Algemeen.....	90
7.2 Technische eisen MKE-procedure.....	95

APPENDIX - Het programma MKEM

BIBLIOGRAFIE

LIJST VAN TABELLEN

	blz.
2.1 Kardinale effectenmatrix IODZH	11
2.2 Rangorde-effectenmatrix IODZH	12
2.3 Klassen-effectenmatrix IODZH	13
2.4 Gewichtenmatrix IODZH	14
3.1 Goals achievement matrix voor alternatief i	18
3.2 Matrix $P_1$ van kardinale effectwaarden	20
3.3 Matrix $P_2$ van ordinale effectwaarden en gewichtenvectoren $w_1$ en $w_2$	20
	en 23
3.4 Matrix $P_3$ van rangordescores voor $P_2$ en $R_1$ t/m $R_6$	24
3.5 Voorbeeld van een alternatievenvergelijkingsmatrix voor criterium $K_j$	26
4.1 Specificatie van onderzochte standaardisatiemethoden	41
4.2 Illustratie van het effect van schrappen van het slechtst scorende alternatief bij het gebruik van interne extremen bij de standaardisatie	47
4.3 Rangorde van alternatieven bij het gebruik van verschillende standaardisatiemethoden in combinatie met MKE-methode SOM	52
5.1 Impliciete gewichten van standaardisatiemethoden	58
5.2 Gewichtenmatrix voor drie criteria van het IODZH-onderzoek	68
5.3 Gestandaardiseerde gewichtenmatrix voor drie criteria van het IODZH-onderzoek	69
6.1 Vergeleken evaluatie- en standaardisatiemethoden	72
6.2 Totaalscores van de IODZH-alternatieven voor verschillende methoden	74
6.3 Korrelatie van resulterende rangorden	75
6.4 Rangorden voor verschillende gewichtenvectoren bij ENTROPIE	79
6.5 Rangorden voor verschillende gewichtenvectoren bij SAATY met SOM	80
6.6 Spreiding in rangorden voor verschillende gewichtenvectoren bij variatie van methoden	83
6.7 Verdeling van posities in de eindrangschikking met 2% verstoring	86
6.8 Verdeling van posities in de eindrangschikking met $\frac{1}{2}$ % verstoring	87



## LIJST VAN FIGUREN

	blz.
4.1 De relatie tussen scores voor en na standaardisatie: criterium 1 (leveringszekerheid)	49
4.2 De relatie tussen scores voor en na standaardisatie: criterium 6 (gemiddelde produktiekosten)	49
4.3 De relatie tussen scores voor en na standaardisatie: criterium 7 (natuurschade)	50
4.4 De relatie tussen scores voor en na standaardisatie: criterium 9 (bodembelasting)	50
5.1 Interne weging bij standaardisatiemethode NUTS : criterium 1	59
5.2 Trade-off en Pareto-optimaliteit	63
5.3 Trade-off volksgezondheid - natuurschade : kriteria 3 en 7	67
5.4 Trade-off produktiekosten - natuurschade : kriteria 6 en 7	67
5.5 Trade-off volksgezondheid - produktiekosten : kriteria 3 en 6	68
6.1 Grafische vergelijking van evaluatie- en standaardisatiemethoden	77
6.2 Verdeling van posities in de eindrangschikking met $\frac{1}{2}$ % verstoring	89

## 1 Inleiding

In dit derde deel van het onderzoek naar toepassingsmogelijkheden van multi-kriteria evaluatiemethoden (MKE-methoden) in het waterbeheer wordt een toetsing van deze methoden op een praktijkgeval beschreven. Dit deel gaat daarbij inhoudelijk nader in op enkele frekwent gebruikte methoden, waarbij met name de technische aspecten belicht zullen worden.

Gezien de doelstelling van dit onderzoek (zie deel I) bevat dit verslag geen uitgebreide analyse van verschillende MKE-methoden, maar betreft het een verkenning van deze methoden. Voor een uitgebreide analyse van deze en andere methoden wordt verwezen naar de literatuur, bijvoorbeeld naar de rapportage van het onderzoek EVAPLAN dat in opdracht van de Rijks Planologische Dienst door het Planologisch Studiecentrum TNO is verzorgd [1, 2].<sup>1)</sup>

Als specifiek voorbeeld voor de vergelijking van verschillende methoden hebben de resultaten gediend van het Integraal Onderzoek naar de Drinkwatervoorziening van Zuid-Holland (IODZH) [7,46]. Deze studie heeft plaatsgevonden in opdracht van de Minister van Volksgezondheid en Milieuhygiëne, de Minister van Cultuur, Rekreatie en Maatschappelijk Werk en van het Provinciaal Bestuur van Zuid-Holland. Gekoördineerd door het Rijks Instituut voor Drinkwatervoorziening is het onderzoek uitgevoerd door een aantal instituten. Het WL was voornamelijk betrokken bij het systeemonderzoek.

In het navolgende wordt per hoofdstuk een korte samenvatting gegeven, waarmee tegelijkertijd de opbouw van het verslag verduidelijkt wordt.

In hoofdstuk 2 wordt ingegaan op het hierboven reeds genoemde IODZH, waarvan de resultaten als rekenvoorbeeld zullen worden gebruikt ter illustratie van de verschillende methoden. Een beschrijving wordt gegeven van de gehanteerde strategieën en effecten en de effectenmatrix wordt gepresenteerd. Tevens worden enige visies aangegeven die van belang zijn voor het gevoeligheidsonderzoek.

---

1) De getallen tussen [ ] verwijzen naar de algemene bibliografie van het rapport aan het eind van dit deel.



De hoofdstukken 3, 4 en 5 geven een bespreking van beschikbare methoden en vormen als zodanig de kern van dit rapport.

Hoofdstuk 3 behandelt een zestal typen MKE-methoden, waarvan de relevante formules worden gegeven en hun gebruik wordt toegelicht aan de hand van het IODZH-voorbeeld. Na een overzicht van beschikbare methoden worden achtereenvolgens behandeld:

- i. de goals achievement matrix-methode;
- ii. verwachtingswaardemethoden;
- iii. permutatiemethoden;
- iv. de methode Saaty (en een variant daarop) als exponent van eigenwaardemethoden;
- v. konkordantiemethoden; en
- vi. meerdimensionale schalingsmethoden.

Bij hoofdstuk 4 gaat het om standaardisatiemethoden, dit zijn methoden om (kardinale) gegevens onder dezelfde noemer te brengen, zodat hun waarden onderling vergelijkbaar worden. In totaal wordt een twaalftal methoden uitgewerkt, waaronder een in dit onderzoek ontwikkelde methode. Verder wordt ingegaan op enige problemen die verband houden met standaardisatie.

Hoofdstuk 5 gaat in op de problematiek rond het vaststellen en gebruik van gewichten, toe te kennen aan criteria om visies te representeren. Behalve expliciete weging door gewichtenvectoren wordt ook impliciete weging besproken. De relatie met standaardisatie wordt gelegd, waarbij ook het begrip interne weging wordt geïntroduceerd. Het hoofdstuk wordt besloten met een bespreking van een tijdens het onderzoek ontwikkelde methode WATERLAB.

Aan een vergelijking tussen verschillende van de behandelde methoden wordt in hoofdstuk 6 aandacht besteed. Als uitgangspunt zijn hierbij steeds de IODZH-gegevens uit hoofdstuk 2 en de daar gegeven visies gehanteerd. Op een aantal punten worden gevoeligheidsonderzoeken beschreven; deze hebben betrekking op de keuze van standaardisatie- en evaluatiemethoden, verschillende gewichtenvectoren en onzekerheden in de effektscores.

Het materiaal uit dit deel van het rapport wordt in hoofdstuk 7 samengevat. Hieraan wordt een pakket van eisen in technische zin die aan een MKE-procedure zouden moeten worden gesteld vastgeknoopt, en die kan dienen voor een evaluatie van MKE op technische gronden.

Aan het rapport is een appendix toegevoegd; deze geeft een summiere beschrijving van het computerprogramma MKEM, waarin enige methoden zijn geïmplementeerd en met behulp waarvan de berekeningen voor de hoofdstukken 3, 4, 5 en 6 van dit deel van het rapport zijn uitgevoerd.



## 2 Rekenvoorbeeld

### 2.1 Algemeen

Met het doel inzicht en ervaring te krijgen in het toepassen van MKE-methoden zijn aan de hand van een praktijkgeval voorbeeld-berekeningen gemaakt (zie met name hoofdstuk 6).

Bij het toepassen van MKE moeten keuzen gemaakt worden ten aanzien van de soort uitvoergegevens en de te gebruiken methoden. In het hier gepresenteerde rekenvoorbeeld zal met name aandacht worden besteed aan de keuze met betrekking tot:

- het type data (kardinaal of ordinaal);
- de standaardisatiemethode; en
- de MKE-methode.

Nagegaan zal worden hoe de resultaten van de berekeningen, die de vorm hebben van rangorden van alternatieven (waarderingen), beïnvloed worden door deze keuzen.

Een vergelijking van bestaande en voor toepassing in het waterbeheer in aanmerking komende MKE-methoden aan de hand van praktijkgevallen heeft slechts een beperkte waarde. De voornaamste beperking is wel gelegen in het feit dat elke beslissituatie specifieke kenmerken heeft ten aanzien van de onderkende alternatieven, de criteria en de skores daarop, en de gewichten (zie ook deel I). De gemaakte vergelijking heeft daarom overwegend een illustratief karakter.

Het gekozen praktijkgeval is het Integraal Onderzoek Drinkwatervoorziening Zuid-Holland (IODZH). In paragraaf 2.2 wordt de keuze van dit onderzoek gemotiveerd en paragraaf 2.3 gaat kort in op de probleemstelling van het IODZH. Vervolgens geeft paragraaf 2.4 een beschrijving van de alternatieven en paragraaf 2.5 doet hetzelfde met betrekking tot de criteria. Paragraaf 2.6 presenteert de effektenmatrix, waarna paragraaf 2.7 de gewichtenmatrix bevat, waarmee gewerkt zal worden. Voor meer informatie over het IODZH wordt verwezen naar [7] en [46].

## 2.2 Keuzemotivering

Aan het Integraal Onderzoek Drinkwatervoorziening Zuid-Holland (IODZH) is een effectenmatrix ontleend die betrekking heeft op 24 beleidsalternatieven en 10 beslissingskriteria<sup>1)</sup> De motivering om deze matrix als uitgangspunt voor het rekenvoorbeeld te nemen is drieledig:

- Het betreft hier een voorbeeld uit de praktijk, waardoor resultaten gemakkelijk interpreteerbaar zijn en een gevoel voor de invloed van de diverse methoden kan worden gekweekt.
- Vanwege de betrokkenheid van het Waterloopkundig Laboratorium bij het IODZH is er kennis aanwezig omtrent de gehanteerde criteria, de alternatieven, de gewichten die bij de diverse politieke of bestuurlijke visies horen, de wijze en nauwkeurigheid van de effectenberekeningen, enz.
- De IODZH-effectenmatrix heeft een volledig kardinaal karakter, waaruit naar wens op systematische wijze diverse ordinale effectenmatrices kunnen worden afgeleid. Zodoende is het mogelijk uitkomsten van kardinale en ordinale MKE-methoden met elkaar te vergelijken.

## 2.3 Probleemstelling

De volgende algemene omschrijving en probleemstelling is aan [7] ontleend.

Onder meer bij de herziening van het Struktuurschema Drink- en Industriewater-voorziening 1972 is gebleken dat voor de planning van toekomstige infrastrukturele werken in de provincie Zuid-Holland een integrale benadering noodzakelijk is. Met name betreft dit de plaats, die duininfiltratie hierin dient in te nemen op middellange en lange termijn. In het bijzonder in het duingebied blijkt een tegenstelling aanwezig te zijn tussen de belangen van de drinkwatervoorziening, het natuur- en landschapsbehoud en de recreatie. Deze tegenstellingen manifesteren zich in de diverse beleidsvoornemens op rijks- en pro-

---

1) In de originele effectenmatrix is sprake van 11 criteria, doch bij de beoordeling van de alternatieven in het eindrapport [46] is geen rekening gehouden met het elfde en laatste criterium: de bestuurlijk-juridische implicaties.



vinciaal niveau. Daarnaast is de in het geding zijnde problematiek een aktueel beleidsvraagstuk geworden, doordat aanvragen om vergunningen voor realisatie van onderdelen van het drinkwatersysteem in behandeling zijn. Tenslotte wordt door een onvoldoende integraal inzicht de toetsing bemoeilijkt van Tienjarenplannen voor de openbare watervoorziening aan het rijksoverheids- en provinciale beleid.

Gelet op het voorgaande is het zogeheten Integraal Onderzoek Drinkwatervoorziening Zuid-Holland (IODZH) gestart. Hiertoe is door de Minister van Volksgezondheid en Milieuhygiëne, handelende mede namens de Minister van Verkeer en Waterstaat, de Staatssekretaris van Cultuur, Rekreatie en Maatschappelijk Werk en Gedeputeerde Staten van Zuid-Holland, een stuurgroep ingesteld met als taak (uit [46], p. 7):

- "a. het initiëren, doen uitvoeren en begeleiden van onderzoek met betrekking tot de toekomstige drinkwatervoorziening van Zuid-Holland en daarmee samenhangende belangen in onderling verband;
- b. het uitbrengen van een rapport over mogelijke oplossingen voor de vraagstukken met betrekking tot de toekomstige drinkwatervoorziening met aanduiding van de gevolgen van deze oplossingen voor de daarmee verband houdende belangen."

Op grond van het voorgaande is in het onderzoek de volgende vraagstelling centraal gesteld (uit [46], p. 17):

"In welke mate en onder welke voorwaarden kunnen de duinen een functie vervullen in de drink- en industriewatervoorziening van de provincie Zuid-Holland in relatie tot natuurbehoud en rekreatie. Welke andere mogelijkheden doen zich voor om in de vraag naar water te voorzien en wat zijn de voor- en nadelen daarvan voor alle relevante belangen waarop de drink- en industriewatervoorziening invloed uitoefent".

#### 2.4 Alternatieven

<u>Alternatieven</u>	<u>Beschrijving</u>
1 en 2	<u>Duininfiltratie</u> zeer groot (1) tot maximaal (2). Alternatief 1 stelt de huidige beleidslijn van de drinkwaterleidingbedrijven voor.

<u>Alternatieven</u>	<u>Beschrijving</u>
3 t/m 7	<u>Diepinfiltratie</u> vervangt duininfiltratie geheel of gedeeltelijk (3: vervanging beperkt; 7: vervanging maximaal, 4, 5 en 6 nemen tussenposities in).
8 t/m 12	<u>Oppervlaktewaterzuivering Kralingen</u> vervangt oppervlakkige duininfiltratie geheel of gedeeltelijk. (8: vervanging beperkt, 12: vervanging maximaal, 9, 10 en 11 nemen tussenposities in).
13 t/m 15	<u>Oppervlaktewaterzuivering Kralingen en Zoetermeer</u> vervangen samen geheel of gedeeltelijk de oppervlakkige duininfiltratie (13: beperkte vervanging, 15: gehele vervanging, 14 neemt tussenpositie in).
16 t/m 20	<u>Oppervlaktewaterzuivering Zoetermeer</u> vervangt de oppervlakkige duininfiltratie; levering vanuit Kralingen is alleen toegestaan aan het behoeftegebied Rotterdam (16: beperkte vervanging, 20: gehele vervanging, 17, 18 en 19 nemen tussenposities in).
21 t/m 24	<u>Oeverinfiltratie en eventueel aanvullend brak en zout grondwater</u> (via omgekeerde osmose ontzout) vervangen de oppervlakkige duininfiltratie; geen levering vanuit Kralingen buiten het behoeftegebied Rotterdam; geen oppervlaktewaterzuivering Zoetermeer (21: vervanging beperkt, 24: vervanging zo veel mogelijk, 22 en 23 nemen tussenposities in).

## 2.5 Kriteria

De bij het IODZH betrokken beslissingscriteria zijn alle zo gedefinieerd dat steeds geldt dat hoe lager de score is des te gunstiger het alternatief met betrekking tot het criterium in kwestie moet worden beoordeeld. Ter illustratie van het type criteria waarmee bij een MKE ten behoeve van de waterhuishouding rekening kan worden gehouden wordt hieronder elk criterium in het kort toegelicht.



### 1. Leveringszekerheid

Het gaat hier om de gevolgen van kalamiteuze verontreinigingen aan de bron, storingen in zuiveringssystemen, leidingbreuken, enz. De leveringszekerheid wordt tot uitdrukking gebracht in het promillage van de totale vraag in Zuid-Holland waaraan jaarlijks niet voldaan kan worden.

### 2. Kwaliteit van het drinkwater

De kwaliteit van het drinkwater dat door het voorzieningssysteem wordt geproduceerd, wordt gekarakteriseerd met behulp van een twaalftal geselecteerde parameters die elk 90% van de tijd onder een bepaalde norm moeten blijven. De kwaliteit van het drinkwater wordt tot uitdrukking gebracht in een tussen 0 en 0,4 liggende kwaliteitsindex. Deze index zou gelijk zijn aan 1 als de gemiddelde waarde van elke parameter altijd precies gelijk was aan de norm. Waarden van één of meer parameters beneden de norm leiden tot een lagere index.

### 3. Volksgezondheidsaspecten van het drinkwater

Voor volksgezondheid is een zelfde type index gedefinieerd als de kwaliteitsindex waarbij veertien parameters gekozen zijn die van belang zijn voor de volksgezondheid.

### 4. Energieverbruik

Het energieverbruik is afzonderlijk als criterium opgenomen, omdat grote onzekerheden heersen omtrent de toekomstige ontwikkelingen. Een zo energie-onafhankelijk mogelijke drinkwatervoorziening wordt nagestreefd. Het energieverbruik wordt gemeten in kWh/m<sup>3</sup> drinkwater.

### 5. Investerings

De kontante waarde van de investeringen in het voorzieningssysteem tot aan de planhorizon (het jaar 2010) wordt als criterium gehanteerd met het oog op de in de toekomst te maken aanspraken op de kapitaalmarkt.

## 6. Produktiekosten

Het gaat hier om produktiekosten per m<sup>3</sup> drinkwater. Deze kosten bevatten zowel variabele als vaste kosten van winning, transport en zuivering van ruwwater tot drinkwater.

## 7. Natuurschade

Het natuurlijke milieu zal invloed ondervinden van diverse ingrepen, die ten behoeve van de drinkwatervoorziening moeten worden gedaan. Te denken valt aan inundatie, vergraving, verharding, hydrologische beïnvloeding. Natuurschade wordt uitgedrukt in een schadegetal dat in hectaren aangeeft de hoeveelheid natuurterrein, gewogen naar vegetatiesoort, dat verloren gaat (schadegetal positief) aan de drinkwatervoorziening of wordt geregenereerd (schadegetal negatief) bij stoppen van schadelijke activiteiten. Rekening is tevens gehouden met de duur van de beïnvloeding, c.q. de tijd benodigd voor herstel.

## 8. Direkte produktie van afvalstoffen

In het bereidingsproces van drinkwater komen afvalstoffen vrij (zuiverings-slib). Deze hoeveelheden zijn weergegeven in tonnen voor het jaar 2010.

## 9. Bodembelasting

Voor een eerste beoordeling van grondwaterverontreiniging door lozing van bepaalde gevaarlijke stoffen wordt de hoeveelheid door kunstmatige infiltratie in de ondergrond gebrachte stoffen aangegeven, gemeten in tonnen voor het jaar 2010.

## 10. Bezinking van stoffen in spaar- en procesbekkens

In drinkwaterbekkens vindt een al dan niet gestimuleerde bezinking van stoffen plaats. De omvang van deze bezinking in tonnen per jaar in het jaar 2010 is als maat beschouwd voor de omvang van een potentieel probleem samenhangend met het vrijkomen van slib, waardoor baggeren nodig wordt.



## 2.6 Effektenmatrices

In tabel 2.1 (zie figuur 3.3.2 uit [46]) is de kardinale effektenmatrix weergegeven. Rekening dient te worden gehouden met het feit, dat de kleinste scores overeenkomen met de meest positieve effekten voor een criterium. Zij vormt de basis voor twee ordinale effektenmatrices:

- i. De rangorde-effektenmatrix (tabel 2.2). Per criterium zijn aan de effect-scores van de alternatieven rangnummers gegeven uitgaande van de kardinale scores. Het rangnummer 24 wordt aan de laagste (beste) score en het rangnummer 1 aan de hoogste (slechtste) score gegeven. Aan gelijke scores wordt hetzelfde rangnummer toegekend (het hoogste).
- ii. De klassen-effektenmatrix (tabel 2.3). Per criterium zijn vijf klassen gedefinieerd. Daarna is op grond van de kardinale scores per criterium bepaald in welke klasse een alternatief scoort. Klassegrenzen zijn zodanig bepaald dat gelijkwaardige of bijna gelijkwaardige scores in één klasse worden samengevoegd. Doordat deze klassegrenzen onafhankelijk van de effectskores zijn gedefinieerd, komt het voor dat criteria slechts in enkele klassen skoren (zie bijvoorbeeld criterium 9). De keuze van deze klassen is arbitrair en komt niet overeen met de indeling in [46, tabel 6.2.1].

## 2.7 Gewichtenmatrix

Bij het IODZH zijn verschillende beleidsinstanties betrokken die aan de onderscheiden criteria verschillende gewichten toekennen. Door ambtenaren die aan dit onderzoek hebben meegewerkt, is in het kader van onderhavig onderzoek een zestal alternatieve gewichtenvectoren opgesteld die de visies van deze beleidsinstanties weergeven (zie tabel 2.4).

alternatieven	maximale duininfiltratie		diepinfiltratie als alternatief voor duininfiltratie					zuivering Kralingen als alternatief voor duininfiltratie					Kralingen/Zoetermeer i.p.v. duininfiltratie					zuivering Zoetermeer als alternatief voor duininfiltratie					oevergrondwater en zout grondwater voor duininfiltratie			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
1. leveringszekerheid [0/oo vd vraag]	1,3	1,4	1,6	1,5	1,4	1,3	1,3	1,5	1,7	1,7	1,5	1,4	1,7	1,5	1,3	1,5	1,6	1,6	1,4	1,3	1,7	1,9	2,3	2,1		
2. drinkwaterkwaliteitsindex in 2010 [10 <sup>-3</sup> ]	149	148	149	148	147	146	144	148	153	158	163	164	156	160	161	149	150	150	151	150	150	150	151	151		
3. volkgezondheidsindex in 2010 [10 <sup>-3</sup> ]	6,5	6,4	6,5	6,4	6,3	6,2	6,3	6,5	6,6	6,9	7,1	7,2	6,8	7,3	7,3	6,6	6,7	7,3	7,8	7,6	6,5	6,4	6,4	6,5		
4. energieverbruik in 2010 [kWh/m <sup>3</sup> ]	0,39	0,42	0,39	0,39	0,40	0,39	0,39	0,39	0,39	0,41	0,43	0,43	0,41	0,43	0,43	0,40	0,40	0,42	0,43	0,43	0,38	0,38	0,44	0,54		
5. gediskonteerde investeringskosten [106 gulden]	372	467	334	369	438	489	670	323	304	396	479	506	424	514	550	348	386	499	583	616	324	343	382	390		
6. gemiddelde productie-kosten [gulden/m <sup>3</sup> ]	0,79	0,82	0,78	0,79	0,82	0,84	0,92	0,77	0,73	0,80	0,82	0,82	0,79	0,85	0,87	0,78	0,79	0,85	0,88	0,89	0,77	0,76	0,77	0,80		
7. gevogen natuurschade [ha]	277	384	246	173	-88	-305	-360	239	164	-124	-360	-421	-126	-362	-423	240	155	-129	-367	-429	245	166	64	-103		
8. directe produktie afvalstoffen in 2010 [10 <sup>3</sup> ton]	2,35	1,84	2,34	2,60	2,34	2,31	1,64	2,31	2,29	2,16	2,04	1,38	2,21	2,11	1,46	2,34	2,55	2,32	2,29	1,65	2,32	4,17	5,44	10,89		
9. bodembelasting afvalstoffen in 2010 [10 <sup>3</sup> ton]	0,52	0,16	0,51	0,47	0,49	0,47	0,10	0,50	0,65	0,61	0,57	0,87	0,61	0,57	0,87	0,50	0,58	0,61	0,57	0,87	0,74	1,16	1,41	1,58		
10. afvalstoffen in bekken in 2010 [10 <sup>3</sup> ton]	11,25	13,54	11,25	11,59	11,59	11,59	12,52	11,01	10,61	9,65	8,78	8,98	10,12	9,50	9,76	11,25	11,55	11,22	11,22	11,58	10,73	9,84	9,05	7,66		

Tabel 2.1 Kardinale effectenmatrix IODZH



alternatieven	maximale duininfilt- ratie	diepfiltratie als alternatief voor duinfiltratie	zuivering Kralingen als alternatief voor duinfiltratie	Kralingen/Zoe- termeer i.p.v. duinfiltratie	zuivering Zoetermeer als alternatief voor duinfiltratie	oevergrondwater en zout grondwater voor duinfiltratie
kriteria	1 2	3 4 5 6 7	8 9 10 11 12	13 14 15	16 17 18 19 20	21 22 23 24
1. leverings- zekerheid	24	19	10 15 19 24 24	15 7 7 15 19	7 15 24	7 3 1 2
2. drinkwater- kwaliteits- index	18	21	18 21 22 23 24	21 7 5 2 1	6 4 3	15 15 10 10
3. volksge- zondheids- index	17	21	17 21 23 24 23	17 12 8 7 6	9 5 5	17 21 21 17
4. energiever- bruik in 2010	22	10	22 22 15 22 22	22 22 12 8 8	12 8 8	24 24 2 1
5. gediscon- teerde inverte- ringen	17	10	21 18 11 8 1	23 24 13 9 6	12 5 4	22 20 16 14
6. gemiddelde produktie- kosten	17	11	19 17 11 7 1	22 24 13 11 11	17 6 4	22 23 22 13
7. gewosen natuur- schade	2	1	3 7 12 17 18	6 11 14 19 22	15 20 23	4 8 10 13
8. directe pro- duktie van afvalstof- fen	6	20	9 4 9 13 22	13 15 17 19 24	16 18 23	11 3 2 1
9. bodembelas- ting met af- valstoffen in 2010	16	23	17 22 20 22 24	19 8 11 15 6	11 15 6	7 3 2 1
10. afvalstof- fen in bek- kens in 2010	10	1	10 5 5 5 2	13 15 19 23 22	16 20 18	14 17 22 24

Tabel 2.2 Rangorde-effektenmatrix IODZH

alternatieven	maximale duinfil- tratie	diepinfiltratie als alternatief voor duinfiltratie	zuivering Kralingen als alternatief voor duinfiltratie	Kralingen/zoe- meer i.p.v. duinfiltratie	zuivering Zoetermeer als alternatief voor duinfiltratie	oevergrondwater en zout grondwater voor duinfiltratie																		
kriteria	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Leverings- zekerheid	4	4	3	3	4	4	4	3	3	3	3	4	3	3	4	3	3	3	4	4	3	2	2	2
2. drinkwater- kwaliteits- index in	4	4	4	4	5	5	5	4	3	3	2	2	3	2	2	4	4	4	4	4	4	4	4	4
3. volgsge- zondheids- index	4	4	4	4	5	5	5	4	4	3	2	2	3	2	2	4	3	2	1	1	4	4	4	4
4. energiever- bruik in 2010	4	3	4	4	4	4	4	4	4	3	3	3	3	3	3	4	4	3	3	3	4	4	2	1
5. gediscou- teerde inves- tingen	4	2	5	4	3	2	1	5	5	4	2	1	3	1	1	5	4	2	1	1	5	5	4	4
6. gemiddelde productie- kosten	4	3	4	4	3	3	1	5	5	4	3	3	4	3	2	4	4	3	2	2	5	5	5	4
7. Gewogen natuur- schade	1	1	1	2	3	4	4	1	2	3	4	5	3	4	5	1	2	3	4	5	1	2	2	3
8. directe pro- ductie van afvalstof- fen in 2010	3	4	3	3	3	3	4	3	3	3	3	4	3	3	4	3	3	3	3	4	3	2	2	1
9. bodembelas- ting met af- valstoffen in 2010	4	5	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3
10. afvalstof- fen in bek- kens in 2010	4	3	4	4	4	4	3	4	4	5	5	5	4	5	5	4	4	4	4	4	4	5	5	5

Tabel 2.3 Klasse-effectenmatrix IODZH



beslissingscriteria	gewichtenvectoren					
	A	B	C	D	E	F
1. Leveringszekerheid	0,06	0,04	0,08	0,07	0,05	0,05
2. Drinkwaterkwaliteit	0,12	0,08	0,16	0,13	0,10	0,10
3. Volksgezondheid	0,12	0,08	0,16	0,13	0,10	0,10
4. Energieverbruik	0,06	0,04	0,06	0,07	0,10	0,05
5. Investerings	0,12	0,08	0,12	0,13	0,20	0,10
6. Productiekosten	0,12	0,08	0,12	0,13	0,20	0,10
7. Natuurschade	0,30	0,50	0,30	0,33	0,25	0,50
8. Direkte prod. afvalstoffen	0,03	0,03	0,00	0,00	0,00	0,00
9. Bodembelasting afvalstoffen	0,03	0,03	0,00	0,00	0,00	0,00
10. Afvalstoffen in bekkens	0,03	0,03	0,00	0,00	0,00	0,00

Tabel 2.4 Gewichtenmatrix IODZH

### 3 MKE-methoden

Na een overzicht in paragraaf 3.1 van beschikbare MKE-methoden, waarin een poging wordt ondernomen het veld te structureren behandelt de rest van dit hoofdstuk deze methoden in meer detail. Het gaat hierbij om:

- paragraaf 3.2: goals achievement matrix;
- paragraaf 3.3: verwachtingswaardemethoden;
- paragraaf 3.4: permutatiemethoden;
- paragraaf 3.5: eigenwaardemethoden;
- paragraaf 3.6: konkordantiemethoden;
- paragraaf 3.7: meerdimensionale schalingsmethoden.

Bovenstaande methoden hebben als overeenkomstig kenmerk dat uitgegaan wordt van een (tenminste) tweedimensionale effectiviteitenmatrix waarbij in één richting de criteria zijn uitgezet en langs de andere as de alternatieven. Als derde dimensie kunnen de verschillende maatschappelijke groeperingen genomen worden. Tezamen met een gewichtenvector of -matrix leveren de in dit hoofdstuk behandelde methoden een (eventueel zwakke) ordening op van de alternatieven, dit in tegenstelling tot andere groepen evaluatiemethoden; in deel II worden MKE-methoden gezet in het wijdere perspectief van evaluatiemethoden. In paragraaf 3.8 wordt aandacht besteed aan de mogelijkheid dat de effectenmatrix zowel kardinale als ordinale informatie bevat. In deze paragraaf wordt een methode uiteengezet die expliciet rekening houdt met deze mogelijkheid.

#### 3.1 Overzicht

In grote lijnen kunnen vijf groepen MKE-methoden worden onderscheiden: gewogen sommeringsmethoden, permutatiemethoden, eigenwaardemethoden, konkordantiemethoden en multidimensionale schalingsmethoden.

##### Gewogen sommeringsmethoden

Dit zijn ongetwijfeld de meest gebruikte methoden, wellicht ook door de grote eenvoud en overzichtelijkheid. Ook in andere MKE-methoden komt gewogen somming veelal als onderdeel van de totale methode voor. In matrixnotatie kan



de gewogen someringsmethode als volgt weergegeven worden:

$$P = W.E \quad (3.1)$$

waarbij:

P = waarderingsmatrix

W = gewichtenmatrix

E = effectiviteitenmatrix

In dit hoofdstuk worden de Goals Achievement Matrix (GAM) en de verwachtingswaardemethoden behandeld.

#### Permutatiemethoden

Permutatiemethoden gaan uit van het principe dat iedere mogelijke rangschikking van keuzemogelijkheden getest kan worden op de mate van overeenstemming met de effectenscores en gewichten. In dit hoofdstuk wordt de methode van Paelinck nader toegelicht.

#### Eigenwaardemethoden

Bij deze methoden worden de effektscores van de alternatieven op de criteria en de gewichtenvector van de criteria afgeleid door middel van de bepaling van de eigenvector van paarsgewijze vergelijkingsmatrices. Deze matrices bevatten respectievelijk de onderhavige dominantie van de verschillende alternatieven per criterium afzonderlijk en de dominantie van de criteria ten opzichte van elkaar. In dit hoofdstuk wordt de methode van Saaty besproken, die in Nederland een zekere mate van populariteit bevat [21].

#### Konkordantiemethoden

Deze methoden hebben als kenmerk dat de verschillende alternatieven paarsgewijs met elkaar vergeleken worden teneinde voor ieder alternatief dominantieindices te bepalen [17, 22, 23, 26, 44]. Deze indices kunnen betrekking hebben op relatieve dominantie in termen van de gewichten van de criteria waarop de alternatieven vergeleken worden (konkordantie) en op de mate waarin de alter-

natieven van elkaar verschillen (diskordantie). Onderscheid kan worden gemaakt tussen konkordantie rangordemethoden en konkordantie eliminatiemethoden. Deze laatste methoden leveren als eindresultaat alternatieven op die aan bepaalde drempelwaarden van effektskores voldoen.

### Multidimensionale schalingsmethoden

Schalings technieken zijn ontwikkeld in de psychometrie en hebben tot doel grote verzamelingen van gegevens weer te geven in een ruimte met zo klein mogelijke dimensies. Hiervoor is een breed scala aan multivariate technieken beschikbaar: bijvoorbeeld principale componentenanalyse, faktoranalyse, interdependentie-analyse. Toepassing ten behoeve van multikriteria evaluatie waarbij voornamelijk ordinale informatie beschikbaar is, is relatief nieuw [25].

Gewerkt wordt vanuit de hypothese dat de criteria beïnvloed worden door slechts een beperkt aantal achterliggende factoren. Aan deze hypothese is in de praktijk lang niet altijd voldaan. Een kans op het met succes toepassen is voornamelijk aanwezig bij een effektenmatrix van redelijk grote omvang<sup>1)</sup>.

Een eerste stap is de klassifikatie van keuzemogelijkheden op basis van het beperkte aantal achterliggende factoren. Een tweede stap is de waardering van de resulterende klassen met behulp van de criterium-gewichten.

Een nadeel van de methode kan zijn dat het met name bij problemen met een groot aantal alternatieven en criteria onduidelijk kan zijn hoe de uiteindelijke rangschikking tot stand komt.

### 3.2 Goals achievement matrix (GAM)

Bij de goals achievement matrixmethode (GAM) worden gekwantificeerde effekten (ordinaal of kardinaal) gedefinieerd in relatie tot de onderscheiden criteria. Het is gebruikelijk per criterium de effekten (voor- en nadelig) op eenzelfde schaal te meten (ratio- of intervalschaal) en in eenzelfde eenheid (bijvoorbeeld na standaardisatie). Verondersteld wordt nu dat het criteriapakket bekend is en dat de relatieve waarde (gewicht, prioriteit), die aan elke doel-

1) [9] stelt als vuistregel tenminste 5 alternatieven en 5 criteria.



stelling wordt toegekend, is bepaald. Vervolgens wordt verondersteld dat het gewicht dat aan verschillende maatschappelijke groeperingen moet worden toegekend, ook vastgesteld is. Voor elk alternatief wordt dan een goals achievement matrix opgesteld, waarvan de structuur is als in tabel 3.1 is aangegeven; zie ook [24].

kriterium (j)	X <sub>1</sub>			X <sub>2</sub>		
gewicht (a <sub>j</sub> )	2			3		
groepen (q)	groeps- gewicht (b <sub>q1</sub> )	kost (K <sub>q1i</sub> )	baat (B <sub>q1i</sub> )	groeps- gewicht (b <sub>q2</sub> )	kost (K <sub>q2i</sub> )	baat (B <sub>q2i</sub> )
a	1	A	-	5	-	J
b	2	B	D	4	-	K
c	1	C	E	3	-	L
d	4	-	F	1	H	M
e	2	-	G	4	I	N
totaal		...	...		...	...

Tabel 3.1 Goals achievement matrix voor alternatief i

In deze tabel worden de criteria aangeduid met X<sub>j</sub> (j=1,...,J). De verschillende maatschappelijke groeperingen a,b,c,... worden geïdentificeerd aan de hand van de betrokkenheid bij de onderkende planalternatieven. De letters A,B,C,... geven de kosten en baten weer, die zowel in financiële, als niet-financiële eenheden weergegeven kunnen worden. Door nu voor ieder kriterium de scores per groep (al dan niet gewogen) op te tellen en vervolgens deze gesommeerde scores per kriterium voor de gehele goals achievement matrix op te tellen (wederom al dan niet gewogen) kan een "index of goals achievement" bepaald worden. Formeel kan de index als volgt voorgesteld worden:

$$Z_i = \sum_{j=1}^J [a_j \sum_{q=1}^Q \{b_{qj}(B_{qji} - K_{qji})\}] \quad (3.2)$$

waarbij:

$Z_i$  = gewogen index voor doelsukses voor alternatief  $i$  ( $i=1, \dots, I$ )

$a_j$  = gewicht van criterium  $j$  ( $j=1, \dots, J$ )

$b_{qj}$  = groepsgewicht voor groep  $q$  ( $q=1, \dots, Q$ ) en criterium  $j$

$B_{qji}$  = baten van alternatief  $i$  t.a.v. criterium  $j$  voor groep  $q$

$K_{qji}$  = kosten van alternatief  $i$  t.a.v. criterium  $j$  voor groep  $q$

Het is duidelijk dat een van de fundamentele moeilijkheden van de goals achievement matrixmethode gelegen is in de aggregatieprocedure. Het optellen van ongewogen en ongenormeerde effectwaarden leidt tot weinig zinvolle uitkomsten, terwijl het bepalen van de gewichten en normwaarden daarentegen stuit op het probleem van subjectieve waarderingen. De methode is meer geschikt voor het presenteren dan voor het uitvoeren van een evaluatie van de verschillende alternatieven. Dit is ook de mening die uit [18] naar voren komt.

### 3.3 Verwachtingswaardemethoden

Een andere, vrij eenvoudige, MKE-methode is de "expected value" methode, ook wel de "rank-based expected value" methode genoemd. Aan de criteria worden (per visie) gewichten gegeven in de vorm van een rangorde door middel van discrete getallen ( $1, 2, \dots, J$ ), waarbij het criterium waaraan het minste belang wordt gehecht de rangorde 1 krijgt toegekend. Vervolgens wordt per criterium een rangorde aan de alternatieven gegeven op basis van de effectenwaarde. Indien twee of meer alternatieven dezelfde score voor een criterium hebben, dan neemt men gewoonlijk het rekenkundige gemiddelde van de opeenvolgende ordinale waarden die daarvoor gegeven zouden moeten worden. De hoogste rangorde ( $n$  bij  $n$  alternatieven) wordt toegekend aan het alternatief met de hoogste score in geval van een batenkriterium en in geval van een kostenkriterium, aan het alternatief met de laagste score. Eventueel kan nog een correctie plaatsvinden voor de waarschijnlijkheid dat een plan ook werkelijk zal worden uitgevoerd dan wel voor de haalbaarheid van een alternatief op grond van andere overwegingen zoals toestemming van de overheid, externe financiering e.d. [31]. De rank-based expected value wordt als volgt bepaald:

$$V_i = p_i \sum_{j=1}^J w_j^r r_{ji} \quad (3.3)$$



waarbij:

$V_i$  = de waarde van alternatief  $i$  ( $i = 1, \dots, n$ )

$P_i$  = de kans dat alternatief  $i$  uitgevoerd wordt c.q. de haalbaarheid van alternatief  $i$

$w_j$  = de rangorde van het criterium  $j$  ( $j = 1, \dots, J$ )

$r_{ji}$  = de rangorde van alternatief  $i$  t.a.v. criterium  $j$

Aan de hand van de uitwerking van een cijfervoorbeeld zal deze methode geïllustreerd worden. Stel dat de gegevens van tabel 3.2 beschikbaar zijn:

kriteria	alternatieven			aard kriteria	eenheid
	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>		
K <sub>1</sub>	12	16	38	kosten	mln. f
K <sub>2</sub>	0,3	0,15	0,05	grondwaterdaling	m
K <sub>3</sub>	105	60	84	ruimtebeslag	ha

Tabel 3.2 Matrix  $P_1$  van kardinale effectwaarden

Op basis van matrix  $P_1$  kan de volgende ordinale matrix  $P_2$  opgesteld worden:

kriteria	alternatieven			gewichtenvectoren	
	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	w <sub>1</sub>	w <sub>2</sub>
K <sub>1</sub>	3	2	1	3	2
K <sub>2</sub>	1	2	3	2	3
K <sub>3</sub>	1	3	2	1	1

Tabel 3.3 Matrix  $P_2$  van ordinale effectwaarden en gewichtenvectoren  $w_1$  en  $w_2$

In het geval van de ongewogen variant ( $w_j = 1$  voor iedere  $j$ ) en zonder rekening te houden met uitvoeringskansen ( $p_i = 1$  voor iedere  $i$ ) worden de waarden  $V_i$  (verwachtingswaarden):

$V_1 = 3 + 1 + 1 = 5$	De preferentie-
$V_2 = 2 + 2 + 3 = 7$	volgorde is:
$V_3 = 1 + 3 + 2 = 6$	2 - 3 - 1

Als wel gebruik wordt gemaakt van gewichten, wordt de matrix van ordinale effecten voorvermenigvuldigd met de gewichtenvector:

- gewichtenvector  $w_1$ :

$V_1 = (3 \times 3) + (2 \times 1) + (1 \times 1) = 12$	De preferentie-
$V_2 = (3 \times 2) + (2 \times 2) + (1 \times 3) = 13$	volgorde is:
$V_3 = (3 \times 1) + (2 \times 3) + (1 \times 2) = 11$	2 - 1 - 3

- gewichtenvector  $w_2$ :

$V_1 = (2 \times 3) + (3 \times 1) + (1 \times 1) = 10$	De preferentie-
$V_2 = (2 \times 2) + (3 \times 2) + (1 \times 3) = 13$	volgorde is:
$V_3 = (2 \times 1) + (3 \times 3) + (1 \times 2) = 13$	$\begin{Bmatrix} 2 \\ 3 \end{Bmatrix} - 1$

Zoals het bovenstaande voorbeeld uitwijst, kunnen afhankelijk van de gekozen gewichtenvector verschillende rangorden van verwachtingswaarden resulteren. Als sprake is van gelijke verwachtingswaarden zoals bij toepassing van gewichtenvector  $w_2$ , wordt gesproken van "intransitiviteit": het is dan onmogelijk een preferentievolvergadering aan te geven. De verwachtingswaardemethode geeft in zo'n geval een zogenaamde zwakke ordening tussen de alternatieven aan. Dit fenomeen kan zich overigens ook bij andere MKE-methoden voordoen.

Een bezwaar van deze methode is, dat op geen enkele wijze rekening gehouden wordt met de omvang van de verschillen. Het is denkbaar dat alternatief  $A_1$  voor bijvoorbeeld twee criteria veel en veel aantrekkelijker is dan alternatief  $A_2$ , maar toch afgewezen wordt doordat  $A_2$  voor andere criteria steeds net iets beter scoort.



De toevoeging van een kansparameter  $p_i$  is in feite geen wezenlijke verbetering van de methode. De waarschijnlijkheid dat een projekt al of niet uitgevoerd zal worden, hangt af van de kenmerken van het projekt (dus de effektskores). Het waarschijnlijkheidselement zit in feite opgesloten in die skores, tenzij er meer beslissingskriteria, dan in de effektenmatrix staan vermeld, een doorslaggevende rol spelen. Is dit het geval, dan zouden deze criteria toegevoegd moeten worden aan de reeds onderkende criteria in plaats van een impliciete waardering te geven door middel van een kansparameter.

Mede op grond van bovengenoemde overwegingen zullen verwachtingswaardemethoden met grote voorzichtigheid moeten worden toegepast; de methode is een zeer rigide procedure, die de aanwezige informatie niet optimaal benut.

Er zijn veel varianten op de bovengeschetste rank-based expected value methode ontwikkeld. Zo kunnen naast gewichtenvectoren van ordinale waarden bij het bepalen van de prioriteitsvolgorde per criterium (de essentiële eigenschap van de verwachtingswaardemethode) ook gewichtenvectoren gebruikt worden die op een andere basis gefundeerd zijn, zoals op basis van de som = 100 of 1. Deze laatste variant wordt in wat volgt als SOM aangeduid (zie ook tabel 2.4). Voor de volledigheid wordt ook nog gewezen op de "ordinal comparison method" van Holmes [27], het skoremodel van Schimpeler en Grecco [30] en de nutsbenadering van Stanley [32].

### 3.4 Permutatiemethoden

Permutatiemethoden gaan uit van het principe dat iedere mogelijke rangschikking van keuzemogelijkheden getest kan worden op de mate van overeenstemming met de effectiviteitsskores en criteriumgewichten. Door middel van een simultane behandeling van preferentieskores en ordinale uitkomsten via een reeks van permutaties kan de meest voorkomende rangorde van de alternatieven worden afgeleid.

In Nederland is de methode van Paelinck [29] het meest bekend. Deze methode stelt de volgende voorwaarden aan het gegevensmateriaal:

1. ieder alternatief kan gewaardeerd worden ten aanzien van een aantal onderkende criteria;

2. de waarden van de effecten met betrekking tot deze criteria kunnen worden gemeten in ordinale schalen, die niet dezelfde behoeven te zijn; en
3. voor de waardering van de belangrijkheid van elk criterium in de uiteindelijke beslissing worden geordende (ordinale) gewichten gebruikt.

Deze methode zal worden uitgewerkt aan de hand van het voorbeeld in paragraaf 3.3. Hiertoe wordt de matrix  $P_2$  uit tabel 3.3 gebruikt.

criteria	alternatieven			gewichtenvectoren	
	$A_1$	$A_2$	$A_3$	$w_1$	$w_2$
$K_1$	3	2	1	3	2
$K_2$	1	2	3	2	3
$K_3$	1	3	2	1	1

Tabel 3.3 Matrix  $P_2$  van ordinale effectwaarden en gewichtenvectoren  $w_1$  en  $w_2$

Bij een probleem met  $n$  alternatieven zijn er indien afgezien wordt van intran-  
sitiviteit (zie paragraaf 3.3)  $n!$  mogelijkheden om de alternatieven te rang-  
schikken. In dit geval zijn er  $3! = 6$  rangorden:

$$R_1 : 1 > 2 > 3$$

$$R_2 : 1 > 3 > 2$$

$$R_3 : 2 > 1 > 3$$

$$R_4 : 2 > 3 > 1$$

$$R_5 : 3 > 1 > 2$$

$$R_6 : 3 > 2 > 1$$

Hierbij betekent  $1 > 2$  dat 1 beter is dan 2. Met behulp van matrix  $P_2$  kunnen nu alle mogelijke rangorden onderling getest worden. Daartoe wordt aan iedere rangorde een waardering toegekend op basis van de vergelijking van partiële rangorden en de gegevens in matrix  $P_2$ . Is de partiële rangorde ( $1 > 2$ ) in overeenstemming met de ordinale skorerangorde in matrix  $P_2$ , dan is de waarde-



ring +1, is dit niet het geval dan wordt aan de partiële rangorde -1 toegekend. De waardering voor de totale rangorde is de som van de waarderingen van de partiële rangorden. Voor bijvoorbeeld rangorde  $R_1$  ten aanzien van criterium  $K_3$  houdt dit in:

- 1 > 2 niet in overeenstemming met de gegevens uit  $P_2$ : waarde -1
- 1 > 3 niet in overeenstemming met de gegevens uit  $P_2$ : waarde -1
- 2 > 3 wel in overeenstemming met de gegevens uit  $P_2$ : waarde +1 +  
 waardering voor  $R_1$  t.a.v.  $K_3$ : waarde -1

Deze procedure kan voor alle rangorden en alle criteria uitgevoerd worden. Eventueel wordt de waarde 0 toegekend voor gelijkskorende alternatieven (zie onder). De resultaten staan vermeld in tabel 3.4:

rangorden	criteria		
	$K_1$	$K_2$	$K_3$
$R_1$	+3	-3	-1
$R_2$	+1	-1	-3
$R_3$	+1	-1	+1
$R_4$	-1	+1	+3
$R_5$	-1	+1	-1
$R_6$	-3	+3	+1

Tabel 3.4 Matrix  $P_3$  van rangordeskores voor  $P_2$  en  $R_1$  t/m  $R_6$

Vervolgens kan bepaald worden welke rangorde het meest robuust is bij toepassing van de verschillende gewichtenvectoren ( $w_1$  en  $w_2$ ; al dan niet genormaliseerd). Voor gewichtenvector  $w_1=(3,2,1)$  betekent dit via een matrixvermenigvuldiging:

$$P_3 \cdot w_1 = \begin{bmatrix} (+3 \times 3) + (-3 \times 2) + (-1 \times 1) \\ (+1 \times 3) + (-1 \times 2) + (-3 \times 1) \\ (+1 \times 3) + (-1 \times 2) + (+1 \times 1) \\ (-1 \times 3) + (+1 \times 2) + (+3 \times 1) \\ (-1 \times 3) + (+1 \times 2) + (-1 \times 1) \\ (-3 \times 3) + (+3 \times 2) + (+1 \times 1) \end{bmatrix} = \begin{bmatrix} +2 \\ -2 \\ +2 \\ +2 \\ -2 \\ -2 \end{bmatrix}$$

Bij gewichtenvector  $w_1$  hebben de rangorden  $R_1$ ,  $R_3$  en  $R_4$  de hoogste skore. Een eenduidige indikatie welke rangorde de beste is, is op basis van deze methode voor dit voorbeeld niet te geven.

Een moeilijkheid ontstaat als alternatieven ten aanzien van een bepaald criterium gelijk gewaardeerd worden. De rangorden kunnen namelijk dan geen maximale waardering krijgen, doordat bij een partiële gelijkheid geen +1 of -1 geskoord wordt, maar 0. In dit geval kan een korrektie toegepast worden door de skores van de rangorden bij het betreffende criterium  $K_j$  te vermenigvuldigen met een faktor  $\frac{n(n-1)}{2 \cdot \max_j}$ , waarbij  $n$  het aantal alternatieven is en  $\max_j$  de maximale skore bij criterium  $K_j$  ( $\frac{n(n-1)}{2}$  is het aantal gemaakte vergelijkingen).

Op de methode van Paelinck zijn nog meer aanpassingen mogelijk, zoals de mogelijkheid rekening te houden met het feit, dat meer informatie beschikbaar is dan alleen een rangorde van de alternatieven. Zo kan rekening gehouden worden met intervalskores van de onderscheiden alternatieven voor een bepaald criterium (men weet niet alleen dat alternatief  $i$  beter is dan alternatief  $i'$ , maar ook hoeveel beter). De methode van Paelinck kent ook een aanpassing voor bewerking van gekwantificeerde informatie (vgl. de methode van Jacquet-Lagrez, [28]). Ook als alleen waarschijnlijkheden bekend waarmee bepaalde rangschikkingen juist zijn, kan de methode van Paelinck toegepast worden. Als resultaat krijgt men dan de kansen dat de verschillende alternatieven het beste zullen zijn.

Een nadeel van deze methode is echter dat in het geval veel alternatieven beoordeeld moeten worden, het aantal permutaties snel zeer groot wordt: immers, dit aantal is gelijk aan  $n!$ . Dit kan rekentechnisch voor de nodige problemen zorgen. Hoewel Paelinck een manier gevonden heeft om het aantal rangorden te reduceren tot een klein percentage van het oorspronkelijke aantal, blijft het aantal nog steeds zeer aanzienlijk.



### 3.5 Eigenwaardemethoden

Een methode, die in Nederland bekendheid heeft verworven door haar toepassing ten behoeve van de waterhuishouding in Gelderland, is de methode van Saaty [21]. Ook bij het voormalige Rijksinstituut voor Drinkwatervoorziening (RID) te Leidschendam is ervaring met deze methode opgedaan.

De methode van Saaty is gebaseerd op enerzijds een paarsgewijze vergelijking van de verschillende alternatieven per criterium afzonderlijk en anderzijds op een paarsgewijze vergelijking van de criteria onderling. Bij het bepalen van de graad van dominantie hanteert Saaty een schaal van 1 tot 9, waarbij de cijfers als volgt zijn gedefinieerd:

1. even belangrijk
3. iets belangrijker
5. veel belangrijker
7. duidelijk veel belangrijker
9. absoluut veel belangrijker.

Even waarderingscijfers worden gebruikt om tussenliggende oordelen uit te spreken. De waarderingscijfers worden per criterium  $K_j$  in een matrix  $A_j$  ingevuld, waarbij er voor zorg wordt gedragen dat deze matrix reciprook is (dat wil zeggen dat  $A_j(i,k) \cdot A_j(k,i) = 1$  ( $i, k = 1, \dots, n$ )). Een voorbeeld is gegeven in tabel 3.5.

alternatieven	alternatieven					
	1	2	3	.	.	n
1	1	2	1/5	.	.	.
2	1/2	1	1/5	.	.	.
3	5	5	1	.	.	.
.	.	.	.	.	.	.
I	.	.	.	.	.	.

Tabel 3.5 Voorbeeld van een alternatievenvergelijkingsmatrix voor criterium  $K_j$

Onder de veronderstelling, dat voor het criterium in kwestie de elementen van de matrix de verhouding vormen van achterliggende skores van de alternatieven op dit criterium (zo is meteen de reciprociteit van de matrix gerechtvaardigd), verenigd in een vector  $v_j$  per criterium  $K_j$ , blijkt eenvoudig, dat

$$A_j v_j = n v_j \quad (j=1, \dots, J) \quad (3.4)$$

De oplossing van  $v_j$  uit (3.4) met bekende  $A_j$  staat bekend als een eigenwaardeprobleem; hier is  $n$  een eigenwaarde en  $v_j$  de bijbehorende eigenvector. Normaal gesproken zal de invulling van de matrix niet consistent, dat wil zeggen conform de achterliggende effectenscores, plaatsvinden. Als gevolg daarvan zal de (grootste) gevonden eigenwaarde niet gelijk zijn aan  $n$  en de daarbij horende eigenvector verschillen van  $v_j$ . Deze laatste is echter onbekend, reden waarom de bij de grootste eigenwaarde van  $A_j$  behorende eigenvector als de uit  $A_j$  resulterende vector van (relatieve) skores op criterium  $K_j$  wordt beschouwd; deze vectoren vormen dan de effectenmatrix voor het onderhavige probleem.

Dezelfde procedure kan worden gevolgd met betrekking tot de gewichtentoekenning: uitgaande van een reciproke matrix, die het relatieve belang bevat op eenzelfde schaal met waarden van 1 tot 9 van alle criteria ten opzichte van elkaar, wordt tot een eigenwaardeprobleem gekomen. Idealiter is de grootste eigenwaarde daarvan gelijk aan  $J$  (het aantal criteria) en de eigenvector daarbij is de achterliggende gewichtenvector.

Het rekenvoorschrift, dat bij het RID wordt gehanteerd, komt neer op het berekenen van een vector waarvan elk element het meetkundige gemiddelde is van de elementen van de overeenkomstige rij van de paarsgewijze vergelijkingsmatrix. Het voordeel van de RID-methode boven de oorspronkelijke Saaty-methode is de mogelijkheid statistische beschouwingen te gebruiken bij het interpreteren van de resultaten. Daarnaast kost zij wat minder rekenwerk en kan zij ook worden gebruikt als de paarsgewijze vergelijkingsmatrix niet volledig is ingevuld, zij het dat het rekenvoorschrift dan wat minder eenvoudig in woorden is uit te drukken. Als alle paarsgewijze vergelijkingen exakt overeenkomen met de verhoudingen van de skores, leveren de beide varianten exakt dezelfde resultaten. In praktische situaties zijn de numerieke verschillen in de uitkomsten van beide methoden te verwaarlozen.



In bovenstaande is feitelijk sprake van een combinatie van preferentiemethode en standaardisatiemethode (zie subparagraaf 4.2.1): uit kwalitatieve gegevens (veel belangrijker dan, enzovoort) worden door oplossing van  $(J+1)$  eigenwaardeproblemen een effektenmatrix en een gewichtenvector opgesteld. Deze kunnen als basismateriaal voor de toepassing van verwachtingswaardemethoden worden gebruikt. De methode Saaty kan ook als voorbewerking op de gegevens worden toegepast, waarna een der andere evaluatiemethoden de preferenties oplevert.

### 3.6 Konkordantiemethoden

Bij konkordantiemethoden wordt getracht door middel van een paarsgewijze vergelijking van de alternatieven aan te geven welk alternatief dominant is boven alle andere alternatieven. Konkordantiemethoden zijn in principe gebaseerd op de door Benayoun e.a. [26] ontwikkelde Electre-methode (Elimination et choix traduisant la réalité).

Het toepassen van konkordantiemethoden valt in de volgende stappen uiteen:

- a. Opstelling van de effektenmatrix  $P$ , waarbij de elementen  $p(j,i)$  de uitkomst van alternatief  $i$  ( $i = 1, \dots, n$ ) voor het  $j$ -de criterium ( $j = 1, \dots, J$ ) voorstellen. De elementen van deze matrix (d.w.z. de skores van de alternatieven op de criteria) dienen per criterium gestandaardiseerd te worden, zodat alle rijvectoren van  $P$  onderling vergelijkbaar worden.
- b. De gewichtenvector van de criteria wordt opgesteld; deze is een vector  $w$  met elementen  $w(j)$ , waarbij  $w(j)$  het relatieve prioriteitsgewicht van criterium  $j$  voorstelt.
- c. Bepaling van de konkordantie- en diskordantieverzameling. De konkordantieverzameling wordt bepaald op grond van een paarsgewijze vergelijking van elk van de planalternatieven. Voor elk paar alternatieven kunnen namelijk de criteria worden onderverdeeld in twee deelverzamelingen. De eerste deelverzameling, aangeduid als  $k(i,i')$ , bestaat uit alle criteria, waarvoor alternatief  $i$  geprefereerd wordt boven of een gelijke waardering krijgt als alternatief  $i'$ . Deze deelverzameling wordt de konkordantieverzameling van  $i$  ten opzichte van  $i'$  genoemd. Voor elk paar alternatieven kan een dergelijke

verzameling bepaald worden. De hieraan komplementaire deelverzameling wordt de diskordantieverzameling  $d(i,i')$  genoemd. Geformaliseerd is de konkordantieverzameling als volgt gedefinieerd:

$$k(i,i') = \{j:p(j,i) > p(j,i')\} \quad (3.5)$$

Op analoge wijze wordt de diskordantieverzameling  $d(i,i')$  opgesteld:

$$d(i,i') = \{j:p(j,i) < p(j,i')\} \quad (3.6)$$

- d. Vervolgens wordt de konkordantie-index  $k_{ii'}$  bepaald. De konkordantie-index is een maat die behoort bij de konkordantieverzameling. Deze index geeft voor elk paar alternatieven  $i$  en  $i'$  de relatieve prioriteit van  $i$  ten opzichte van  $i'$  aan; zij wordt berekend door optelling van de gewichten  $w(j)$  van alle criteria  $j$  die in de konkordantieverzameling  $k(i,i')$  zitten.
- e. Berekening van de diskordantie-index  $d_{ii'}$ . Deze geeft van elk paar alternatieven  $i$  en  $i'$  de relatieve mate weer waarin binnen de diskordantieverzameling alternatief  $i$  een slechtere uitkomst oplevert dan  $i'$ . Deze index kan worden berekend als de som van het absolute verschil tussen de effecten van matrix  $P$  voor elk paar alternatieven, voorzover de bijbehorende criteria in de diskordantieverzameling  $d(i,i')$  zitten, genormeerd door deling door het grootste absolute verschil:

$$d_{ii'} = \sum_{j \in d(i,i')} \frac{p(j,i) - p(j,i')}{d_{\max}(j)} \quad (3.7)$$

waarbij  $d_{\max}(j)$  gelijk is aan het maximale absolute verschil tussen de scores der alternatieven voor criterium  $j$ :

$$d_{\max}(j) = \max_{1 < i, i' < I} |p(j,i) - p(j,i')| \quad j \in d(i,i') \quad (3.8)$$

Bij sommige toepassingen van de diskordantie-index wordt gebruik gemaakt van een gewogen index, d.w.z. dat de relatieve verschillen vermenigvuldigd worden met het bijbehorende criteriumgewicht:



$$d_{ii'}^* = w(j) \cdot d_{ii'} \quad (3.9)$$

Op grond van het bovenstaande kan een diskordantiematrix van diskordantieindices voor alle paren projecten ingevuld worden.

f. Tenslotte dienen de konkordantie- en diskordantiematrices onderzocht te worden, opdat het alternatief dat ten opzichte van de andere alternatieven een maximale netto konkordantie en een minimale diskordantie bezit, getraceerd kan worden. De netto-konkordantie-dominantie-index  $kd(i)$  wordt als volgt berekend:

$$kd(i) = \sum_{\substack{i'=1 \\ i' \neq i}}^I k_{ii'} - \sum_{\substack{i'=1 \\ i' \neq i}}^I k_{i'i} \quad (3.10)$$

In woorden:  $kd(i)$  is de som van de relatieve preferenties van alternatief  $i$  ten opzichte van alternatief  $i'$  verminderd met de som van de relatieve preferenties van  $i'$  ten opzichte van  $i$ . M.a.w. de netto-konkordantie-dominantie-index van alternatief  $i$  is het verschil van rijtotaal en kolomtotaal van de konkordantiematrix. Op analoge wijze wordt de netto-diskordantie-dominantie-index  $dd(i)$  bepaald:

$$dd(i) = \sum_{\substack{i'=1 \\ i' \neq i}}^I d_{ii'} - \sum_{\substack{i'=1 \\ i' \neq i}}^I d_{i'i} \quad (3.11)$$

Het zal duidelijk zijn, dat een alternatief  $i$  het gunstigst is, als  $kd(i)$  het grootst is en  $dd(i)$  het kleinst. Als er geen eenduidige oplossing uit deze laatste vraagstelling komt, kan deze procedure toch gebruikt worden om minder gunstige plannen te elimineren. In de originele Electre-methode wordt gebruik gemaakt van drempelwaarden voor de konkordantie- en diskordantieindices. De achterliggende idee is, dat een bepaald alternatief een grotere kans maakt om geëlimineerd te worden, als zijn konkordantieindices ten opzichte van andere alternatieven relatief groot zijn. De konkordantiemethoden die geen gebruik van drempelwaarden maken worden ook wel konkordantie-rangschikkingsmethoden genoemd. Deze methoden leveren als eindresultaat een (zwakke) ordening van de alternatieven op, in tegenstelling tot konkordantie-eliminatiemethoden (zoals Electre) met behulp waarvan een reductie wordt aangebracht van het aantal

alternatieven waaruit gekozen kan worden. Voor een uitgebreid overzicht van een groot aantal varianten op het bovengeschetste raamwerk van konkordantie-methoden wordt verwezen naar [1] en [2]. Deze varianten hebben voornamelijk betrekking op aanpassing van de basismethode aan de kenmerken van de effectenmatrix en de gewichtenvector (zoals ten behoeve van bewerking van kwalitatieve gegevens).

### 3.7 Meerdimensionale schalingsmethoden

Bij het analyseren van meerdimensionale gegevens (d.w.z. één alternatief scoort op meer dan één criterium) kan men zich vaak niet aan de indruk onttrekken dat er tussen de verschillende grootheden een samenhang bestaat. In sommige gevallen is dat direkt te doorzien: als alternatieven worden geëvalueerd naar zowel de netto kontante waarde van hun kosten of opbrengsten, doch tevens naar investeringslasten en jaarlijkse kosten, dan is het duidelijk dat deze drie criteria teruggebracht kunnen worden tot twee.

In complexere situaties is dit echter niet direkt duidelijk. Bijvoorbeeld is er onduidelijke samenhang tussen een criterium als daling grondwaterstand en opbrengstverliezen landbouw uitgedrukt in inkomen van de boer en schade aan het milieu. Het kan dan wenselijk zijn middels wiskundige technieken na te gaan of het aantal verschillende criteria kan worden teruggebracht tot een kleiner aantal, dat wel onafhankelijk is. Men moet immers bij de analyse vermijden dat men bepaalde aspecten teveel gewicht geeft alleen al door het aantal criteria. Daarnaast wordt de analyse vereenvoudigd als er minder criteria simultaan moeten worden bekeken.

In de statistiek zijn diverse methodieken ontwikkeld om kardinale gegevens in deze zin te onderzoeken. Zij worden o.a. toegepast bij oudheidkundig onderzoek en in de medische wetenschap bij het stellen van diagnose en kunnen zonder veel moeite worden gebruikt bij kardinale multikriteria evaluaties. Een simpele toepassing bij de vergelijking tussen verschillende evaluatie- en standaardisatiemethoden is opgenomen in paragraaf 6.3 (zie figuur 6.1).

Voor ordinale gegevens zijn zij echter niet zonder meer geschikt. Daarom zijn recentelijk rekenmethodieken ontwikkeld, die de ordinale gegevens transforme-



ren naar kardinale gegevens van lagere dimensie. Daarbij moeten de kardinale gegevens zo goed mogelijk in overeenstemming blijven met de oorspronkelijke ordinale gegevens. In het bijzonder moet er op worden gelet dat de rangorde van de kardinale gegevens gelijk is aan die van de ordinale uitgangsgegevens.

Door dr. J.H. Voogd (TH-Delft) is een ordinale geometrische evaluatiemethode (OGEM) ontwikkeld, die zowel voor de alternatieven als voor de criteria gebruik maakt van een ordinale schalingsmethode zoals hierboven geschetst. Zowel de alternatieven als de criteria worden als punten in een meetkundige ruimte met zo laag mogelijke dimensie afgebeeld. De punten worden zo in deze ruimten geplaatst, dat hun afstanden zo goed mogelijk overeenkomen met hun rangorden. Daarbij wordt de uiteindelijke totaalrangorde van de alternatieven bepaald op grond van hun afstand tot een hypothetisch "ideaalpunt": het ideaalpunt is een hypothetische keuzemogelijkheid, waaraan boven elk ander bestaand alternatief de voorkeur wordt gegeven. Als de criteria niet tegenstrijdig zouden zijn, zou dit ideaalpunt samenvallen met de criteriumpunten zelf. Dit komt echter zelden voor. Het ideaalpunt wordt daarom zo geplaatst, dat het dichterbij criteriumpunten met een groot gewicht dan bij criteriumpunten met een klein gewicht.

De numerieke oplossing wordt verkregen middels een complex optimalisatieproces. Dit vereist voor alle toepassingen ruime computerfaciliteiten, terwijl interpretatie en gevoeligheidsanalyse niet eenvoudig zijn. Voor grote aantallen alternatieven is de OGEM daarom minder geschikt. Er wordt echter wel op een volledig verantwoorde wijze omgegaan met ordinale gegevens: bij andere MKE-methoden worden ordinale gegevens vaak (verdekt) op heuristische wijze omgezet in kardinale gegevens.

Bij de Ordinaal Geometrische Evaluatiemethode wordt voor het gewicht van de afzonderlijke criteria kardinale informatie gebruikt. Indien ook voor de criteria slechts ordinale informatie beschikbaar is, dient nog te worden nagegaan wat de invloed van het toekennen van gewichten aan de criteria is. Daarvoor gebruikt Voogd de "extreme punten" benadering. Wat dit inhoudt, kan het best aan de hand van een voorbeeld duidelijk worden gemaakt. Stel dat er drie criteria zijn; bekend is dat voor de gewichten geldt:  $w(1) > w(2) > w(3)$ , d.w.z. ze zijn geordend in afnemende belangrijkheid. Worden  $w(1)$ ,  $w(2)$  en  $w(3)$  langs

drie orthogonale assen uitgezet, dan liggen alle mogelijke gewichtskombinaties binnen een viervlak met als hoekpunten de oorsprong en de volgende drie extreme gewichtspunten (merk op, dat voor de som van de gewichten 1 wordt genomen):

1.  $w(1) = 1$  ,  $w(2) = 0$  ,  $w(3) = 0$
2.  $w(1) = 1/2$  ,  $w(2) = 1/2$  ,  $w(3) = 0$
3.  $w(1) = 1/3$  ,  $w(2) = 1/3$  ,  $w(3) = 1/3$

Voor deze drie extreme situaties wordt dan de berekening uitgevoerd, en de resultaten worden simultaan geïnterpreteerd. Het is echter niet zeker of dit voldoende is; evenals bij de methode Paelinck kunnen naast de extreme situaties beter ook alle tussenverdelingen worden genomen, d.w.z.

- 1.2.  $w(1) = 3/4$  ,  $w(2) = 1/4$  ,  $w(3) = 0$
- 2.3.  $w(1) = 5/12$  ,  $w(2) = 5/12$  ,  $w(3) = 2/12$
- 1.3.  $w(1) = 4/6$  ,  $w(2) = 1/6$  ,  $w(3) = 1/6$
- 1.2.3.  $w(1) = 11/18$  ,  $w(2) = 5/18$  ,  $w(3) = 2/18$

Bovenstaande vier tussenverdelingen zijn de exacte middens, die kunnen worden afgeleid uit de gebruikte extreme gewichtspunten.

Voor de wiskundige formulering wordt verwezen naar [2].

### 3.8 MKE-methoden en soort gegevens

De hierboven beschreven MKE-methoden gaan alle van de veronderstelling uit dat de effectenmatrix hetzij kardinale gegevens hetzij ordinale gegevens bevat. De meeste methoden zijn voor beide gevallen geschikt gemaakt. Specifiek ordinale methoden zoals permutatiemethoden kunnen ook in geval dat een kardinale effectenmatrix ter beschikking staat toegepast worden door middel van een vertaling van de kardinale gegevens (vergelijk tabel 2.1) naar ordinale gegevens (vergelijk tabel 2.2). Daarbij moet bedacht worden dat bij deze vertaling een belangrijk deel van de informatie verloren gaat. Het verdient dan ook de voorkeur om indien kardinale gegevens beschikbaar zijn gebruik te maken van kardinale MKE-methoden.



In de praktijk blijkt echter niet altijd aan genoemde vooronderstellingen voldaan te zijn. Vaak worden effektenmatrices aangetroffen die zowel kwantitatieve skores bevatten als kwalitatieve schattingen. Dit probleem kan overvragen worden door kardinale gegevens in ordinale gegevens om te zetten (zie ook paragraaf 4.1). Dit betekent wel dat de voor sommige criteria wel beschikbare kwantitatieve informatie in de evaluatie slechts gedeeltelijk gebruikt wordt, namelijk alleen de rangorde-eigenschappen van de alternatieven.

In de Evamix-benadering [12, 44] wordt expliciet rekening gehouden met de mogelijkheid dat de effektenmatrix zowel kardinale als ordinale informatie bevat. De effektenmatrix wordt daarbij in twee deelmatrices gesplitst: een kardinale effektenmatrix en een ordinale effektenmatrix. Voor iedere matrix afzonderlijk wordt met behulp van konkordantie-analyse de dominantie-skore per paar alternatieven bepaald. Hierbij wordt gebruik gemaakt van een gewichten-vector (zie paragraaf 3.6). Vervolgens worden de dominantie-indices gestandaardiseerd bijvoorbeeld door de afzonderlijke skores te delen door de som van alle skores (in absoluut-termen) van een dominantiematrix (respektievelijk de "kardinale" dominantiematrix en de "ordinale" dominantiematrix). Op deze wijze worden twee gestandaardiseerde dominantiematrices verkregen die onderling vergeleken kunnen worden door middel van weging met de kriteriagewichten. De totale dominantie-indices worden berekend door de gestandaardiseerde "ordinale" dominantieskores te vermenigvuldigen met de som van de gewichten van de criteria waarvoor slechts ordinale informatie beschikbaar is en door de gestandaardiseerde "kardinale" dominantieskores te vermenigvuldigen met de som van de gewichten van de kardinale criteria. Uit deze totale konkordantiematrix kan voor ieder alternatief de netto-konkordantie-dominantie-index  $kd(i)$  bepaald worden (zie formule (3.10)).

#### 4 Standaardisatie

In dit hoofdstuk wordt aandacht besteed aan de standaardisatie van met name kardinale effektscores. Paragraaf 4.1 bevat een inleiding omtrent doel en werkwijze van standaardisatie. Paragraaf 4.2 behandelt de in het onderzoek betrokken standaardisatiemethoden; formules worden gegeven en aan de hand van figuren en voorbeelden worden hun eigenschappen toegelicht. Enig materiaal met betrekking tot het gebruik van externe extremen is te vinden in paragraaf 4.3, waarna paragraaf 4.4 een vergelijking presenteert tussen verschillende standaardisatiemethoden aan de hand van het IODZH-voorbeeld van hoofdstuk 2.

##### 4.1 Algemeen

In het kort is het doel van standaardisatie het onderling vergelijkbaar maken van scores die aan verschillende effecten worden toegekend. Door standaardisatie van scores in de effectenmatrix wordt een zogenaamde effektiviteitenmatrix verkregen.

Aan de effektiviteitenmatrix moet de eis gesteld worden dat voor een bepaalde effektiviteitsscore ( $y$ ) voor alle criteria ( $k_j$ ) eenzelfde interpretatie geldt:

$$y \text{ op } k_1 \equiv y \text{ op } k_2 \quad (4.1)$$

Bij de meeste toepassingen van MKE (zoals ook in dit rapport eenvoudigheds-halve gebeurd is) wordt gebruik gemaakt van één en dezelfde standaardisatiemethode voor de gehele effectenmatrix. Dit doet veelal geen recht aan de specifieke aard van de gegevens.

Daarom moet de effectenmatrix in principe op de volgende wijze gestandaardiseerd worden:

1. kies op basis van de aard van de gegevens de meest geëigende standaardisatiemethode per criterium, waarbij geldt dat voor alle criteria "hoogst = best" of andersom; en
2. normeer de rijen van de effektiviteitenmatrix, zodanig dat (4.1) geldt.



Het is voor een goed begrip van belang drie zaken gescheiden te houden:

- (i) Bij het toepassen van ordinale MKE-methoden moet men beschikken over een volledig ordinale effektenmatrix. Doorgaans echter is het uitgangsmateriaal een combinatie van kardinale en ordinale gegevens. Indien geen gebruik gemaakt wordt van methoden die expliciet met het voorkomen van een gemengde effektenmatrix rekening houden (zoals in paragraaf 3.8), moeten de kardinale gegevens vertaald worden naar ordinale gegevens. Hoewel vaak niet zo aangeduid, is bij het omzetten van kardinale in ordinale gegevens sprake van een standaardisatiestap. Deze stap houdt in: het vertalen in een rangorde van kardinale (cijfermatige) skores of ook wel het indelen van zulke skores in categorieën als "goed", "redelijk" en "slecht".
- (ii) Met standaardisatie binnen een criterium wordt in het vervolg bedoeld het herschalen van de oorspronkelijke (kardinale) effektskores naar een tevoren gespecificeerd interval, waarbij tevens een interpretatie van deze skores plaatsvindt door verschillen in effektskores te representeren in grote of kleine verschillen in effectiviteitskores. Een goed voorbeeld van de standaardisatie binnen een criterium is het geven van proefwerkcijfers op grond van het aantal fouten in een vertaling: het interval loopt hier van 1 tot 10, waarbij 10 de gunstigste skore is en de interpretatie enerzijds plaatsvindt door een grens tussen 5 en 6 vast te leggen en anderzijds door eventueel de cijfers niet recht evenredig met het aantal fouten te doen zijn (bijvoorbeeld 1 fout  $\rightarrow$  9, 3 fouten  $\rightarrow$  8, 6 fouten  $\rightarrow$  7....24 (of meer!) fouten  $\rightarrow$  1).
- (iii) Ook tussen criteria kan standaardisatie plaatsvinden; hierbij worden voor alle criteria de effektskores herschaald naar éénzelfde interval: hun waarden worden dan onderling vergelijkbaar. Dit proces kan goed worden vergeleken met het dimensieloos maken van grootheden (of in dezelfde eenheid uitdrukken van waarden van één grootheid). Deze vorm van standaardisatie is wat doorgaans met standaardisatie wordt bedoeld.

Met nadruk wordt gewezen op het feit, dat de standaardisatie binnen een criterium tevens een standaardisatie tussen criteria inhoudt, terwijl de genoemde interpretatie van effektskores doorslaggevend kan zijn voor het uiteindelijke resultaat van de MKE (zie de bespreking van dit aspekt in hoofdstuk 5).

## 4.2 Methoden

In deze paragraaf komen de gebruikte standaardisatiemethoden aan de orde. Om de methoden te kunnen identificeren heeft elke methode een naam gekregen, die kan verschillen van benamingen zoals die in de literatuur voorkomen: van een eenheid in naamgeving is in de literatuur klaarblijkelijk geen sprake.

Achtereenvolgens gaat subparagraaf 4.2.1 in op methoden om kardinale gegevens om te zetten in ordinale (RANGORDE, KLASSE, SAATY, QUALI), subparagraaf 4.2.2 op methoden, waarbij een eenvoudig verband bestaat tussen de effektscores en de (kardinale) effectiviteitscores (TOTAAL, MAX, MAXMIN, MIN), subparagraaf 4.2.3 geeft methoden, waarbij een keuzeparameter aanwezig is (NUTS, INDIF) en subparagraaf 4.2.4 behandelt de meer geavanceerde methoden (ENTROPIE, WATER-LAB).

Een poging alle in de literatuur voorkomende standaardisatiemethoden te behandelen is niet ondernomen. De gegeven lijst kan zonder veel moeite worden uitgebreid met andere bekende of onbekende methoden.

### 4.2.1 Niet-kardinale gegevens

#### Rangorde (RANGORDE)

De eenvoudigste en meest voor de hand liggende methode om kardinale gegevens om te vormen tot ordinale scores is het geven van rangnummers. Hierbij kan bijvoorbeeld het gunstigste alternatief voor een bepaald criterium rangorde 1 krijgen, het daaropvolgende rangorde 2, enzovoort. Indien gewenst kan de volgorde ook omgekeerd worden (vergelijk tabel 2.2). Een moeilijkheid ontstaat als meer dan één alternatief eenzelfde score bezit voor een bepaald criterium: het toegekende rangnummer kan dan het kleinste of het grootste (tabel 2.2) in aanmerking komende rangnummer worden ofwel kan het gemiddelde rangnummer worden toegekend. Hierbij zij op het feit gewezen, dat sommige evaluatiemethoden het effect van deze keuzemogelijkheid in de uiteindelijke preferentievolvergadering van de alternatieven tot uitdrukking laat komen!



Tabel 2.2 is een voorbeeld van een rangordestandaardisatie, hier toegepast op de IODZH-effektenmatrix van tabel 2.1; voor alle criteria is het principe toegepast, dat de beste skore het hoogste rangnummer krijgt (in dit geval 24).

#### Klassenindeling (KLASSE)

Doordat bij RANGORDE slechts naar de volgorde van de verschillende skores op een criterium wordt gekeken wordt in het geheel niet met de grootte van skoreverschillen rekening gehouden. Dit kan worden geïllustreerd aan de hand van de skores voor criterium 2 in tabel 2.1. Een vluchtige beschouwing leert reeds, dat de skoreverschillen minimaal zijn op een paar waarden na (voor de alternatieven 10 tot en met 15). Een gevolg daarvan is dat het verschil tussen 153 (alternatief 9) en 156 (alternatief 13) in tabel 2.1 slechts één rangnummer verschil in tabel 2.2 oplevert, terwijl 150 (alternatief 22) en 151 (alternatief 23) maar liefst vijf plaatsen in de rangorde uitmaken.

Een manier om hier wat aan te doen is het werken met klassen van effektskores; hierbij worden skores, die dicht bij elkaar liggen in eenzelfde klasse ondergebracht ter onderscheiding van (groepen) skores waartoe grotere verschillen gelden. Dit nu is gedaan bij de samenstelling van tabel 2.3, waarin slechts de klassenummers 1 tot en met 5 voorkomen (die niet steeds alle gebruikt zijn). Alternatieven 22 en 23 (en nog maar liefst 12 andere) komen met betrekking tot criterium 2 samen in klasse 4. De grens tussen klassen 2 en 3 is op 159 gesteld, waardoor alternatieven 9 en 13 beide in klasse 3 terechtkomen; een grens van 155 had betekend dat het relatief grote verschil tussen alternatieven 9 en 13 tot uiting was gekomen in een verschil in klassenindeling.

Daarmee is meteen duidelijk gemaakt dat door de klassenindeling een subjektief element is toegevoegd: er wordt gestandaardiseerd binnen een criterium (zie de opmerking in paragraaf 4.1) op grond van grenzen tussen klassen die arbitrair worden getrokken. Ook het aantal klassen is arbitrair: als één uiterste kan gelden de indeling in twee klassen (of zelfs geen indeling) en als ander uiterste een indeling waarbij elke mogelijke schaalwaarde een eigen klasse vormt en waardoor in feite een kardinale effectiviteitenmatrix wordt verkregen. Derhalve kan KLASSE als een tussenvorm tussen puur kardinale en puur ordinale standaardisatiemethoden worden beschouwd.

Een algemeen punt met betrekking tot ordinale (standaardisatie)methoden is nog dat een informatiereductie plaatsvindt: de informatie immers, die besloten ligt in de skoreverschillen, is niet in de rangorde terug te vinden en evenmin volledig uitgedrukt in de klassenindeling. Een voorbeeld van dit laatste is het rangnummer- en klasseverschil van 1 voor de alternatieven 23 en 24 met betrekking tot criterium 4 ondanks het relatief tot de andere skoreverschillen erg grote verschil tussen 0,44 en 0,54.

### SAATY

De methode SAATY is een evaluatiemethode (zie de bespreking als zodanig in paragraaf 3.5), waar op twee plaatsen gesproken kan worden van standaardisatie. Ten eerste vereist de methode de beschikbaarheid van waarderingsmatrices (per criterium), waarin slechts een beperkt aantal schaalwaarden mag voorkomen. Deze schaalwaarden lopen van 1 (even gunstig/belangrijk) tot 9 (extreem veel gunstiger/belangrijker); de omgekeerden worden gebruikt in het omgekeerde geval: 5 (veel gunstiger) wordt  $\frac{1}{5}$  (veel ongunstiger). Ook is het mogelijk slechts de schaalwaarden 1 en 3 te gebruiken, daarmee alleen onderscheid makend tussen even gunstige en ongelijkwaardige alternatieven. In feite gaat het hier om een tussenvorm tussen een ordinale en een intervalschaal.

Ten tweede worden tijdens de berekeningen, die bij de methode SAATY toe worden gepast, en waarbij een eigenwaardeprobleem opgelost wordt, skores gevonden die positief zijn en gesommeerd 1 opleveren: de skores worden dus door de methode in gestandaardiseerde vorm opgeleverd. Hierbij moet worden bedacht dat aan het begin van de analyse geen effektskores per criterium en effekt beschikbaar zijn: het enige wat dan bekend is, is de preferentie per criterium van alternatieven ten opzichte van elkaar.

### Kwalitatieve aanpak (QUALI)

Bij de omvorming van kardinale gegevens tot kwalitatieve gegevens wordt aan de verschillende alternatieven per criterium een waarde gegeven op een schaal, waarop begrippen als "goed", "matig" en "slecht" voorkomen of ook (voor criteria) "zeer belangrijk", "belangrijk" en "niet (zo) belangrijk". Deze waarden kunnen ook gepresenteerd worden door schaalwaarden in getalvorm weer te geven



mits deze op kwalitatieve wijze geïnterpreteerd worden: operaties op deze getallen, zoals bij de gewogen sommeringsmethoden (zie paragraaf 3.3) passen dan ook niet in deze kontekst.

In het onderzoek is QUALI slechts gekombineerd met een konkordantiemethode (zie paragraaf 3.6) toegepast. Deze keuze is gemaakt om te voorkomen dat de kwalitatieve gegevens toch ordinaal of zelfs kardinaal geïnterpreteerd worden door een evaluatiemethode, die zulke informatie vooronderstelt. Belangrijk bij de toepassing van kwalitatieve gegevens en modellering is dan ook de opmerking, dat geen echte rangvolgorde van alternatieven het resultaat kan vormen van de MKE: de MKE kan niet meer informatie opleveren dan in de uitgangsgegevens is vervat, daarentegen kan een MKE wel informatie verloren doen gaan door de beschikbare informatie niet optimaal te benutten.

#### 4.2.2 Eenvoudige kardinale methoden

De standaardisatiemethoden in deze subparagraaf hebben als gemeenschappelijk element, dat er een zeer eenvoudig functioneel verband bestaat tussen de skores voor en na de standaardisatie: de skores zijn kardinaal van aard en worden door (TOTAAL, MAX, MAXMIN) of op (MIN) een vast getal (per criterium) gedeeld. Van al deze methoden en van die in subparagraaf 4.2.3 bestaan in principe twee varianten: hoogste skore voor beste alternatief of omgekeerd. In de literatuur wordt per methode slechts één van de varianten gegeven; in tabel 4.1 waarin voor de methoden in deze en de volgende subparagraaf de functionele verbanden zijn opgenomen zijn echter beide varianten gegeven.

#### TOTAAL

Waarschijnlijk de meest voorkomende standaardisatiemethode voor kardinale gegevens bestaat in het delen van de effektskores door het totaal van alle skores. Dit levert een effectiviteitsskore tussen 0 en 1 op, waarbij de skore geïnterpreteerd kan worden als het aandeel dat het effect van een bepaald alternatief op een bepaald criterium uitmaakt van het totaal der effecten op dat criterium. Elke eenheid van een bepaald criterium wordt bij de methode TOTAAL steeds gelijk gewaardeerd.

nummer	naam methode	variant		p.
		hoogst=best	hoogst=slechtst	
1	TOTAAL	$\frac{\text{skore}_{ij}}{\text{tot}_j}$	$\frac{\text{tot}_j - \text{skore}_{ij}}{\text{tot}_j}$	40
2	MAX	$\frac{\text{skore}_{ij}}{\text{max}_j}$	$\frac{\text{max}_j + \text{min}_j - \text{skore}_{ij}}{\text{max}_j}$	42
3	MAXMIN	$\frac{\text{skore}_{ij} - \text{min}_j}{\text{max}_j - \text{min}_j}$	$\frac{\text{max}_j - \text{skore}_{ij}}{\text{max}_j - \text{min}_j}$	42
4	MIN	$\frac{\text{max}_j + \text{min}_j}{\text{max}_j} - \frac{\text{min}_j}{\text{skore}_{ij}}$	$\frac{\text{min}_j}{\text{skore}_{ij}}$	42
5	NUTS	$\frac{\text{skore}_{ij} - \text{min}_j}{\lambda_j \cdot \text{min}_j - \text{skore}_{ij}} \cdot C$  $C = \frac{\lambda_j \cdot \text{min}_j - \text{max}_j}{\text{max}_j - \text{min}_j}$  $0 < \lambda_j < 1$	$\frac{\text{max}_j - \text{skore}_{ij}}{\lambda_j \cdot \text{max}_j - \text{skore}_{ij}} \cdot C$  $C = \frac{\lambda_j \cdot \text{max}_j - \text{min}_j}{\text{max}_j - \text{min}_j}$  $\lambda_j > 1$	43
6	INDIF	derdegraadsfunctie <sup>1)</sup> $F(\text{skore}_{ij})$	1 - derdegraadsfunctie <sup>1)</sup> $F(\text{skore}_{ij})$	43

Betekenis symbolen:

$\text{skore}_{ij}$ : skore van alternatief i op criterium j;

$\text{tot}_j$  : totaal der skores op criterium j;

$\text{max}_j$  : maximale waarde van criterium j;

$\text{min}_j$  : minimale waarde van criterium j;

$\lambda_j$  : keuzeparameter waarmee kan worden geregeld hoeveel één extra eenheid van criterium j minder gewaardeerd wordt dan de voorgaande eenheid.

1) namelijk  $F(\text{skore}_{ij}) = A \cdot (\frac{1}{3} \cdot \text{skore}_{ij}^3 + B \cdot \text{skore}_{ij}^2 + C \cdot \text{skore}_{ij} + D)$   
zodanig dat  $F(\text{max}_j) = 1$ ,  $F(\text{min}_j) = 0$ ,  $F'(\text{max}_j) = \lambda_j F'(\frac{1}{2}(\text{min}_j + \text{max}_j))$  ( $0 < \lambda_j < 1$ )

Tabel 4.1 Specificatie van onderzochte standaardisatiemethoden



### MAX

Bij de methode MAX wordt elke score gedeeld door het maximum der scores, waarbij weer scores tussen 0 en 1 ontstaan: de gunstigste score krijgt effectiviteit 1 en de ongunstigste score het quotiënt van de laagste en de hoogste score. Doordat hiertussen lineair wordt geïnterpoleerd wordt ook hier elke eenheid van een criterium steeds even zwaar meegenomen.

### MAXMIN

De gunstigste score krijgt bij de methode MAXMIN als waardering 1 en de ongunstige score als waardering 0. De tussenliggende scores worden via lineaire interpolatie gewaardeerd; deze interpolatie vindt plaats middels deling van alle scores door het verschil tussen de grootste en de kleinste score. Ook hier worden extra eenheden van een bepaald criterium steeds gelijk gewaardeerd.

### MIN

Bij de methode MIN worden de uitersten op dezelfde wijze gewaardeerd als bij de methode MAX; de scores worden echter gedeeld op de laagste waarde. Als gevolg van het gehanteerde rekenvoorschrift neemt de effectiviteit voor een eenheid extra van een criterium omgekeerd kwadratisch af met de grootte van de effektscore.

#### 4.2.3 Kardinale methoden met keuzeparameter

Bij de keuze van één der methoden uit de vorige subparagraaf ligt de waardering van eenheden op een criterium in verschillende delen van de meetschaal ten opzichte van elkaar vast: bij de eerste drie methoden zijn de waarderingen gelijk, terwijl de waardering bij de laatste methode omgekeerd kwadratisch met de grootte van de effektscore afneemt. In deze subparagraaf worden twee methoden behandeld, waarbij via een keuzeparameter invloed kan worden uitgeoefend op de onderlinge verhouding van de effectiviteiten in verschillende deelgebieden van de meetschaal.

## NUTS

Bij de methode NUTS krijgt de gunstigste skore als waardering 1; aan de ongunstigste skore wordt de waarde 0 toegekend. De tussenliggende skores zijn zodanig bepaald, dat iedere eenheid van een criterium in gunstige richting steeds minder wordt gewaardeerd. Deze aanname komt overeen met de uitgangspunten van de nutstheorie.

Uit tabel 4.1 blijkt, dat de methode NUTS een faktor toevoegt aan de formule van de methode MAXMIN, die voor skores vlak bij de maximumskore dicht bij 0 ligt en er verder zorg voor draagt dat bovenvermelde skore afhankelijke waardering van extra eenheden wordt geëffektueerd; de keuzeparameter  $\lambda$  bepaalt de mate waarin zulks plaatsvindt. Dit wordt geïllustreerd in de figuren 4.1 tot en met 4.4, welke verder besproken zullen worden in paragraaf 4.4.

## INDIF

Een andere methode, waarbij rekening is gehouden met skore afhankelijke waardering van extra eenheden op een criterium is de methode INDIF. Bij deze methode krijgt de kleinste skore wederom als waardering 0 en de grootste skore wordt met 1 gewaardeerd. De tussenliggende skores worden nu gewaardeerd door de koëfficiënten van een derdegraads veelterm zo te kiezen, dat de verhouding tussen de (gelijke) helling van de funktie bij de minimum- en de maximumskore tot de helling midden tussen deze waarden in gelijk is aan een in te stellen parameter  $\lambda$ . Hiermee wordt bereikt, dat schaalwaarden in het middengebied tussen de hoogste en de laagste skore sterker differentiëren dan extreme skores. Impliciet worden waarden in het middengebied daardoor sterk gewogen (zie hoofdstuk 5).

### 4.2.4 Geavanceerde methoden

## ENTROPIE

Het entropieconcept stamt uit de informatietheorie, die het op haar beurt ontleent aan de natuurkunde. Simpel gezegd is entropie een maat voor de aanwezigheid van wanorde. In de informatietheorie duidt wanorde op afwezigheid van een



goede voorspellingsmethode van een volgend gegeven op basis van de voorafgaande gegevens; is de voorspelbaarheid groot, dan is de hoeveelheid (nieuwe) informatie die in een volgend gegeven is vervat klein: informatie-inhoud en entropie zijn daardoor nauw verweven concepten.

Zowel de effectenmatrix als de gewichtenvectoren kunnen worden beschouwd als boodschappen, waarvan de entropie de mate van diversifikatie aangeeft. Met behulp van een konventionele entropiemaatstaf kunnen effectiviteiten worden bepaald die aangepast zijn aan de gemiddelde intrinsieke informatie, die in de verzameling uitkomsten van alternatieven met betrekking tot een bepaald criterium is vervat.

De totale hoeveelheid informatie, die een criterium  $j$  toelevert aan de evaluatie kan uitgedrukt worden in de entropiemaat

$$h_j = \frac{1}{\log I} \sum_{i=1}^I (-e_{ji} \log e_{ji}) \quad (4.2)$$

waarbij  $e_{ji}$  de gestandaardiseerde effectwaarde voorstelt van alternatief  $i$  voor criterium  $j$ ; als standaardisatie is hier door het totaal van de skores gedeeld (methode TOTAAL), waardoor de  $e_{ji}$  als kansen kunnen worden geïnterpreteerd. Naarmate  $h_j$  kleiner is werkt criterium  $j$  in de evaluatie minder diskriminerend. Door per criterium  $1-h_j$  als gewicht te hanteren, wat in feite neerkomt op standaardisatie per criterium, waarbij toch ook een bewerking over de criteria heen plaatsvindt, wordt in feite een additionele veronderstelling ingebracht. Er wordt namelijk verondersteld, dat het wenselijk is, criteria te wegen naar mate van hun capaciteit tussen alternatieven te diskrimineren. In [2] besteedt Voogd enige opmerkingen aan deze kwestie.

#### WATERLAB

Zoals in hoofdstuk 5 zal blijken, impliceert het gebruik van een standaardisatiemethode een weging tussen criteria: het terugbrengen van effectskores die tussen, bijvoorbeeld, 10 en 200 liggen, tot effectiviteiten tussen de grenzen 0 en 1 houdt in, dat (globaal) aan de effectiviteiten 200 maal het gewicht moet worden toegekend, dat bij de effectskores past om de weging die veroorzaakt wordt door de standaardisatie te niet te doen. Met deze impliciete

weging is echter bij de vaststelling door de beslissingsnemer van expliciete gewichten geen rekening gehouden. In feite zou deze bij het bepalen van gewichtenvectoren met de te hanteren standaardisatiemethoden (in principe voor elk criterium apart!) rekening moeten houden, voorwaar een exactheid waaraan onmogelijk te voldoen valt.

Door het WL is steunend op eerder opgedane ervaring een methode WATERLAB ontwikkeld die de gesignaleerde problematiek te lijf gaat. De benadering is gebaseerd op het door middel van trade-offs tussen de onderscheiden criteria bepalen van gestandaardiseerde gewichten in plaats van gestandaardiseerde effectskores zoals bij de traditionele methoden het geval is. Naar is gebleken zijn beslissingsnemers zeer wel in staat met behulp van trade-offs afwegingen tussen criteria te maken.

In hoofdstuk 5 komen enige andere aspecten van de methode WATERLAB naar voren. Gezien het feit, dat zij voornamelijk als preferentiemethode moet worden gezien vindt daar ook een uiteenzetting van haar technische toepassing plaats. Aan het begrip trade-off is aandacht besteed in deel II.

#### 4.3 Externe extremen

Bij de in tabel 4.1 gegeven standaardisatieformules zijn de grootheden  $\min_j$  en  $\max_j$  intern gedefinieerd (als respectievelijk de kleinste en de grootste optredende schaalwaarde van criterium  $j$ ). Dat wil zeggen, dat  $\min_j$  en  $\max_j$  afhankelijk zijn van de in de MKE meespelende alternatieven. In vele evaluatieprocedures zullen gedurende de verschillende overleggronden alternatieven worden geschrapt en misschien nieuwe alternatieven worden geïntroduceerd. Hierdoor kan het heel wel gebeuren dat voor één of meer der gebruikte criteria de minimum- en/of maximum skore verandert; vrijwel zeker zal dit ook met het totaal der skores gebeuren. Het gevolg hiervan is, dat de standaardisatieformules met andere parameters werken, waardoor uiteindelijk de rangorde van de overgebleven alternatieven na toepassing van de MKE-methode in kwestie totaal kan verschillen van de rangorde die verkregen wordt met de oorspronkelijke verzameling alternatieven.



Het is door de hierboven geschetste werkwijze niet duidelijk of inderdaad steeds de minst belovende alternatieven worden uitgesloten van verdere analyse. Met een voorbeeld kan dit effect worden gedemonstreerd. Uitgangspunt is de effektskorematrix van het IODZH-voorbeeld (tabel 2.1), waaruit via de standaardisatiemethode MAX (subparagraaf 4.2.2) en de evaluatiemethode SOM (gewogen somming: paragraaf 3.3) onder gebruikmaking van gewichtenvector C uit tabel 2.4 de volgorde van alternatieven wordt bepaald. Na weglaten van het slechtst skorende alternatief wordt een nieuwe rangorde bepaald, waarna de procedure wordt herhaald tot twee alternatieven resteren.

De resultaten hiervan zijn opgenomen in tabel 4.2: boven de streep staat het nummer van het alternatief en elke rij bevat de rangorde der overgebleven alternatieven. Met behulp van bovenstaande procedure is eenvoudig een indruk te verkrijgen aangaande de gevoeligheid van de gevonden rangorde in een MKE voor het laten vervallen van alternatieven: er wordt geen betere volgorde gevonden maar de stabiliteit van de rangorde kan worden nagegaan.

Uit een aantal van dergelijke experimenten met het IODZH-materiaal blijkt dat de standaardisatiemethoden MAX, MAXMIN en NUTS matig gevoelig zijn voor het weglaten van alternatieven. Methode MIN is het minst gevoelig. Andere methoden voor standaardisatie zoals ENTROPIE en TOTAAL bleken op dit punt extreem gevoelig te zijn. Dat methode MIN het minst gevoelig blijkt, is niet verwonderlijk aangezien bij het IODZH-voorbeeld voor alle criteria geldt "hoogst = slechtst". Omgekeerd mag verwacht worden, dat in het geval dat voor alle criteria geldt "hoogst = best" de methode MAX het minst gevoelig is voor het weglaten van alternatieven.

Om aan dit probleem van intern gedefinieerde maxima en minima (en totalen) te ontkomen lijkt het aanbevelenswaardig door een team van deskundigen of door de beslisser(s) zelf maximum- en minimumwaarden per criterium te laten vaststellen (of in het geval van toepassing van de methode TOTAAL de gemiddelde (of totale) skore). Aan deze methode kleeft echter het bezwaar dat het criterium, waarbij deze extern gedefinieerde maxima en minima relatief dicht bij de intern gedefinieerde maxima en minima gekozen zijn, ten opzichte van andere criteria in belangrijkheid toeneemt, en daardoor diskriminerender gaat werken.

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

17	14	10	5	1	2	6	3	8	4	7	12	9	15	13	18	20	19	21	16	22	23	24
16	14	12	5	1	2	7	3	8	4	6	11	9	13	15	18	21	19	20	17	22	23	
12	16	13	5	2	1	7	4	10	3	6	15	8	11	14	19	21	18	20	17	22		
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13	20	21	17	19	18			
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13	20		17	19	18			
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13			17	19	18			
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13			17		18			
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13			17					
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13			17					
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13			17					
4	14	15	5	2	1	8	7	12	3	6	16	9	10	13			17					
5	14		6	2	1	8	7	12	3	4		10	9	13								
5			6	2	1	8	7	12	3	4		10	9	13								
5			6	2	1	8	7	12	3	4		10	9									
5			6	2	1	8	7		3	4		10	9									
5			6	2	1	8	7		3	4		10	9									
5			6	2	1	8	7		3	4			9									
5			6	2	1		7		3	4												
5			6	2	1				3	4												
5				2	1				3	4												
				2	1				4	3												
				2	1					3												
				2	1																	

Toelichting:

Boven de streep staan de nummers van de gehanteerde alternatieven; daaronder staat in elke rij de positie van het betreffende alternatief in de verkregen rangorde, waarbij naar beneden toe voor elke rij één alternatief (het slechtste van de rij erboven) minder bij de MKE is betrokken. Toegepast zijn de standaardisatiemethode MAX (subparagraaf 4.2.2) en de evaluatiemethode SOM (gewogen somming: paragraaf 3.3).

Tabel 4.2 Illustratie van het effect van schrappen van het slechtst skorende alternatief bij het gebruik van interne extremen bij de standaardisatie.



Dit is hetzelfde effect dat optreedt bij het geven van rapportcijfers in het onderwijs: daar diskrimineert een leraar die cijfers van 2 tot 10 geeft meer dan zijn/haar collega die niet verder gaat dan de cijfers 5, 6 en 7: de eerste heeft een grotere invloed op de overgangsbepaling met betrekking tot zijn/haar leerlingen. Bij de standaardisatiemethoden van subparagraaf 4.2.2 worden kleine scoreverschillen bij een criterium gelegen tussen nauwe grenzen, gegeven door de externe extremen, door de methode opgeblazen tot het volledige schaalgebied tussen, bij voorbeeld, 0 en 1.

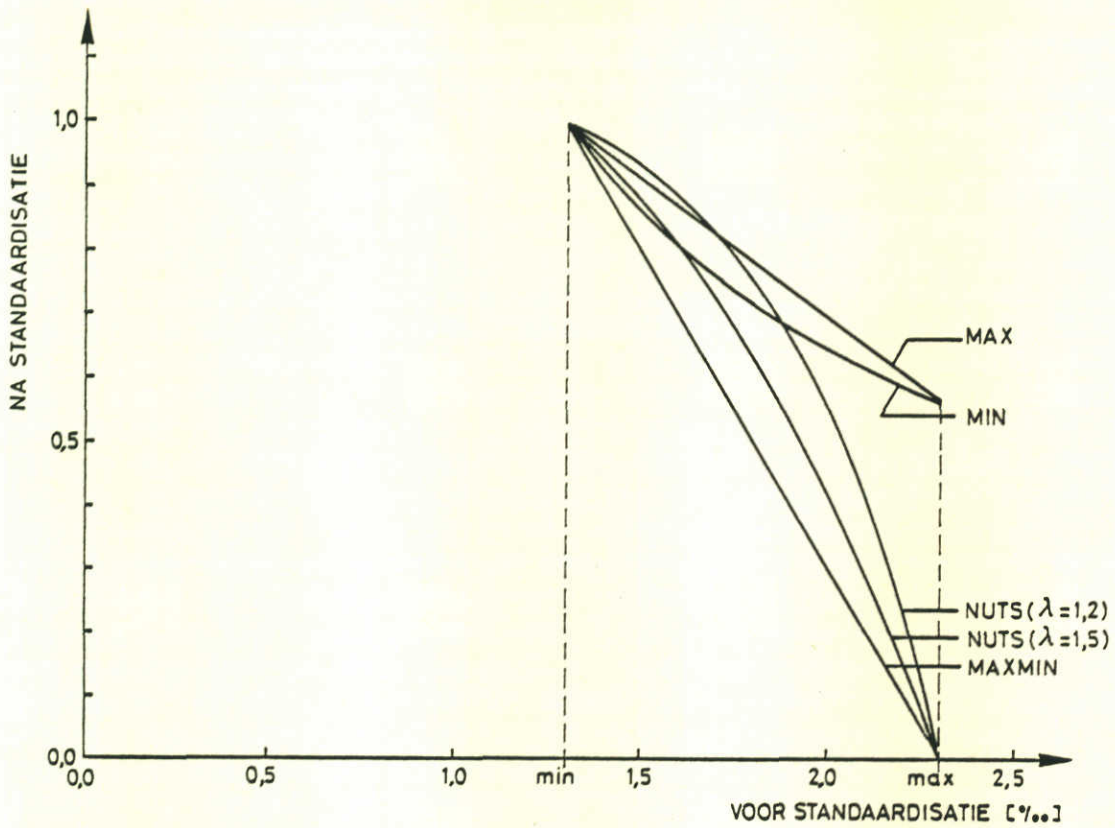
In principe is dit probleem op te lossen door gewichtenvectoren samen te stellen die afhankelijk zijn van de gekozen externe extremen. In de praktijk lijkt dit echter een onmogelijkheid. Het complexe karakter van het bepalen van de gewichtenvectoren wordt dan nog complexer.

#### 4.4 Vergelijking

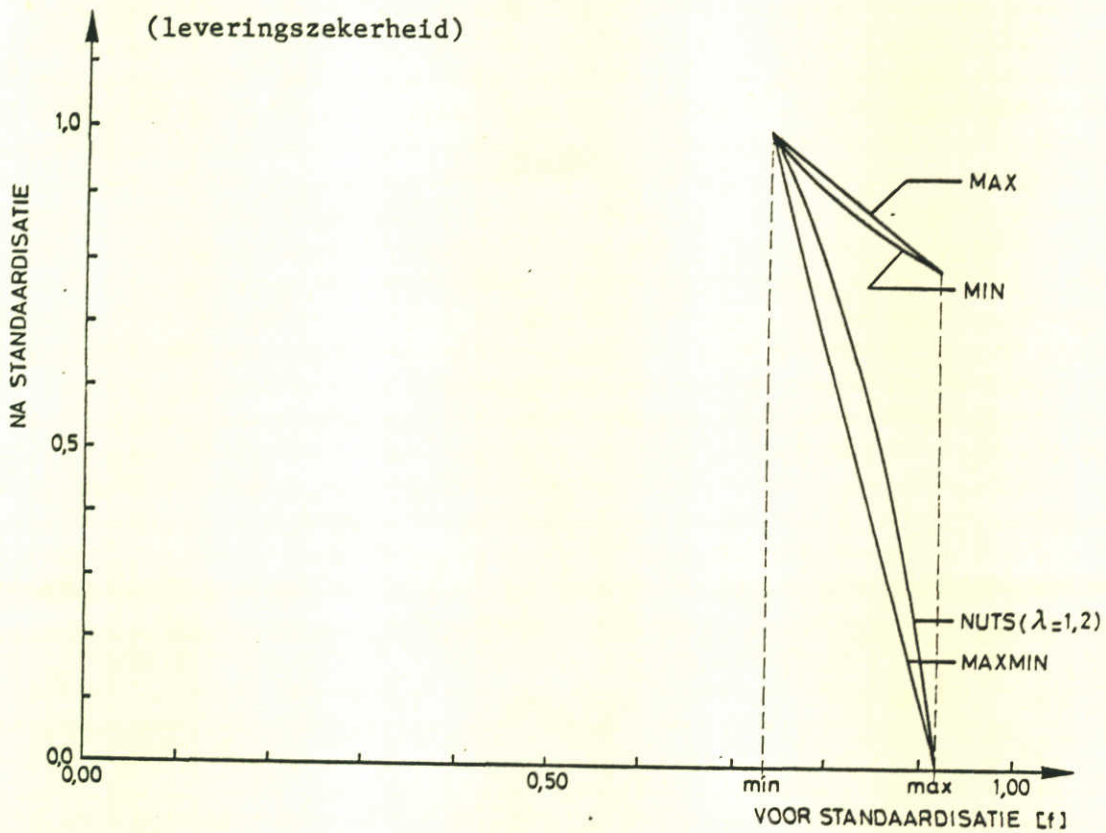
In de figuren 4.1 tot en met 4.4 is voor de criteria "leveringszekerheid", respectievelijk "gemiddelde produktiekosten", "bodembelasting" en "natuurschade" (ontleend aan het IODZH voorbeeld) en voor de standaardisatiemethoden van tabel 4.1 (en subparagrafen 4.2.2 en 4.2.3) de relatie tussen effectskores voor en effectiviteiten na standaardisatie grafisch weergegeven.

Uit de figuren 4.1, 4.2 en 4.4 blijkt dat voor niet zeer extreme situaties de methoden MIN en MAX op elkaar gelijken evenals MAXMIN en NUTS (het effect van NUTS kan bovendien worden beïnvloed door de keuze van  $\lambda$ ); TOTAAL en INDIF liggen dicht bij deze methoden.

Alle methoden behalve MIN gaan meer op elkaar gelijken naarmate minimum- en maximumwaarde verder van elkaar liggen, terwijl MIN dan steeds verder af gaat wijken. Figuur 4.3 illustreert duidelijk dat de methode MIN ten enen male ongeschikt is voor standaardisatie bij skores, die positieve en negatieve waarden kunnen aannemen: in de buurt van nul is het resultaat zelfs onbegrensd; bij alleen negatieve skores kan MIN ook niet zonder meer worden gebruikt, omdat de gunstigste skore getransformeerd wordt naar de ongunstigste waarde: beter kan dan van teken worden gewisseld en alleen met positieve skores worden gewerkt. De familierelatie tussen MAXMIN en NUTS is duidelijk te zien;

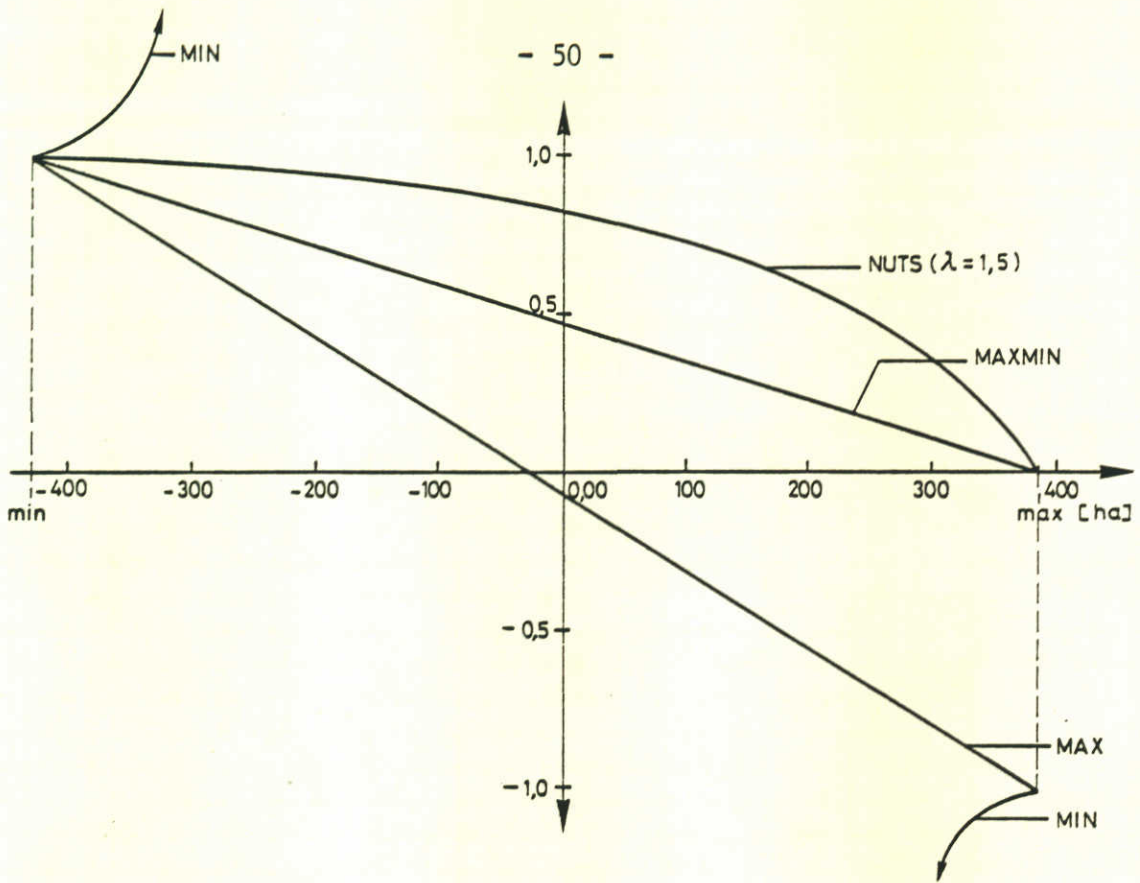


Figuur 4.1 De relatie tussen scores voor en na standaardisatie: criterium 1 (leveringszekerheid)

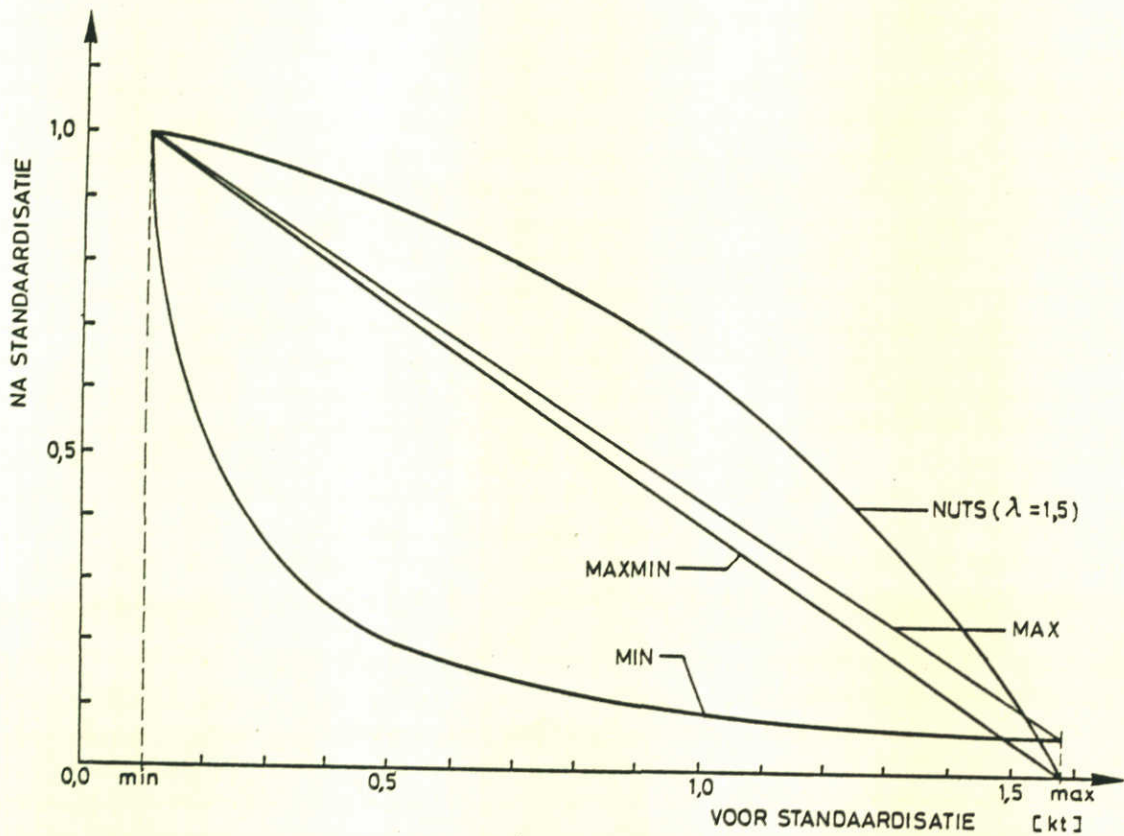


Figuur 4.2 De relatie tussen scores voor en na standaardisatie: criterium 6 (gemiddelde produktiekosten)





Figuur 4.3 De relatie tussen skores voor en na standaardisatie: criterium 7 (natuurschade)



Figuur 4.4 De relatie tussen skores voor en na standaardisatie: criterium 9 (bodembelasting)

NUTS kan beter niet worden gebruikt in dit geval, omdat hier in feite een intervallschaal is gebruikt, waardoor niet duidelijk is of en zo ja in welke mate extra schade extra sterk moet worden gewaardeerd.

Een andere, hiermee samenhangende, vraag is in hoeverre de keuze van een andere standaardisatiemethode een andere uiteindelijke rangorde tot gevolg heeft. In tabel 4.3 zijn de volgorden weergegeven die resulteren bij de toepassing van verschillende standaardisatiemethoden op de IODZH-gegevens bij konstant houden van de evaluatiemethode (SOM: paragraaf 3.3) en gewichtenvector (C uit tabel 2.4).

Duidelijk blijkt, dat de overgang van één standaardisatiemethode op een andere (op alle criteria is dezelfde standaardisatiemethode toegepast) een grote verschuiving teweeg brengt in de resulterende rangorden; men zie de alternatieven 12, 15 en 20.

De konklusie, het zij nogmaals vermeld, luidt, dat de keuze van een standaardisatiemethode een grote invloed heeft op de eindresultaten van een MKE. In deze paragraaf lijkt dit genoegzaam aangetoond, terwijl nog niet eens gekeken is naar de relatie tussen gewichtentoekenning en standaardisatie; dit is het onderwerp van hoofdstuk 5, waar het onder andere gaat over impliciete weging door standaardisatie tussen criteria en interne weging binnen criteria.



alternatieven	standaardisatiemethode			
	MAX	MIN	MAXMIN	NUTS
1	23	19	13	17
2	24	24	23	22
3	19	12	8	12
4	14	17	7	6
5	6	13	3	3
6	2	3	2	2
7	18	10	18	21
8	12	7	5	9
9	1	1	1	1
10	5	14	12	7
11	4	5	15	14
12	3	2	16	16
13	7	16	9	5
14	9	8	19	19
15	8	4	20	20
16	20	18	10	13
17	21	22	14	11
18	22	23	21	18
19	16	15	24	23
20	11	6	22	24
21	17	9	6	10
22	10	11	4	4
23	15	20	11	8
24	13	21	17	15

Tabel 4.3 Rangorde van alternatieven bij het gebruik van verschillende standaardisatiemethoden in combinatie met MKE-methode SOM

## 5 Gewichten

### 5.1 Inleiding

In deel II is al summier ingegaan op het gebruik van gewichten bij MKE. Kort gezegd geven gewichten het relatieve belang van criteria ten opzichte van elkaar weer; als zodanig zijn zij de weerspiegeling van de mening of voorkeur (preferentie) van een beslissingsnemer of een maatschappelijke groepering. Gewichten kunnen ook worden toegepast in een laatste fase van het afwegingsproces, daar waar de rangorden van alternatieven die het resultaat zijn van de voorgaande MKE per beslissingsnemer of groepering moeten worden samengesmolten tot één uiteindelijke rangorde. Hierbij representeren de gewichten de relatieve invloed die elk der groeperingen op het eindresultaat heeft.

In de MKE worden de gewichten gebruikt om skores op een belangrijk criterium een relatief grote invloed op de uiteindelijke rangorde van de alternatieven te geven. In de literatuur, o.a. [2], wordt er op gewezen, dat de keuze van de gewichten grote invloed op de uiteindelijke rangschikking kan hebben: de gevoeligheid van de MKE voor de gewichten is derhalve groot, iets wat ook in het onderhavige onderzoek bleek. In paragraaf 5.2 is ingegaan op methoden ter verkrijging van betrouwbare gewichten, de zogenaamde preferentiemethoden.

Reeds eerder werd op het feit gewezen, dat toekenning van gewichten aan de criteria niet los gezien mag worden van de toegepaste standaardisatie: door de standaardisatie worden effektskores getransformeerd naar effectiviteitskores op een dikwijls geheel verschillende schaal, waardoor sprake kan zijn van het toekennen van (extra) gewichten aan de skores van de alternatieven. Aan het verschijnsel van deze impliciete gewichten wordt aandacht besteed in paragraaf 5.3.

Een aantal standaardisatiemethoden brengt behalve een schaalverandering voor een bepaald criterium tevens een ongelijkmatige interne weging binnen dat criterium teweeg: hiermee wordt bedoeld, dat gelijke skoreverschillen in verschillende delen van de meetschaal voor het criterium ongelijke verschillen in de toegekende effectiviteitskores impliceren. Hiermee wordt een skore-afhankelijke weging geïntroduceerd, die het onderwerp is van paragraaf 5.4.



Een poging recht te doen aan de onderlinge verwevenheid tussen gewichtentoe-  
kenning en standaardisatie heeft geleid tot een methode WATERLAB, waaraan in  
subparagraaf 4.2.4 al enige woorden zijn gewijd. Paragraaf 5.5 gaat in op  
aspecten van WATERLAB in verband met bovenstaande problematiek.

## 5.2 Preferentiemethoden

Onder de term preferentiemethoden worden in dit rapport verstaan methoden, die  
de voorkeur van beslissingsnemers helpen uitkristalliseren en vertalen in een  
gewichtenvector voor gebruik in een MKE. Het gaat hierbij om meet- en scha-  
lingsmethoden in complexiteit variërend van het niveau van vuistregels tot dat  
van de methode Saaty (zie paragraaf 3.5). Essentieel bij de toepassing van  
zulke preferentiemethoden is de grote mate van interactie tussen MKE-analist  
en beslissingsnemer die is vereist om tot betrouwbare resultaten te kunnen  
komen.

Preferenties kunnen op de volgende manieren vastgesteld worden:

- op grond van waargenomen gedrag (revealed preference);
- door het poneren van alternatieve beleidsrichtingen;
- op grond van enquêtes of interviews;
- interactieve beslissingstechnieken.

Alleen op de laatstgenoemde groep van methoden zal hier worden ingegaan. Voor  
de overige typen methoden wordt naar de literatuur [2] verwezen.

Bij interactieve beslissingstechnieken worden dikwijls in samenspraak tussen  
beslissingsnemer en analist iteratief alternatieven geëlimineerd en gewichten  
vastgesteld; dat wil zeggen, dat gewichten en alternatieven tegelijk worden  
behandeld. Dit laatste bergt het gevaar in zich, dat nieuwe criteria impliciet  
in het proces hun intrede doen, waarmee de MKE in sterke mate ondoorzichtig  
wordt en één van de voordelen van het verrichten van een MKE teniet wordt  
gedaan. Een zodanige MKE vergt dan ook een zorgvuldige bewaking van de proce-  
dure.

Aan twee methoden die tot de groep van de interactieve preferentiemethoden  
gerekend zouden kunnen worden, wordt in dit hoofdstuk aandacht besteed: het  
gaat hier om de methode Saaty en de methode WATERLAB.

Bij de methode Saaty worden, analoog aan de procedure van de paarsgewijze vergelijking van de alternatieven per criterium, de criteria paarsgewijs onderzocht op hun onderlinge verhouding in belangrijkheid. Hierbij kan ook gebruik gemaakt worden van de door Saaty ontworpen schaal van dominantie, lopend van 1 tot 9 (zie ook paragraaf 3.5). Door oplossing van het eigenwaardeprobleem wordt een eigenvector verkregen met de gestandaardiseerde gewichten van de criteria.

Indien de methode Saaty wordt toegepast bij kardinale gegevens, kan er sprake zijn van impliciete standaardisatie: als oplossing van het eigenwaardeprobleem of een variant daarvan, worden skores gevonden die positief zijn en gesommeerd 1 opleveren. De skores zijn dus automatisch gestandaardiseerd. De bedoelde impliciete standaardisatie vindt echter plaats bij het toekennen van schaalwaarden bij de paarsgewijze vergelijkingen. Hierdoor bestaat de mogelijkheid van 'dubbel tellen' in die zin dat bij criteria die als 'belangrijk' worden aangemerkt er grotere schaalwaarden worden gebruikt dan bij minder belangrijke criteria, zonder dat het cijfer- of gevoelsmateriaal zelf daar aanleiding toe zou geven. Ook is er een mogelijkheid tot manipulatie, doordat extreme schaalwaarden impliciet aanleiding geven tot een schijnbaar groter gewicht van de betreffende criteria. Het is daarom vereist dat degenen die schaalwaarden toekennen, hun keuzen kunnen verantwoorden. Dankzij het feit dat in het algemeen meer paarsgewijze vergelijkingen worden uitgevoerd dan minimaal noodzakelijk, is het echter mogelijk te toetsen op duidelijk inkonsekvent of manipulatief toekennen van schaalwaarden. In zeer "gevoelig" liggende toepassingen kan men eventueel helemaal afzien van schaalwaarden en alleen de schaalwaarden 1 (gelijk) of 3 (belangrijker) benutten.

Bij de methode WATERLAB worden van de beslissingsnemer(s) standpunten gevraagd over de gewenste positie op trade-offs tussen criteria. In zekere zin kan dit beschouwd worden als een variant op de informatie, waarom bij de toepassing van de Saaty-methode wordt gevraagd. Het verschil is echter, dat bij deze standpuntsbepaling expliciet rekening wordt gehouden met de skores op de criteria. Verder is er sprake van een standaardisatie van de effektskores die tegelijkertijd plaatsvindt. Voor een verdere beschrijving van de methode zelf wordt verwezen naar paragraaf 5.5.



### 5.3 Impliciete gewichten

Verschillende standaardisatiemethoden - en zelfs verschillende varianten van dezelfde standaardisatiemethode - leiden tot onderling verschillende rangschikking van de alternatieven, zelfs als zij worden gekombineerd met één en dezelfde MKE-methode en dezelfde gewichtenvector.

Dit verschijnsel kan als volgt verklaard worden. Bij standaardisatie wordt de effectskore  $x_{ij}$  (van alternatief  $i$  op criterium  $j$ ) met behulp van een standaardisatiefunctie  $S$  omgezet in een effectiviteit  $S(x_{ij})$ . In het geval van een lineaire standaardisatiefunctie kan deze als volgt geschreven worden:

$$S(x_{ij}) = a_j x_{ij} + b_j \quad (5.1)$$

Bij gewogen somming is de bijdrage van  $x_{ij}$  aan de totale skore van alternatief  $i$

$$w_j \cdot S(x_{ij}) = w_j (a_j x_{ij} + b_j) \quad (5.2)$$

waarbij  $w_j$  = gewicht van criterium  $j$

Twee verschillende standaardisatiemethoden  $S$  en  $S^*$  leveren eenzelfde rangschikking van de alternatieven op als geldt

$$w_j \cdot S(x_{ij}) - w_j \cdot S^*(x_{ij}) = \text{konstant} \quad (5.3)$$

voor iedere  $x_{ij}$ . Uit (5.2) kan afgeleid worden dat (5.3) geldt indien:

$$a_j = a_j^* \quad (5.4)$$

Met andere woorden, de rangschikking van alternatieven die verkregen is met standaardisatiemethode  $S$ , kan ook verkregen worden met methode  $S^*$  als geldt

$$\frac{d S(x_{ij})}{d x_{ij}} = \frac{d S^*(x_{ij})}{d x_{ij}} \quad (5.5)$$

voor ieder criterium  $j$ . Indien de rangschikking, bij methode  $S^*$  verschilt van de rangschikking verkregen met methode  $S$ , kan alsnog de rangschikking van methode  $S$  verkregen worden bij methode  $S^*$  door de effektscores te vermenigvuldigen met het quotiënt van de afgeleiden:

$$\frac{d S(x_{ij})/d x_{ij}}{d S^*(x_{ij})/d x_{ij}} = \frac{a_j}{a_j^*} \quad (5.6)$$

De vector met als elementen  $a_j/a_j^*$  kan aangeduid worden als de vector met zogenaamde impliciete gewichten.

Experimenten met het IODZH-materiaal hebben tabel 5.1 opgeleverd. In deze tabel is aangegeven met welke impliciete gewichtenvector de standaardisatiemethode MAXMIN (met extern gespecificeerde extremen), MAX (interne extremen) en TOTAAL moeten worden doorgerekend om de rangorde van MAXMIN (nu met interne extremen) te verkrijgen. Tevens is de verhouding van het grootste en het kleinste gewicht in elke gewichtenvector aangegeven. In dit voorbeeld is gebruik gemaakt van MKE-methode SOM en van de gewichtenvector C uit tabel 2.4.

Dat deze impliciete gewichten een grote invloed op het uiteindelijke resultaat uitoefenen is in één oogopslag duidelijk gezien de grote waarde in de onderste rij van tabel 5.1. Nadere beschouwing van deze tabel leert dat de standaardisatiemethode TOTAAL het criterium kwaliteit (criterium 5) zes maal zo zwaar meetelt als het criterium natuur (criterium 2), terwijl MAX juist het criterium natuur twee keer zo zwaar weegt.

De konklusie moet luiden dat impliciete gewichten gezien hun grootte van groot belang zijn. Ten behoeve van het IODZH zijn (externe) gewichtenvectoren vastgesteld, waar in het meest extreme geval de gewichten voor natuur en kwaliteit in een onderlinge verhouding van vier ten opzichte van elkaar staan. De impliciete gewichten blijken derhalve belangrijker te zijn.



standaardisatiemethode	MAXMIN	MAX	TOTAAL
kriterium	(extern)	(intern)	
1	7,8	2,3	37
2	18,0	0,82	166
3	5,0	4,9	100
4	11,3	3,4	62
5	2,5	1,8	29
6	13,7	4,8	102
7	2,1	1,7	29
8	2,1	1,2	7
9	2,7	1,1	11
10	3,4	2,3	43
grootste/kleinste	9	6	24

Opm. Getallen geven aan met welke gewichten de effektscores moeten worden vermenigvuldigd om dezelfde rangorde van alternatieven te verkrijgen als verkregen is met MAXMIN (interne extremen).

Tabel 5.1 Impliciete gewichten van standaardisatiemethoden

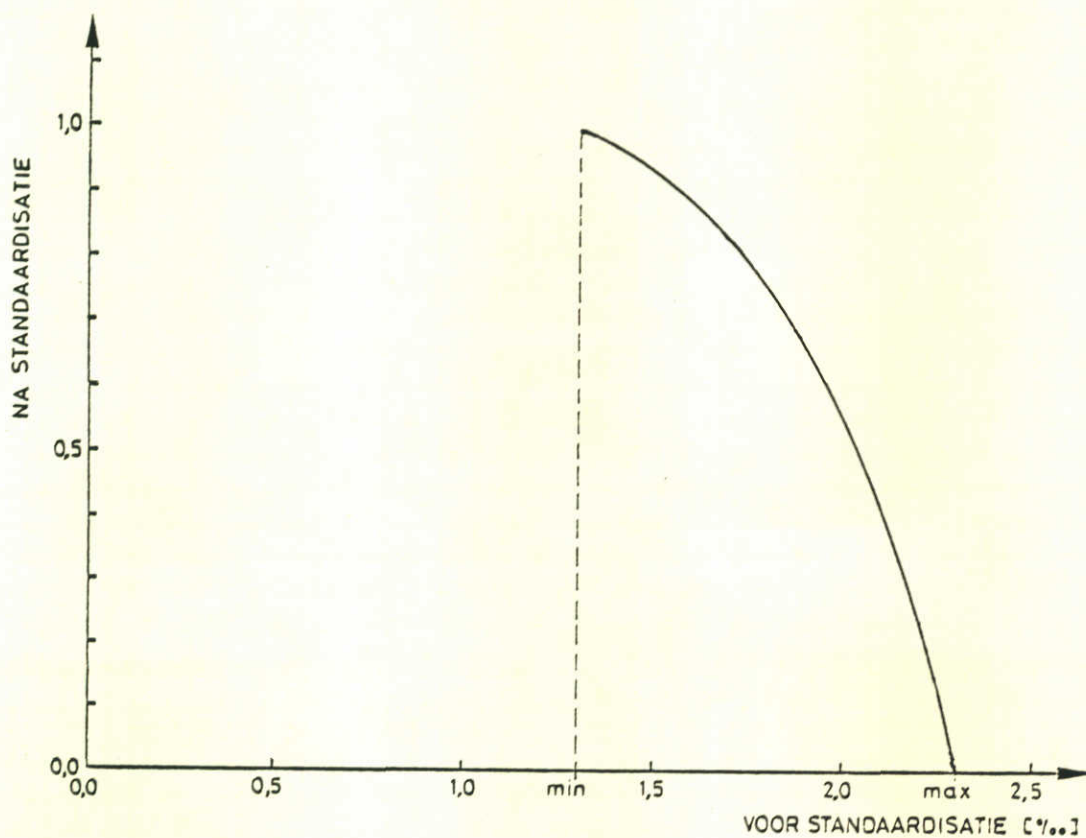
Het gewetensvol bepalen van gewichtenvectoren door beslissingsnemers kan dan ook alleen zinvol zijn als de standaardisatieproblematiek nauwgezet mede in ogenschouw wordt genomen.

Andere experimenten met het IODZH-materiaal tonen aan dat ook het verschil in uitkomsten bij het gebruik van kardinale en ordinale evaluatiemethoden op het konto van impliciete gewichtentoekenning kan worden geschreven; zoals in hoofdstuk 4 is opgemerkt wordt in dit rapport, mede op grond van bovenstaand verschijnsel, ook het toekennen van rangnummers aan en met name de indeling in klassen van de effektscores als standaardisatie betiteld: een andere indeling in klassen levert andere uitkomsten!

#### 5.4 Interne weging

Met interne weging wordt hier bedoeld de mate waarin skoreverschillen binnen één criterium tot uitdrukking worden gebracht in verschillen tussen de gestandaardiseerde effektskores. Deze interne weging ontstaat bij het gebruik van niet-lineaire standaardisatiemethoden. Bij deze methoden is de afgeleide geen konstante maar een functie van  $x_{ij}$ .

Met een voorbeeld wordt getracht bovenstaande te verduidelijken. De illustratie zal plaatsvinden aan de hand van de standaardisatiemethode NUTS (met  $\lambda = 1,2$ ) en IODZH-kriterium 1 ("leveringszekerheid"); in figuur 5.1 hieronder (zie ook figuur 4.1) is de effectiviteit als functie van de effektskore uitgezet (voor de variant "hoogst = slechtst").



Figuur 5.1 Interne weging bij standaardisatiemethode NUTS: criterium 1



Duidelijk blijkt uit de figuur dat skoreverschillen bij grote schaalwaarden (in de buurt van de maximale effektskore) een grotere invloed hebben op de effektiviteitskore dan dito bij kleine schaalwaarden. Mathematisch geformuleerd kan bovenbedoelde interne weging worden uitgedrukt als de afgeleide van de effektiviteit naar het effekt; dat wil zeggen, de interne weging van een skore  $x_{ij}$  door een standaardisatiemethode  $S$  kan worden uitgedrukt als  $d S(x_{ij})/d x_{ij}$ . In figuur 5.1 kan deze waarde teruggevonden worden uit de helling van de raaklijn ter plaatse  $x_{ij}$ : de interne weging is de tangens van deze helling.

Voor het geval van figuur 5.1 (NUTS met  $\lambda = 1,2$  en kriterium 1) loopt de interne weging van  $-0,32$  bij de kleinste tot  $-3,17$  bij de grootste schaalwaarde. Hierbij is het minteken een gevolg van de specificatie "hoogst = slechtst" die in feite de schaalwaarden omkeert; het is verder van geen belang. Wat wel van groot belang is, is de verhouding tussen de beide gespecificeerde interne wegingen: skoreverschillen in de buurt van de grootste schaalwaarde zijn tot tien maal zo effectief als die aan het andere einde van het spektrum, daarmee de grote invloed van de interne weging demonstrerend.

De methode NUTS heeft zoals gezegd een keuzeparameter  $\lambda$  die juist geïntroduceerd is om de mate van interne weging te kunnen sturen. Wordt in plaats van  $1,2$  voor  $\lambda$  de waarde  $1,5$  genomen, dan variëren de interne gewichten van  $-0,53$  tot  $-1,87$ , een verhouding van "slechts"  $3\frac{1}{2}$ .

Het zal duidelijk zijn dat deze interne gewing in feite als een specifiek geval van toekenning van impliciete gewichten gezien kan worden. Bij toepassing van niet-lineaire standaardisatiemethoden kan zowel sprake zijn van impliciete gewichten (resultierend in verschillende rangschikkingen van de alternatieven bij gebruik van verschillende standaardisatiemethoden) als van een interne weging die tot een verschillende rangschikking kan zorgen. Op dezelfde wijze als in paragraaf 5.3 is uiteengezet kan het effekt van de standaardisatiemethode (zowel lineair als niet-lineair) op de rangschikking bestudeerd worden. Voor dat doel wordt een matrix  $A$  [ $N \times J$ ] opgesteld met als elementen  $a_{ij}$  de quotiënten van de afgeleiden van de funkties  $S(x_{ij})$  en  $S^*(x_{ij})$ . Op de rijen van matrix  $A$  staan dan de interne wegingsfactoren vermeld.

Het zal tevens duidelijk zijn dat het van groot belang is zorgvuldig de wenselijkheid van deze interne weging af te wegen. Het is natuurlijk, en de nutstheorie onderstreept dit, in verschillende delen van de effektschaal verschillende waarden te hechten aan skoreverschillen, met andere woorden skore-afhankelijke weging toe te passen. Dit is ook wat de methode WATERLAB in paragraaf 5.5 beschrijft. Van belang is echter de relatie tussen extern (door de beslissingsnemer(s)) opgegeven gewichten, impliciete gewichten (paragraaf 5.3) en interne weging in ogenschouw te nemen. Bij de meeste methoden is het inzicht in deze relatie slechts moeizaam te verkrijgen en het risico bestaat, ook gezien ervaringen met de gebruikte literatuur, dat een MKE wordt uitgevoerd zonder de hier genoemde aspecten in de analyse te betrekken.

## 5.5 WATERLAB

### 5.5.1 Inleiding

Gezien de sterke relatie tussen standaardisatie en het toekennen van gewichten aan criteria is door het WL onderzoek gedaan naar het gezamenlijk behandelen van standaardisatie en gewichtentoedeling, waarbij is aangesloten bij de economische nutstheorie. Van de beslissingsnemer wordt gevraagd de door hem/haar als ideaal beschouwde positie aan te geven op iso-nutscurven voor paren criteria; dit wil zeggen dat gevraagd wordt welke verhouding de skores van alternatieven op twee criteria idealiter moeten hebben indien gekozen moet worden tussen combinaties van skores die hetzelfde nut opleveren (op dezelfde trade-off liggen).

In deze paragraaf zullen in het kort de werkwijze en achtergrond van WATERLAB uit de doeken worden gedaan. Achtereenvolgens komt aan de orde:

- 5.5.2: achtergrond, trade-off en Pareto-optimaliteit;
- 5.5.3: formulering van de methodiek;
- 5.5.4: toepassing op het IODZH-praktijkgeval; en
- 5.5.5: enige kanttekeningen bij de methode.



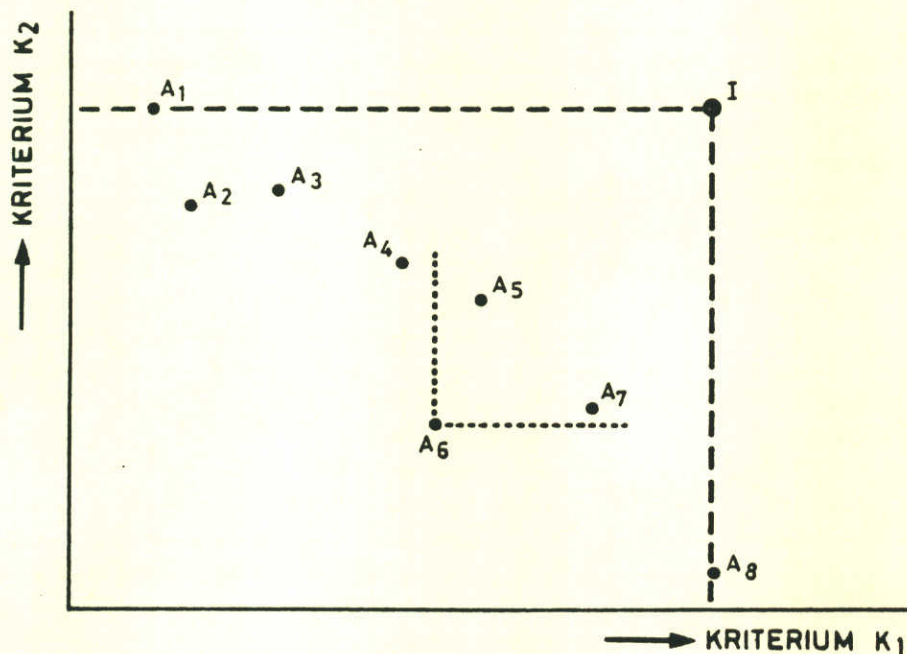
### 5.5.2 Achtergrond

Bij de in de literatuur voorkomende MKE-methoden komt meestal de procedure schematisch neer op een opeenvolging van de volgende stappen:

- 1 - probleemdefinitie;
- 2 - vaststelling criteria;
- 3 - vaststelling effecten;
- 4 - vaststelling gewichten;
- 5 - standaardisatie effektscores;
- 6 - evaluatie;
- 7 - presentatie resultaten.

Deze sequentiële opvolging suggereert, dat de verschillende stappen onafhankelijk van elkaar worden uitgevoerd. In de praktijk zal dit vaak ook zo zijn, ondanks de onderlinge relaties die tussen de verschillende stappen bestaan. Eén van de belangrijkste konklusies van het onderzoek is wel het feit, dat iedere standaardisatiemethode impliciet een weging tussen criteria in zich bergt (paragraaf 5.3). Met deze impliciete weging (en in geval van het gebruik van niet-lineaire standaardisatiemethoden ook met de interne weging van paragraaf 5.4) wordt bij de expliciet door de beslissingsnemer vastgestelde wegingsvectoren geen rekening gehouden.

De in het kader van deze studie bestudeerde literatuur is op dit punt nagenoeg blanco (zie echter [34]). De benadering, die bij het ontwikkelen van de methode WATERLAB is toegepast, is gebaseerd op het door middel van trade-offs tussen de onderscheiden criteria bepalen van gestandaardiseerde gewichten in plaats van gestandaardiseerde effektscores zoals bij de traditionele methoden het geval is. In termen van bovenstaande opsomming worden de stappen 4 en 5 tezamen uitgevoerd, waarbij tevens rekening wordt gehouden met stap 3. WATERLAB is te combineren met elke gewenste evaluatiemethode: zij is te beschouwen als een gekombineerde preferentie- en standaardisatiemethode.



Figuur 5.2 Trade-off en Pareto-optimaliteit

De WATERLAB-methode maakt gebruik van de begrippen trade-off en Pareto-optimaliteit. Deze begrippen zullen worden toegelicht aan de hand van figuur 5.2. In deze figuur zijn de scores van de alternatieven  $A_1$  tot en met  $A_8$  op de criteria  $K_1$  en  $K_2$  uitgezet voor een hypothetisch MKE-probleem. Uitgaande van het principe "hoogst = best" zegt de ideaalpuntbenadering [2] dat een ideale combinatie van scores op criteria, een ideaal project dus, de maximale scores op  $K_1$  en  $K_2$  moet behalen; in de figuur is dat punt I. In de praktijk zal een verbetering in een criterium gepaard gaan met een verslechtering in een ander criterium; dit heeft tot gevolg, dat er projecten moeten bestaan, waarbij een verbetering in een richting niet mogelijk is zonder een verslechtering in een andere richting. Waren slechts de alternatieven  $A_1$  tot en met  $A_8$  van figuur 5.2 mogelijk, dan waren dat de alternatieven  $A_1$ ,  $A_3$ ,  $A_4$ ,  $A_5$ ,  $A_7$  en  $A_8$ ; zulke projecten heten Pareto-optimaal. De projecten  $A_2$  en  $A_6$  echter worden gedomineerd door respectievelijk  $A_3$ , en  $A_5$  en  $A_7$  samen: voor project  $A_7$  is zowel de score op  $K_1$  als die op  $K_2$  gunstiger dan de overeenkomstige score van  $A_6$ . In figuur 5.2 is dit te zien doordat  $A_7$  in het met stippellijnen aangegeven kwadrant rechts boven  $A_6$  ligt.



In bijna alle praktische gevallen bestaat er in principe een continu spectrum van alternatieven; de nutstheorie leert dat ook dan Pareto-optimale alternatieven bestaan; deze zullen op een continue lijn liggen, die bij de specificatie "hoogst=best" concaaf is naar de oorsprong. Zo een lijn wordt aangeduid met de term transformatiecurve of trade-off curve. In het geval dat geldt "hoogst=slechtst" zoals in het IODZH-voorbeeld, is de trade-off curve convex naar de oorsprong (vergelijk de figuren 5.3, 5.4 en 5.5).

Door voor elk paar criteria zulke trade-offs vast te stellen en in de ruimte uit te zetten ontstaat een meerdimensionaal lichaam op de rand waarvan de Pareto-optimale oplossingen dienen te worden gezocht. De dimensie hiervan is afhankelijk van de onderlinge onafhankelijkheid van de gebruikte criteria: waren in figuur 5.2 bijvoorbeeld de criteria "natuur" en "landschap" uitgezet dan hadden de assen niet onder een hoek van 90° mogen worden getekend gezien de grote mate van samenhang tussen beide criteria. Op dit onderwerp zal verder niet worden ingegaan; volstaan wordt met de opmerking dat meerdimensionale schalingstechnieken (zie paragraaf 3.7) en clusteranalyse ontwikkeld zijn juist om deze onderlinge afhankelijkheden op te sporen en het bestand aan criteria zover te reduceren (en om te vormen) dat een minimale onderlinge onafhankelijkheid van de uiteindelijk overblijvende criteria resteert.

### 5.5.3 Werkwijze

Bij een MKE met behulp van de methode WATERLAB worden door de analist trade-offs vastgesteld op grond van de gegevens uit de effectenmatrix (zie de vorige subparagraaf). Laat de trade-off voor criteria  $K_i$  en  $K_j$  gegeven worden door:

$$K_j = T_{ij}(K_i) \quad (\text{en } K_i = T_{ji}(K_j), \text{ dus } T_{ji} = T_{ij}^{-1}) \quad (5.7)$$

Hiermee wordt bedoeld, dat de trade-off  $T_{ij}$  een functie van waarden op de schaal van criterium  $K_i$  is, die bij elke waarde van  $K_i$  de bijbehorende Pareto-optimale waarde voor criterium  $K_j$  oplevert.

Van de beslissingsnemer wordt nu verwacht, dat hij/zij de gewenste positie op deze trade-off aangeeft, dat wil zeggen een punt  $(K_i^*, K_j^*)$  bepaalt, dat op de

trade-off ligt en voor hem/haar de beste combinatie van beide criteria weergeeft, gegeven het maximaal haalbare resultaat, zoals gegeven door de trade-off. Deze positie (het "standpunt" van de beslissingsnemer) levert het relatieve gewicht  $r_{ij}$  van criterium  $K_i$  ten opzichte van criterium  $K_j$  volgens:

$$r_{ij} = - \frac{dT_{ij}}{dK_i} (K_i^*, K_j^*) \quad (5.8)$$

In woorden: het relatieve gewicht van criterium  $K_i$  ten opzichte van criterium  $K_j$  wordt (op een minteken na) gelijk gesteld aan de richtingscoëfficiënt van de raaklijn<sup>1)</sup> aan de trade-off in het gewenste punt  $(K_i^*, K_j^*)$ .

Als motivatie van deze formule kan worden opgemerkt, dat bij het gewenste punt door zijn optimaliteit de marginale winst door toename van criterium  $K_i$  en het marginale verlies door afname van criterium  $K_j$  elkaar in evenwicht houden. De verhouding van deze marginale veranderingen nu is gelijk aan deze afgeleide. Het gewicht  $r_{ij}$  geeft dus aan hoeveel toename in  $K_i$  tegenover een afname van een eenheid  $K_j$  moet staan. Het minteken is geïntroduceerd om positieve gewichten te krijgen: doordat een toename van  $K_i$  een afname van  $K_j$  impliceert is de afgeleide zelf negatief.

De term relatief gewicht voor  $r_{ij}$  vindt mede zijn verklaring in het feit, dat  $r_{ij}$  de dimensies van  $K_i$  en  $K_j$  respecteert en derhalve automatisch rekening houdt met de schaalwaarden van  $K_i$  en  $K_j$ . Het is derhalve niet nodig een aparte standaardisatie van effektscores toe te passen: deling op elkaar van ongelijksoortige grootheden is zinvol waar dit niet geldt voor somming, zoals die veelvuldig bij MKE wordt toegepast.

In geval er meer dan twee criteria bij een evaluatie in het spel zijn moet voor elke combinatie van een tweetal criteria een trade-off worden gemaakt en door de beslissingsnemer worden beoordeeld. Op deze wijze ontstaat een matrix  $R$  van relatieve gewichten, die door haar wijze van ontstaan reciprook is, wat wil zeggen dat voor alle  $i$  en  $j$ :

---

1) Bewust is afgezien van een meerdimensionale formulering met vectorvelden en raakvlakken



$$r_{ij} = \frac{1}{r_{ji}} \quad (\text{en } r_{ii} = 1) \quad (5.9)$$

Door alle  $r_{ij}$  te bepalen kan ook gecontroleerd worden op inkonsistenties in de matrix R. Immers, doordat het relatieve gewicht  $r_{ij}$  van criterium  $K_i$  ten opzichte van  $K_j$  en dat van  $K_j$  ten opzichte van  $K_k$  ( $r_{jk}$ ), zijn bepaald, is ook een schatting  $r_{ij} \cdot r_{jk}$  van  $r_{ik}$  beschikbaar, welke kan worden vergeleken met  $r_{ik}$  zoals die rechtstreeks bepaald is. Een manier om een globale gewichtenvector W uit R af te leiden is meetkundige middeling per kolom van R, dus:

$$W = (w_1, \dots, w_J) \quad (5.10)$$

met

$$w_j = (r_{1j} r_{2j} \dots r_{Jj})^{1/J} \quad (5.11)$$

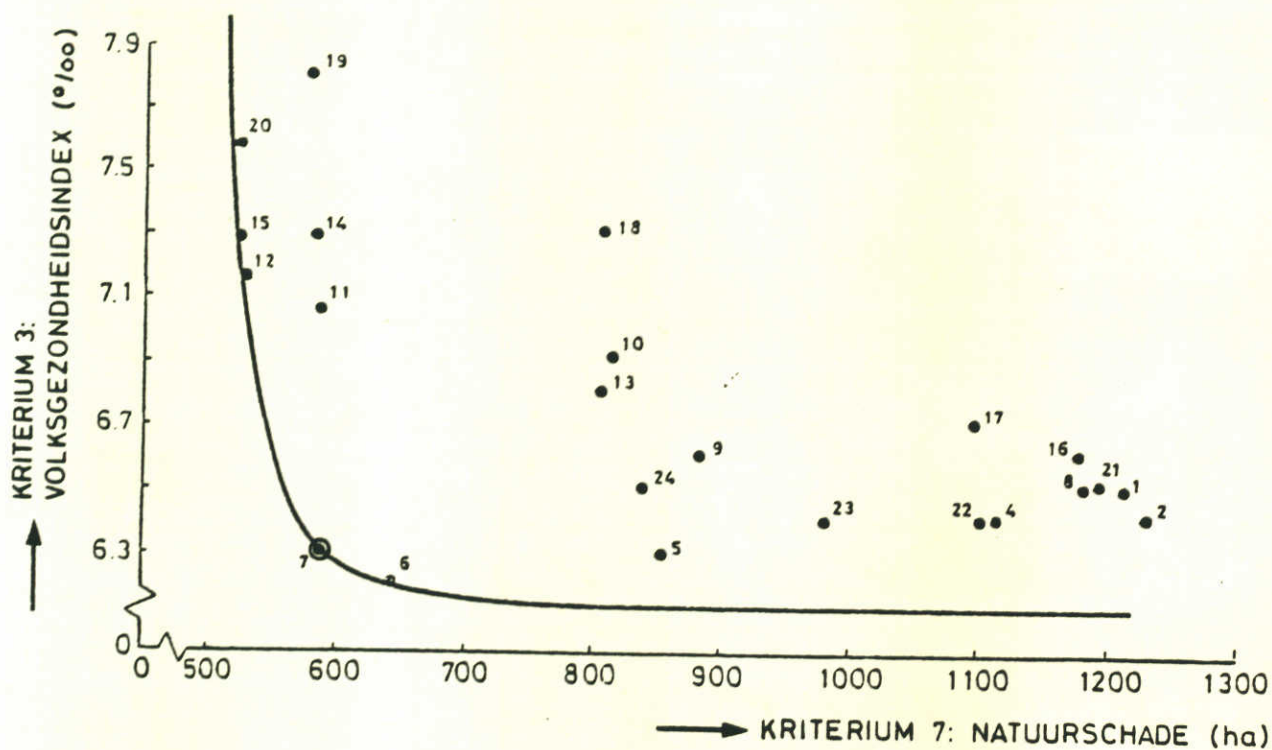
Als alternatief kan worden aangesloten bij het voorschrift dat Saaty hiervoor hanteert (zie paragraaf 3.5).

#### 5.5.4 Toepassing

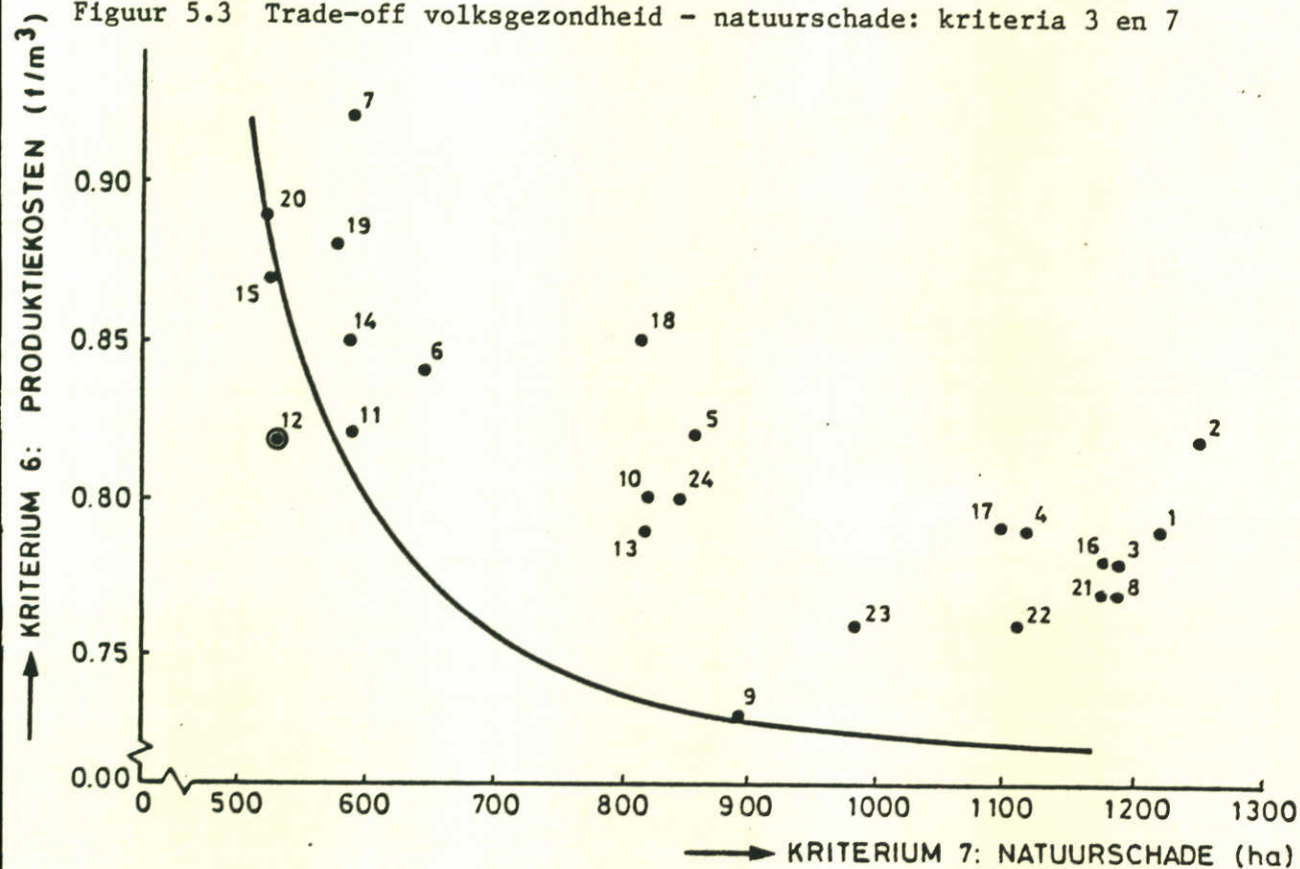
In deze subparagraaf wordt de methode WATERLAB toegepast op IODZH-gegevens. Hierbij wordt voor de eenvoud slechts uitgegaan van drie van de voornaamste criteria, te weten "volksgezondheid" (criterium 3), "produktiekosten" (criterium 6) en "natuurschade" (criterium 7); zie ook tabel 2.4.

Deze laten zich op drie wijzen tot paren criteria combineren en geven daardoor aanleiding tot de volgende trade-offs:

volksgezondheid	- natuurschade	(figuur 5.3)
produktiekosten	- natuurschade	(figuur 5.4)
volksgezondheid	- produktiekosten	(figuur 5.5)

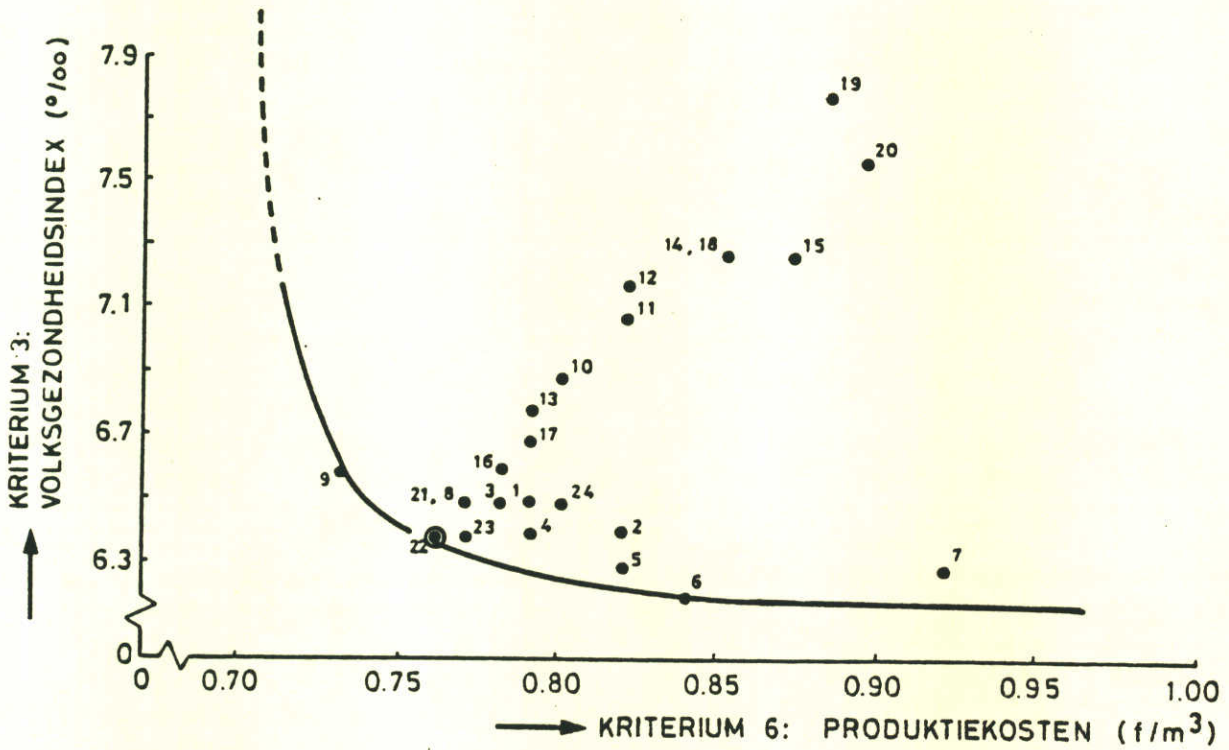


Figuur 5.3 Trade-off volksgezondheid - natuurschade: kriteria 3 en 7



Figuur 5.4 Trade-off productiekosten - natuurschade: kriteria 6 en 7





Figuur 5.5 Trade-off volksgezondheid - produktiekosten: kriteria 3 en 6

Op deze trade-offs zijn ook de meest gewenste posities (cirkels) aangegeven. De gewichtenmatrix R voor de gebruikte kriteria ziet er als volgt uit (de bij de diverse grootheden behorende eenheden zijn voor de eenvoud weggelaten).

kriteria	3 volksgezondheid	6 kosten	7 natuurschade
3	1	4,29	0,004
6	0,23	1	0,0007
7	250	1429	1

Tabel 5.2 Gewichtenmatrix voor drie kriteria van het IODZH-onderzoek

Standaardisatie van bovenstaande gewichten door deling door de som per rij (standaardisatiemethode TOTAAL) geeft een indruk van de onderlinge consistentie van de gekozen posities op de trade-offs (tabel 5.3).

kriteria	3 volksgezondheid	6 kosten	7 natuurschade
3	0,1889	0,8104	0,0008
6	0,1869	0,8125	0,0006
7	0,1488	0,8506	0,0006

Tabel 5.3 Gestandaardiseerde gewichtenmatrix voor drie criteria van het IODZH-onderzoek

Deze tabel geeft duidelijk aan, dat de posities op de trade-offs zeer consistent zijn gekozen: er ontstaan drie sterk op elkaar gelijkende rijen; als consistentiemaat zouden korrelatiekoefficienten kunnen worden gebruikt.

Zoals in subparagraaf 5.5.3 (zie (5.11)) is aangegeven kan uit tabel 5.2 door meetkundige middeling van de kolommen een gewichtenvector worden afgeleid; in dit geval wordt deze:

$$W = (3,86 \quad 18,3 \quad 0,014) \quad (5.12)$$

Ook hierbij zijn de betreffende eenheden weggelaten; deze zijn nu echter van geen belang meer: de gewichtenvector kan evengoed genormeerd worden door deling door de som van zijn elementen, wat neerkomt op een standaardisatie met de methode TOTAAL; het resultaat luidt:

$$W^* = (0,174 \quad 0,825 \quad 0,0006) \quad (5.13)$$

Dit resultaat kan ook direkt uit tabel 5.3 worden afgeleid. Het grote verschil tussen de gewichten is terug te voeren op het feit, dat de effektskores niet zijn gestandaardiseerd: zij moeten derhalve in samenhang worden gezien.



### 5.5.5 Kanttekeningen

- Uitgaande van criteria waarvoor geldt "hoe groter de score des te slechter" moeten trade-offs een convexe vorm hebben. Indien dit niet tot uitdrukking komt in de scores van de onderkende alternatieven, moet de konklusie worden getrokken dat belangrijke alternatieven niet zijn opgenomen in de evaluatie. Dit is bijvoorbeeld het geval met de trade-off kosten-natuurschade in het IODZH. Uit figuur 5.4 kan afgelezen worden dat er wellicht een alternatief bedacht zou kunnen worden dat qua eigenschappen tussen de alternatieven 9 en 12 in zou moeten liggen. Alternatief 13 voldoet daar niet geheel aan (zie de afstand tot de trade-off curve). In de eindfase van deze studie kunnen op deze wijze nog andere alternatieven worden geïdentificeerd. Het in de evaluatie betrekken van trade-offs heeft dus ook een controlefunctie (op de volledigheid van de alternatievenverzameling).
  
- In plaats van het op de trade-off aangeven met een punt van de meest gewenste positie kan dit ook gebeuren door het markeren van een stuk van de trade-off. Gestandaardiseerde gewichten worden dan niet meer in één getal maar in een range uitgedrukt. Bijvoorbeeld criterium i is  $1\frac{1}{2}$  à 3 keer zo belangrijk als criterium j. Op basis hiervan kunnen verschillende rangorden worden bepaald en wordt, in geval er meer visies bij de evaluatie betrokken zijn, de onderhandelingsruimte zichtbaar.
  
- Bij het toenemen van het aantal criteria neemt ook het aantal te konstrueren trade-offs toe (bij J criteria  $J(J-1)/2$  trade-offs). Gezien de grote mate van consistentie waarmee via trade-offs gewichten kunnen worden bepaald, lijkt het voldoende in plaats van alle combinaties van criteria alleen de belangrijkste in beschouwing te nemen. Het "belangrijkste" wordt hier bedoeld, dat combinaties van als belangrijk geziene (let wel: gewichten zijn nog niet vastgesteld) criteria met andere criteria worden beschouwd. Wellicht verdient het aanbeveling met een soort referentiekriterium te werken waarbij elk criterium tenminste met het referentiekriterium wordt vergeleken.

## 6 Toepassing

### 6.1 Inleiding

In het voorliggende hoofdstuk wordt verslag gedaan van een aantal experimenten op met name het IODZH-materiaal van hoofdstuk 2. Hierbij komen methoden aan de orde die in de hoofdstukken 3, 4 en 5 zijn behandeld. Een deel van deze methoden is geïmplementeerd in het computerprogramma MKEM. Dit programma wordt in de appendix besproken; dit hoofdstuk behandelt slechts de resultaten van uitgevoerde berekeningen.

In paragraaf 6.2 wordt het effect van standaardisatiemethode en MKE-methode op de ontstane rangorde gezien. Het materiaal van deze paragraaf wordt in paragraaf 6.3 gebruikt voor een clusteranalyse op grond van korrelatie tussen de gevonden rangorden. Een beschouwing van de invloed van gewichten wordt in paragraaf 6.4 ondernomen, terwijl in paragraaf 6.5 gekeken wordt naar de invloed van onzekerheden in het uitgangsmateriaal.

Met nadruk wordt op het feit gewezen dat het materiaal in dit hoofdstuk door zijn exemplarische karakter niet meer dan een illustratie geeft en dat getrokken konklusies noodgedwongen tentatief zijn.

### 6.2 Invloed van de keuze van standaardisatie- en MKE-methode

Bij de berekeningen ter illustratie van het effect van standaardisatie- en MKE-methoden zijn de combinaties als in tabel 6.1 toegepast. Het rekenwerk is verricht met het computerprogramma MKEM, waarover in de appendix meer informatie is gegeven.

Zoals uit tabel 6.1 blijkt zijn bij de vergelijking twee evaluatiemethoden betrokken namelijk de gewogen sommeringsmethode (SOM; paragraaf 3.3) en de konkordantie-analyse (KON; paragraaf 3.6). Deze twee methoden worden in de praktijk frekwent gebruikt.

De gewogen sommeringsmethode wordt toegepast uitgaande van zowel de kardinale als de twee ordinale (zie subparagraaf 4.2.1) effectenmatrices uit hoofdstuk 2. Op de kardinale effectenmatrix zijn vier bij het onderzoek betrokken



type data	standaardisatie methode	evaluatiemethode	
		SOM	KON
ordinaal	KLASSE	1)	4)
	SAATY	2)	5)
	RANGORDE	3)	6)
kardinaal	MAX	7)	6)
	MIN	8)	
	MAXMIN	9)	
	NUTS	10)	

Tabel 6.1 Vergeleken evaluatie- en standaardisatiemethoden

standaardisatiemethoden toegepast namelijk MAX, MIN en MAXMIN (subparagraaf 4.2.2) en NUTS (subparagraaf 4.2.3).

Ook de konkordantie-analyse is zowel met de kardinale als met de twee ordinale effektenmatrices doorgerekend. Uitgegaan is van de variant waarbij per criterium wordt vastgesteld of de skore van een alternatief groter is dan de skore van een ander alternatief zonder dat de "afstand" tussen deze skores in de beschouwingen is betrokken. Als gevolg hiervan is de uitkomst van deze konkordantie-analyse onafhankelijk van de wijze van standaardisatie van de kardinale gegevens. Bij de bepaling van de netto-diskordantie-dominantie-indices is deze afhankelijkheid wel aanwezig (vergelijk paragraaf 3.6). In dit onderzoek is deze methode echter niet toegepast. De bepaling van netto-konkordantie-dominantie-indices op basis van kardinale data levert hetzelfde resultaat op als die op basis van de rangorde-effektenmatrix (kombinatie 6).

Resumerend is voor de vergelijking het volgende uitgangsmateriaal aanwezig:

- i. drie effektenmatrices (1 kardinaal: tabel 2.1; 2 ordinaal: tabellen 2.2 en 2.3) bestaande uit 24 alternatieven en 10 criteria;

- ii. één gewichtenmatrix (tabel 2.4) bestaande uit 6 visies en 10 criteria; en
- iii. tien combinaties van standaardisatie- en MKE-methoden.

Ten behoeve van vergelijking van de rangorderresultaten zijn volgens het volgende systeem punten toegekend aan de rangorde van de alternatieven zoals die voor de afzonderlijke visies zijn bepaald:

rangorde	punten
1- 5	4
6-10	3
11-15	2
16-20	1
21-24	0

Door sommatie van de zes punttoekenningen die per alternatief gevonden worden, is een totaalscore berekend. Skoort een alternatief 6 keer in de rangordeklasse 1-5 dan is de totaalscore maximaal namelijk 24. De minimale score is 0.

In tabel 6.2 is een overzicht gegeven van de totaalscores van de 24 alternatieven voor de 10 combinaties van standaardisatie- en MKE-methoden die resulteren uit bovenstaande aanpak; hierbij zijn de resultaten als volgt in vier klassen ingedeeld:

puntentotaal	klasse
19-24	++
13-18	+
7-12	-
0-6	--

Een vrij grote overeenkomst tussen deze scores is zichtbaar, maar de alternatieven 11, 14, 15, 19, 21 en 22 krijgen afhankelijk van de gebruikte methode sterk verschillende uitkomsten.



alternatief	type data	ordinaal						kardinaal			
	MKE-methode	SOM			KON			SOM			
	stand.-methode	klas SAATY rang			klas SAATY rang			MAX	MIN	MAXMIN	NUTS
	nummer methode	1	2	3	4	5	6	7	8	9	10
1	--	--	--	-	--	-	--	--	--	--	--
2	--	--	--	--	--	--	--	--	--	--	--
3	-	-	-	-	--	-	--	--	-	--	--
4	+	-	+	+	+	+	-	-	+	+	+
5	++	+	++	++	++	++	+	+	++	++	++
6	++	++	++	++	++	++	++	++	++	++	++
7	++	++	++	++	++	++	++	++	++	++	++
8	+	+	+	+	-	+	-	-	+	-	-
9	++	+	++	++	+	++	+	++	++	++	++
10	+	-	+	+	+	+	+	+	-	+	+
11	-	-	-	-	-	-	++	++	+	+	+
12	+	++	+	+	++	+	++	+	++	++	+
13	-	--	-	-	-	-	+	-	+	++	++
14	--	-	--	--	-	--	++	+	-	-	-
15	-	++	-	-	+	-	++	++	+	-	-
16	-	-	-	-	--	-	--	--	--	--	--
17	-	--	-	-	-	-	--	--	--	-	-
18	--	--	--	--	--	--	-	--	--	-	-
19	--	+	--	-	+	-	+	+	-	--	--
20	+	++	+	+	++	+	++	++	+	-	-
21	+	+	+	+	-	+	--	-	-	--	--
22	++	+	++	++	+	++	-	-	+	+	+
23	-	--	-	-	-	-	-	--	-	+	+
24	+	-	+	+	-	+	-	-	-	-	+

Tabel 6.2 Totaalscores van de IODZH-alternatieven voor verschillende methoden

### 6.3 Vergelijking

De rangorden als uitkomst van de verschillende combinaties van standaardisatie- en evaluatie-methoden zijn onderling vergeleken via een regressieanalyse. Als maat voor de overeenstemming is het kwadraat van de korrelatiekoefficiënt gehanteerd ( $r^2$ ). In tabel 6.3 zijn hiervan de resultaten opgenomen.

type data		ordinaal						kardinaal			
MKE-methode		SOM			KON			SOM			
stand.-meth.	nr.	KLAS	SAATY	RANG	KLAS	SAATY	RANG	MAX	MIN	MAXMIN	NUTS
		1	2	3	4	5	6	7	8	9	10
SOM	KLASSE	1	0,48		0,90	0,76	0,69	0,22	0,29	0,67	0,57
	SAATY	2	1	0,48	0,54	0,70	0,53	0,43	0,72	0,55	0,15
	RANG	3		1	0,96		0,95	0,06	0,22	0,63	
KON	KLASSE	4			1	0,75	0,78	0,13	0,25	0,61	0,41
	SAATY	5				1	0,61	0,51	0,65	0,73	0,44
	RANG	6					1	0,12	0,26	0,69	0,36
SOM	MAX	7						1	0,78	0,55	0,43
	MIN	8							1	0,61	0,29
	MAXMIN	9								1	0,71
	NUTS	10									1

Tabel 6.3 Korrelatie van resulterende rangorden



De hoogte van de  $r^2$  kan als volgt worden geïnterpreteerd:

$0 < r^2 < 0,5$  de vergeleken rangorden lijken niet op elkaar

$0,5 < r^2 < 0,75$  er is duidelijk enige gelijkheid te onderkennen in de vergeleken rangorden

$0,75 < r^2 < 1$  de vergeleken rangorden lijken sterk tot volkomen op elkaar

Teneinde het trekken van conclusies te vergemakkelijken is van tabel 6.3 een ruimtelijke presentatie gemaakt (figuur 6.1). De mate van overeenstemming tussen twee methoden is in deze figuren weergegeven als afstand tussen de verschillende combinaties. De relatie tussen deze afstand  $d$  en de waarde van  $r^2$  is als volgt gekozen:

$$d = 15 (1-r^2) \quad (\text{in cm})$$

Resulteert de vergelijking van twee combinaties in een  $r^2$  van 0,80 dan liggen deze combinaties dus 3 cm uit elkaar. Naarmate een  $r^2$  kleiner is en de uitkomsten van de vergeleken methoden dus minder op elkaar lijken wordt de afstand  $d$  groter. Figuur 6.1 is met behulp van een meerdimensionale schalingsmethode tot stand gekomen.

Herhaald wordt dat het rekenvoorbeeld slechts illustratieve betekenis kan hebben. De gepresenteerde rekenresultaten laten zich in grote lijnen als volgt samenvatten:

- i. Er is een groot verschil tussen de resultaten die bereikt worden op basis van ordinale en die op basis van kardinale invoergegevens.
- ii. De berekende rangorden op basis van de twee ordinale effectenmatrices (klassen en rangorden) komen goed overeen (1 en 3, 4 en 6).
- iii. Een goede mate van overeenstemming in de waarderingen is gevonden tussen SOM en KON, mits de bewerking van de uitvoergegevens dezelfde is (combinaties 1 en 4, 2 en 5, 3 en 6).



▲ KON-methoden

■ ordinale SOM-methoden

\* kardinale SOM-methoden

Figuur 6.1 Grafische vergelijking van evaluatie- en standaardisatiemethoden



iv. Grote verschillen zijn gevonden tussen de gebruikte standaardisatiemethoden. Slechts MAX en MIN (7 en 8) en in iets mindere mate MAXMIN en NUTS (9 en 10) stemmen onderling in beperkte mate overeen ( $r^2 \approx 0,75$ ).

Niettegenstaande de illustratieve waarde van het voorbeeld zal het duidelijk zijn, dat bovenstaande bevindingen tot grote voorzichtigheid nopen bij het toepassen van MKE.

#### 6.4 Invloed van gewichten

Om een indruk te verkrijgen van de invloed die met gewichtenvectoren kan worden uitgeoefend op de uiteindelijke rangvolgorde van de alternatieven zijn tijdens het onderzoek voor vele combinaties van standaardisatie- en MKE-methoden rangvolgorden berekend op grond van de (zes) gewichtenvectoren uit de gewichtenmatrix van tabel 2.4.

Voor elk van deze combinaties van methoden ontstaat zo een tabel van resulterende rangorden per gewichtenvector. De tabellen 6.4 en 6.5 geven hiervan voorbeelden: respectievelijk hebben zij betrekking op het gebruik van ENTROPIE (zie subparagraaf 4.2.4) en de combinatie van SAATY (zie subparagraaf 4.2.1 en paragraaf 3.5) met SOM (gewogen sommeren) (zie paragraaf 3.3). Naast de rangvolgorde van de alternatieven per gewichtenvector is ook de spreiding hierin, geïnterpreteerd als het grootste absolute verschil in rangnummer per alternatief, als kolom opgegeven.

Per (samengestelde) methode kan een maat (gemiddelde relatieve spreiding) voor de gevoeligheid voor gewichten bepaald worden. De gemiddelde relatieve spreiding is de som van de spreiding per alternatief gedeeld door de maximaal mogelijke spreiding (bij  $n$  alternatieven:  $n \cdot (n-1)$ ).

Uiteraard hangt deze maat niet alleen van de methode af, maar ook van de gebruikte alternatieven-effektenmatrix, terwijl vanzelfsprekend ook maar een keuze uit de mogelijke gewichtenvectoren is gemaakt: grofweg kan worden gezegd, dat elke methode bij een "juiste" keuze van gewichtenvectoren elke mogelijke rangvolgorde van de Pareto-optimale alternatieven kan opleveren; zeker kan elk Pareto-optimaal alternatief als beste skoren.

alternatief	visie						spreiding
	A	B	C	D	E	F	
1	23	23	23	23	23	23	0
2	22	22	24	24	24	24	2
3	20	21	22	22	21	22	2
4	15	16	18	18	17	18	3
5	9	12	13	13	13	13	4
6	3	6	7	7	5	8	5
7	1	1	8	8	10	7	9
8	18	19	19	19	19	19	1
9	12	14	11	11	8	14	6
10	10	9	9	9	9	9	1
11	2	2	4	4	2	4	2
12	6	5	1	1	1	1	5
13	11	10	10	10	11	10	1
14	4	3	5	5	4	5	2
15	7	7	2	2	3	2	5
16	19	20	20	21	22	20	3
17	17	17	16	17	18	16	2
18	13	11	14	14	14	11	3
19	5	4	6	6	7	6	3
20	8	8	3	3	6	3	5
21	24	24	21	20	20	21	4
22	21	18	17	16	16	17	5
23	16	15	15	15	15	15	1
24	14	13	12	12	12	12	2

gemiddelde relatieve spreiding 13%

Tabel 6.4 Rangorden voor verschillende gewichtenvectoren bij ENTROPIE



alternatief	visie						spreiding
	A	B	C	D	E	F	
1	18	23	14	14	14	22	9
2	15	20	23	24	23	24	9
3	16	21	12	12	6	20	15
4	19	18	18	17	15	18	4
5	6	9	6	6	12	9	6
6	5	5	1	1	5	4	4
7	2	4	2	3	8	5	6
8	12	15	8	9	3	15	12
9	10	13	10	8	2	12	11
10	14	10	21	21	18	13	11
11	9	7	17	15	21	7	14
12	1	1	4	4	10	2	9
13	23	14	22	22	20	14	9
14	11	8	19	16	22	8	14
15	3	2	5	5	13	3	11
16	17	22	13	13	7	21	15
17	22	24	20	20	17	23	7
18	24	17	24	23	24	17	7
19	8	6	11	11	19	6	13
20	4	3	3	2	9	1	8
21	13	16	9	10	4	16	12
22	7	11	7	7	1	10	10
23	21	19	16	19	11	19	10
24	20	12	15	18	16	11	9

gemiddelde relatieve spreiding 41%

Tabel 6.5 Rangorden voor verschillende gewichtenvectoren bij SAATY met SOM

Niettegenstaande al deze beperkingen is het toch illustratief deze gemiddelde relatieve spreidingen te vergelijken. Bij beschouwing van de tabellen 6.4 en 6.5 blijkt een zeer groot verschil tussen de waarden bij ENTROPIE en SAATY + SOM (respektievelijk 13% en 41%). Hierbij betrekken van de verder nog uitgevoerde berekeningen levert een variatie van de al genoemde 13% tot 48% voor SAATY + KON met een aantal waarden rond de 20% en enige boven de 40%.

Deze gevoeligheid voor gewichten blijkt bij beschouwing van het materiaal met name een gevolg te zijn van de toegepaste standaardisatie: zo blijken de methoden TOTAAL, MAX en MIN (zie subparagraaf 4.2.2) weinig gevoelig te zijn voor gewichtenspecificatie, terwijl MAXMIN (subparagraaf 4.2.2) en SAATY (subparagraaf 4.2.1) grote variaties in de rangorden te zien geven. Toepassing van ENTROPIE (subparagraaf 4.2.4) blijkt (tabel 6.4) erg weinig variatie op te leveren.

Wat moet nu worden gekonkludeerd uit bovenstaande informatie? De bedoeling van het geven van gewichten is het goed tot uitdrukking brengen van de visies van verschillende beslissingsnemers en het is daarom ook juist dat verschillende rangorden het resultaat zijn van de gebruikte visies. De vraag is echter hoeveel variatie het gevolg dient te zijn van de gewichtenspecificatie. Hierop is vooralsnog geen exact antwoord te geven: uitspraken hierover kunnen slechts gedaan worden aan de hand van onderzoek naar de perceptie van mensen die op inherent subjektieve wijze hun visie vertalen in een gewichtenvector. Gekonkludeerd kan worden dat de gevonden spreidingen op zich niet verontrustend kunnen worden genoemd.

Een andere kwestie is of het hanteerbaar is dat de gevonden spreidingen bij de gebruikte methoden onderling zoveel (van 13% tot 48%) moeten verschillen. Op deze laatste vraag past een krachtig neen. Voor de beslissingsnemer moet het duidelijk zijn wat hij/zij bewerkstelligt met een bepaald patroon van invullen van gewichtenvectoren. Op nog een andere wijze dan het materiaal van paragraaf 6.3 aandraagt, ondersteunt bovenstaande de vaak geuite aanbeveling bij een MKE geen toevlucht te zoeken bij slechts één methode maar in een gevoeligheids-onderzoek ook andere methoden toe te passen.



In tabel 6.6 zijn de minimaal en maximaal per alternatief en gewichtenvector behaalde positie op de rangorden samengebracht. Een globale blik op het materiaal leert reeds dat bijna alle alternatieven met een "geschikte" gewichtenvector een plaats bij de eerste acht of bij de laatste zes kan worden gegeven; bij de toch niet zeer extreme gewichtenvectoren van tabel 2.4 kunnen toch maar liefst zeven alternatieven als beste uit de MKE komen en eveneens zeven als slechtste (zie de laatste dubbele kolom)!

Op een soortgelijke wijze als in de tabellen 6.4 en 6.5 is in tabel 6.6 een gemiddelde relatieve spreidingsmaat vastgesteld: deze keer per gewichtenvector waarbij de spreiding door de methoden wordt veroorzaakt en opnieuw gemiddeld is over de alternatieven (kolomsgewijs). Alleen naar deze getallen te oordelen lijken de gewichtenvectoren B en F (21% en 23%) en A, C, D en E (50% tot 53%) onderling sterk op elkaar, waarbij tussen beide groepen een grote kloof zichtbaar is.

De interpretatie van deze gegevens is niet eenvoudig. De gevonden maat stelt namelijk de gevoeligheid van de rangvolgorde per gewichtenvector voor in geval van variatie van MKE-methode. De belangrijkste konklusie op grond van tabel 6.6 is dat bij één effektenmatrix en één gewichtenvector alleen door variatie in de MKE-methode reeds variaties tot een gemiddelde (!) van de helft van de schaal mogelijk zijn; hiermee wordt de afhankelijkheid van de gebruikte MKE-methode wel overduidelijk aangetoond.

Verder blijkt dat de gewichtenvectoren B en F beide veel extreme posities op de rangordeschaal opleveren die afwijken van de posities die overeenkomen met die van de gewichtenvectoren A, C en D; gewichtenvector E geeft juist veel extreme posities in andere richting dan de bij de gewichtenvectoren B en F geconstateerde.

alternatief	VISIES						TOTAAL
	A	B	C	D	E	F	
	min max	min max	min max	min max	min max	min max	min max
1	13 23	21 23	9 23	10 23	10 23	21 23	9 23
2	15 24	20 24	17 24	24 24	19 24	24 24	15 24
3	9 20	20 22	9 22	9 22	6 21	20 22	6 22
4	5 19	15 18	5 18	5 18	4 17	13 18	4 19
5	2 10	6 12	2 13	2 13	3 13	3 13	2 13
6	1 5	1 8	1 7	1 7	2 6	1 8	1 8
7	1 5	2 6	2 8	2 8	8 19	2 7	1 19
8	3 18	12 19	4 19	3 19	1 19	11 19	1 19
9	2 12	9 14	4 11	2 11	1 8	7 15	1 15
10	7 19	9 14	9 21	8 21	5 18	8 16	5 21
11	2 21	2 7	3 22	3 21	2 21	4 10	2 22
12	1 10	1 5	1 17	1 14	1 17	1 5	1 17
13	7 23	10 14	8 22	6 22	7 20	10 15	6 23
14	4 23	3 9	5 24	5 23	7 23	5 13	3 24
15	3 17	2 7	3 20	2 19	3 21	2 6	2 21
16	11 21	19 23	11 20	11 22	7 22	20 23	7 23
17	14 22	17 24	12 23	12 23	11 22	16 23	11 24
18	13 24	11 17	14 24	14 23	14 24	11 17	11 24
19	5 22	4 8	6 22	6 22	7 24	6 10	4 24
20	4 12	3 8	3 13	2 11	6 22	1 4	1 22
21	8 24	16 24	8 21	8 20	4 20	16 21	4 24
22	4 22	11 22	4 17	4 17	1 16	9 17	1 22
23	13 24	15 19	11 22	11 22	10 20	15 19	10 24
24	9 22	11 17	6 17	6 19	12 21	7 18	6 22

50%

21%

51%

50%

53%

23%

67%

gemiddelde relatieve spreiding

Tabel 6.6 Spreiding in rangorden voor verschillende gewichtsvectoren bij variatie van methoden



Bepaling van korrelatiecoëfficiënten (zoals in paragraaf 6.3 voor MKE-methoden reeds werd uitgevoerd) is de geschikte weg om de gewichtenvectoren te klassificeren, maar bovenstaande bevinding dat de gewichtenvectoren B en F sterk gerelateerd zijn wordt ondersteund door te kijken naar tabel 2.4 waaruit blijkt, dat zij vooral door hun natuurvriendelijke accent (kriterium 7) soortgelijk zijn. Vermoedelijk is de nadruk op het kriterium "natuurschade" tevens oorzaak van de verschillen in de bovenvermelde spreidingsmaten: bij gewichtenvectoren B en F is duidelijk sprake van een wat extreme -of eenzijdige- visie, waaronder de gevonden rangorden niet zeer afhankelijk kunnen zijn van de gebruikte methoden. Bij een zeer extreme visie is feitelijk geen sprake meer van multi-kriteria evaluatie!

Tenslotte zij nog gewezen op de globale gemiddelde relatieve spreidingsmaat, die is afgeleid uit de laatste dubbele kolom van tabel 6.6. Deze toont met haar waarde van 67% aan dat door de uitgevoerde variaties in gewichtenvectoren en MKE-methoden gezamenlijk maar liefst tweederde van de schaal wordt bestreken, en dit allemaal uitgaande van een en dezelfde effectenmatrix.

#### 6.5 Invloed van onzekerheden

Hoeveel aandacht ook is besteed aan het consciëntieus vaststellen van de effectenmatrix, in de praktijk zal zij met grote of kleine onnauwkeurigheden zijn behept. Is informatie beschikbaar over de te verwachten orde van grootte van deze onnauwkeurigheden dan kan door middel van een gevoeligheidsonderzoek, waarbij de effectskores naar gelang hun onnauwkeurigheid worden gevarieerd, worden nagegaan in hoeverre de op grond van de effectskores gevonden rangvolgorder vaststaat ondanks de onnauwkeurigheid van deze skores. In deze paragraaf wordt aan de hand van de IODZH-gegevens een dergelijk gevoeligheidsonderzoek uitgevoerd.

Uitgangspunt van het onderzoek is de IODZH-effektenmatrix van tabel 2.1. Door middel van een speciaal hiervoor geschreven (SIMULA67-)programma werden alle effectskores verstoord met een willekeurig bedrag, dat in het te bespreken voorbeeld Gaussisch verdeeld was met verwachtingswaarde nul en standaarddeviatie 2% van de oorspronkelijke skore ( $N(0,2\%)$ -verdeling). Hiervoor werd door toepassing van het (FORTRAN4-)programma MKEM (zie appendix) de preferentie-

volgorde bepaald, in dit geval door toepassing van de gewogen sommeringsmethode (SOM; zie paragraaf 3.3) na standaardisatie met behulp van TOTAAL (zie subparagraaf 4.2.2) en onder gebruikmaking van gewichtenvector C uit tabel 2.4. Telkens opnieuw uitgaande van de oorspronkelijke effectenmatrix werd dit 100 maal uitgevoerd, waarna voor elk alternatief het aantal malen werd geteld, dat het een bepaalde plaats in de preferentierangorde innam.

Het resultaat is weergegeven in tabel 6.7. De kolommen hierin korresponderen met de alternatieven, de rijen met de posities in de eindrangschikking: zo heeft alternatief 3 in 2 gevallen in de eindrangschikking op de 16-de plaats gestaan, 4 keer als 17-de.....18 keer als 23-ste.

Een blik op tabel 6.7 leert onmiddellijk dat de toegepaste verstoring grote gevolgen heeft voor de eindvolgorde: zo zijn er liefst 9 alternatieven (de nummers 1, 3, 4, 8, 16, 17, 21, 22 en 23) die een 16-de plaats bereiken en scoort alternatief 3 alle posities van de 16-de tot en met de 23-ste plaats. Niet alle alternatieven zijn even gevoelig voor de opgelegde onnauwkeurigheid: in feite is alternatief 2 is niet van de 24-ste plaats afgeweest.

Het experiment is herhaald met  $N(0, \frac{1}{2}\%)$ -verdeelde verstoringen; hiervan worden de resultaten in tabel 6.8 gegeven. Men mag verwachten, dat de alternatieven nu een nagenoeg konstante plaats in de eindrangschikking innemen: de kans dat een score nu meer dan  $1\frac{1}{2}\%$  wordt verstoord is nu slechts 0,13%! In werkelijkheid blijkt echter, dat nog steeds een flinke verstoring optreedt van de eindrangschikking; er zijn nu meer alternatieven, die een vaste plaats krijgen toegewezen, maar de verschillen in behaalde posities blijken nog steeds vrij groot te zijn (zie bijvoorbeeld alternatief 21).

Zelfs in een experiment waarbij  $N(0, 0,2\%)$ -verdeelde verstoringen werden toegepast zijn nog steeds alternatieven te onderkennen die een meerduidige plaats krijgen toegewezen; in dit geval zijn dit er echter duidelijk minder.



rangorde positie	alternatieven																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1					98						2													
2					2						12	73		13										
3							1				33	22		2	37					5				
4							6		2		38	3		6	38					7				
5							22		3		11			18	6					40				
6					1		30		16		4			23	3					23				
7					2		28		16		2			27	2					4	19			
8					8		9		43					18	1					17	4			
9					35		4		11	1				6						41	2			
10					45				9	7				5						34				
11					7				45					46						2				
12					2				47					47					1	2				1
13														2					39					59
14				3															58					39
15				50				11									4	2		1	9	22		1
16	1		2	22				19								2	9			2	20	23		
17	2		4	17				14								1	18			4	22	18		
18			10	2				23								6	20			5	18	16		
19	5		10	1				11								11	16			13	20	13		
20	13		17	3				12								10	17			16	8	4		
21	19		18	1				7								23	8			21	1	2		
22	25		21					2								25	7			16	2	2		
23	35		18	1				1								22	1			22				
24	100																							

Tabel 6.7 Verdeling van posities in de eindrangschikking met 2% verstoring

rangorde positie	alternatieven																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1					100																			
2											100													
3										38			62											
4										62			38											
5							47					13						40						
6							38					19						43						
7							15					68						17						
8								100																
9					31													69						
10					69													31						
11										62		38												
12										38		62												
13																		17						83
14																		83						17
15					92					1												1	6	
16					7					7												27	59	
17					1					29								1				41	28	
18										59								4				30	7	
19										4								93				2	1	
20																		23	2			35		
21																		27				35		
22																		46				22		
23																		4				6		
24																								100

Tabel 6.8 Verdeling van posities in de eindrangschikking met  $\frac{1}{2}$ % verstoring

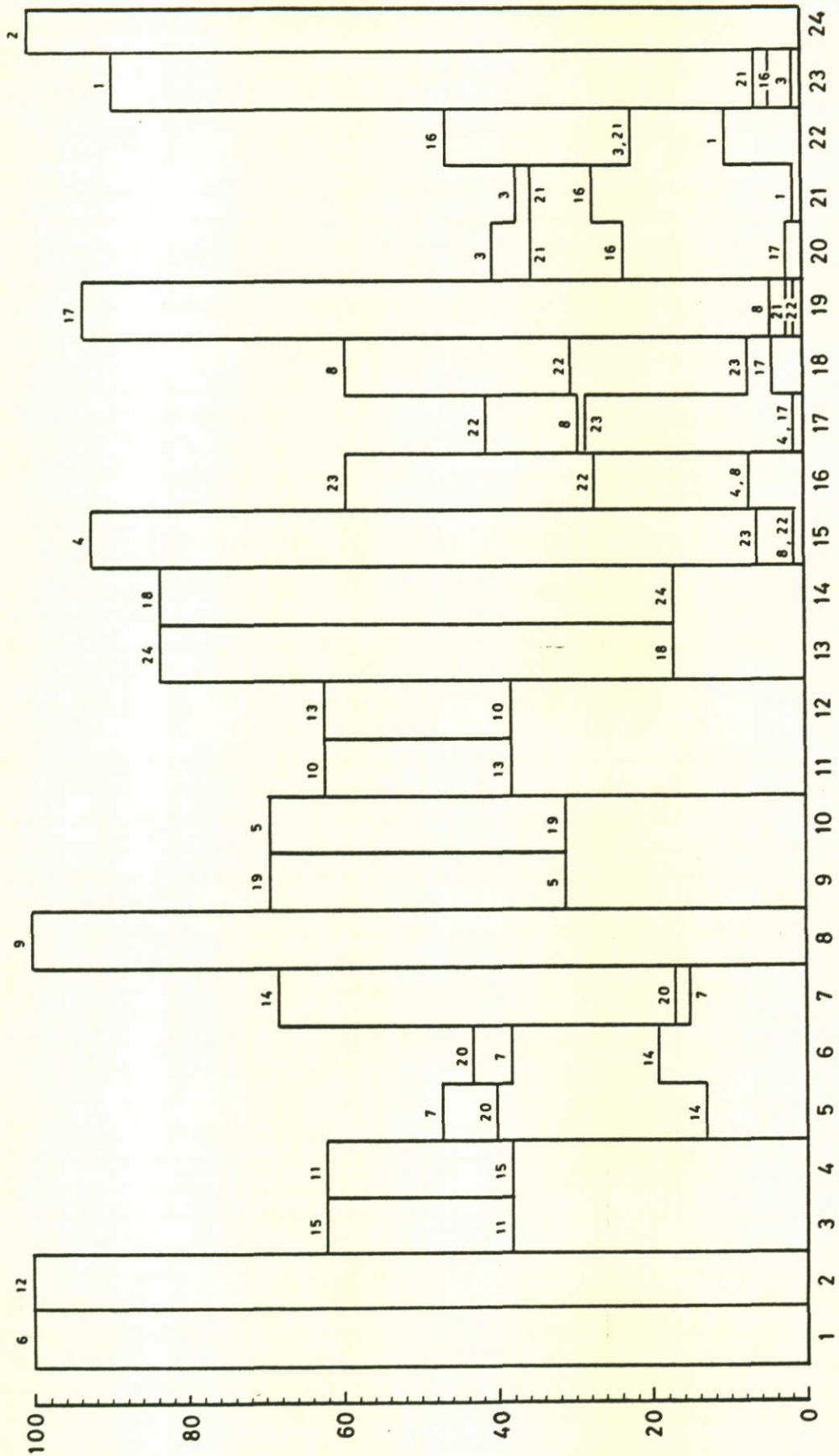


De resultaten van tabel 6.8 zijn tevens weergegeven in figuur 6.2, waarin voor alle alternatieven het aantal malen dat een bepaalde positie werd behaald is uitgezet tegen deze positie. Zo ontstaan 24 grafieken die over elkaar heen zijn geprojecteerd en die onderscheiden kunnen worden aan de hand van het (alternatief-)nummer, waarvan zij zijn voorzien. In figuur 6.2 komen hoge pieken (bij de alternatieven 6, 12, 9, 4, 17, 1 en 2) overeen met alternatieven die door de aangebrachte verstoringen niet of bijna niet van positie zijn veranderd. De tweetallen 11 en 15, 19 en 5, 10 en 13 en 24 en 18 komen op precies twee posities voor waardoor hun positie vrijwel vast ligt: alleen onderling kunnen zij nog van plaats wisselen. Andere alternatieven vertonen meer variatie in positie, zoals de nummers 22, 21 en 16.

De konklusie op grond van dit experiment moet luiden dat de gebruikte standaardisatiemethode (TOTAAL) in combinatie met de evaluatiemethode (SOM) althans voor de gehanteerde gewichtenvector (C) zeer gevoelig is voor onzekerheden in de effectenmatrix: de gehanteerde verstoringen zijn naar alle waarschijnlijkheid een orde kleiner dan de in werkelijkheid optredende onzekerheidsmarges.

Het blijkt al met al van belang te zijn een gevoeligheidsonderzoek uit te voeren waarmee de effecten van onzekerheid in het uitgangsmateriaal kunnen worden getraceerd. Zo'n onderzoek kan worden uitgevoerd volgens de lijnen van het experiment in deze paragraaf, waarmee niet wordt gezegd dat dit de enig mogelijke methode is noch dat een separaat onderzoek naar andere invloeden, zoals die van standaardisatiemethode, gewichtenkeuze en evaluatiemethode in zijn schaduw moet staan. Bij de presentatie van resultaten van een MKE zegt een prioriteitenrangschikking van de onderscheiden alternatieven pas iets als bekend is welke de nauwkeurigheid van de geproduceerde resultaten is in samenhang met de veronderstelde precisie van het uitgangsmateriaal.

aantal maal dat  
alternatief i  
op positie voorkomt



positie in de rangorde

Figuur 6.2 Verdeling van posities in de eindrangschikking met  $\frac{1}{2}\%$  verstoring



## 7 Samenvatting

### 7.1 Algemeen

In deze paragraaf zullen in het kort de in dit deel III behandelde zaken nogmaals de revue passeren, waarbij kritische kanttekeningen worden gemaakt. Tussen haakjes wordt eventueel verwezen naar de plaats waar een en ander behandeld is.

#### MKE-methoden

In dit deel worden zeven MKE-methoden uitgewerkt, die als hoofdvarianten van MKE-methoden gekenschetst kunnen worden. Hiervan werken 6 methoden met een eenvormige effecten- of effectiviteitenmatrix (d.w.z. alleen kardinale of alleen ordinale gegevens):

1. gewogen somming (goals achievement matrix)
2. gewogen somming (verwachtingswaarde)
3. permutatie (Paelinck)
4. eigenwaarde (Saaty)
5. konkordantie (Electre)
6. meerdimensionale schaling (OGEM).

Als zevende methode is een methode gepresenteerd die in staat is een gemengde effectenmatrix te verwerken (Evamix-benadering).

Vrijwel al deze methoden hebben varianten die in staat zijn andersoortige informatie te verwerken. Met name de groep konkordantiemethoden kent veel varianten om bijvoorbeeld kwalitatieve of stochastische informatie te verwerken.

De methode Saaty is een speciaal geval. Deze methode maakt in tegenstelling tot de andere methoden geen expliciet gebruik van een effectentabel. Scores worden verkregen door aan de beslisser(s) te vragen in welke mate een bepaald alternatief een ander alternatief naar zijn/hun inzicht domineert. Dit kan gebeuren met een effectenmatrix als achtergrond, maar dit is niet persé nood-

zakelijk. De methode Saaty kan dus alleen toegepast worden in een situatie waarin interactie met beslissers mogelijk is. Bij de andere MKE-methoden behoeft dit niet het geval te zijn.

Uit experimenten met de IODZH-gegevens blijkt dat er sprake is van methode-onzekerheid. Bij gebruik van eenzelfde effekten- of effectiviteitenmatrix en eenzelfde gewichtenvector leveren verschillende MKE-methoden verschillende rangordes op. Deze methode-onzekerheid neemt af naarmate de criteria onderling een zekere korrelatie bevatten, hetgeen logisch is omdat in dat geval in mindere mate sprake is van multikriteria evaluatie (zie ook [2]). Verschillen in rangschikking tussen verschillende MKE-methoden valt voornamelijk toe te rekenen aan de mate waarin de methode impliciete gewichten toekent aan de effektscores. Hierdoor verschillen met name de rangordes verkregen met MKE-methoden die bij de bewerking van de gegevens gebruik maken van het entropie-koncept (zie subparagraaf 4.2.4) sterk van rangordes verkregen met methoden die geen gebruik maken van het entropie-koncept. Om redenen van methode-onzekerheid is het aan te bevelen bij een MKE gebruik te maken van meerdere MKE-methoden. Overigens is methode-onzekerheid moeilijk los te zien van het probleem van standaardisatie (zoals in feite ook het geval is bij hierboven gesignaleerde onzekerheid bij het gebruik van het entropie-koncept).

### Standaardisatie

Standaardisatie kan worden gezien als voorbewerking van de effektscores met als doel deze onderling (beter) vergelijkbaar te maken. Gekonstateerd moet worden dat in de literatuur de grote invloed van standaardisatie op de uiteindelijke rangschikking van de alternatieven onderbelicht is.

In de experimenten met de IODZH-gegevens is steeds één standaardisatiemethode toegepast op de totale effektenmatrix. Dit is in het algemeen niet aan te bevelen. Het is beter om per criterium te beoordelen welke standaardisatiemethode voor dat criterium gebruikt kan worden. Immers, gestandaardiseerde effektscores kunnen ook geïnterpreteerd worden als effectiviteiten. In dit verband kan effectiviteit opgevat worden als een maat die iets zegt over de mate waarin een alternatief voldoet aan de in de vorm van een criterium gevatte doelstelling. In deze zin opgevat is standaardisatie per definitie een



subjektieve bezigheid, aangezien een ieder en een verschillende mening op na houdt in hoeverre een bepaalde effektskore aflight van het gewenst optimum (= doelstelling). Het behoeft niet waar te zijn, dat de maximale (of minimale) skore van de te evalueren alternatieven ook inderdaat deze optimum-waarde heeft. In feite wordt dit wel verondersteld bij gebruik van respektievelijk de methoden MAX en MIN. Hetzelfde probleem treedt op bij gebruik van de standaardisatiemethode KLASSE. Het aangeven van de klassegrenzen is geen objektieve maar een subjektieve zaak. Uit bovenstaande volgt dat het raadzaam is om, indien interaktie met beslissers mogelijk is, de gestandaardiseerde effektskores te verifiëren met behulp van de methode Saaty.

Om de (per criterium op verschillende wijze) gestandaardiseerde effektskores onderling (tussen kriteria) te kunnen vergelijken moeten de rijen van de effektiviteitenmatrix vervolgens genormeerd worden.

In paragraaf 4.2 worden voorbeelden van methoden ter verkrijging van niet-kardinale (ordinale) gegevens (RANGORDE, KLASSE, SAATY, QUALI), enkele kardinale methoden (MAX, MIN, MAXMIN, TOTAAL), methoden met een keuzeparameter (NUTS, INDIF) en enkele meer geavanceerde methoden (ENTROPIE, WATERLAB) gegeven.

Standaardisatiemethoden die gebruik maken van intern gedefinieerde normeringsbases zijn gevoelig voor wijzigingen in het alternatievenbestand. Het toevoegen of weglaten van "extreme" alternatieven kan grote gevolgen hebben voor de rangschikking van de alternatieven. In het IODZH-voorbeeld blijkt dat MIN het minst gevoelig is voor het weglaten van (extreme) alternatieven, aangezien voor alle kriteria geldt "hoogst=slechtst". NUTS, MAX en MAXMIN blijken matig gevoelig te zijn, terwijl ENTROPIE en TOTAAL het meest gevoelig zijn voor het weglaten van extreme alternatieven. Gebruik van externe referentiepunten, zoals hierboven bedoeld, kan hier een oplossing bieden; echter, ook hieraan kleven bezwaren (paragraaf 4.3).

Standaardisatiemethode MIN is veelal ongeschikt, met name bij skores die dichtbij nul liggen en bij negatieve skores (overigens is het voorkomen van zowel negatieve als positieve skores op een criterium een probleem bij standaardisatie). Daarbij wordt veelal niet ingezien dat MIN een niet-lineaire

standaardisatiemethode is, terwijl het niet de bedoeling is om een dergelijke weging te introduceren.

Gezien de grote invloed van standaardisatiemethoden op de rangschikking van alternatieven is het aan te bevelen meerdere standaardisatiemethoden toe te passen en de verschillen in rangschikking uitvoering te analyseren.

### Gewichten

Gewichten worden geïntroduceerd om het relatieve belang van de gehanteerde criteria aan te geven. In tegenstelling tot andere evaluatiemethoden maken MKE-methoden expliciet gebruik van dergelijke gewichten om tot een rangschikking van alternatieven te komen. In de MKE-procedure moet derhalve veel aandacht besteed worden aan het verkrijgen van deze gewichten.

Methoden om gewichten te bepalen worden preferentiemethoden genoemd. Naast de in dit deel behandelde interactieve preferentiemethoden kunnen onderscheiden worden: preferentievaststelling op grond van waargenomen gedrag (revealed preference), het gebruik van meerdere gewichtsvectoren en het gebruik van enquêtes. Deze methoden hebben het nadeel dat de gewichtsvectoren die de analist opstelt niet geverifieerd worden door de beslisser. In de interactieve benadering is dit wel het geval. Het gebruik van meerdere gewichtsvectoren heeft ook voordelen en wordt wel toegepast bij gevoeligheidsanalyses in het algemeen en bij de methode Paelinck in het bijzonder. Ook de interactieve benadering heeft een nadeel: de procedure kan ondoorzichtig worden door het iteratief elimineren van alternatieven en het iteratief vaststellen van gewichten (paragraaf 5.2).

In dit deel zijn twee interactieve preferentiemethoden verder uitgewerkt, namelijk de methode Saaty en een in het kader van dit onderzoek ontwikkelde methode WATERLAB. Deze laatste methode werkt met door de analist vastgestelde trade-offs, waarop de beslisser zijn/haar standpunt of beleidsruimte aangeeft door middel van een keuze voor één of meerdere alternatieven op de trade-off curve. Op de in paragraaf 5.5 vermelde wijze kan vervolgens een gewichtsvector afgeleid worden.



Aangezien de methode WATERLAB in feite een gekombineerde preferentie- en standaardisatiemethode is, heeft deze methode het voordeel dat het probleem dat veroorzaakt wordt door introductie van impliciete en interne gewichten ten gevolge van standaardisatie van effektscores omzeild kan worden.

Standaardisatiemethoden kunnen immers tot verschillende rangschikkingen van alternatieven komen. Dit kan geïnterpreteerd worden als het toekennen van extra gewichten aan effektscores. Deze extra gewichten worden onderscheiden in impliciete en interne gewichten (zie paragrafen 5.3 en 5.4). Deze kunnen zulke grote waarden aannemen dat bij een vergelijking van verschillende standaardisatiemethoden de expliciete criteriumgewichten daardoor worden overschaduwd.

Het bovenstaande leidt tot de konklusie dat bij een gewichtenvaststelling rekening gehouden moet worden met de gebruikte standaardisatiemethoden en met de gebruikte MKE-methoden: MKE-methoden maken namelijk op een verschillende wijze gebruik van de opgegeven gewichten.

#### Overige onzekerheden

De invloed van kleine onzekerheden in het uitgangsmateriaal (de effectenmatrix) kan grote vormen aannemen; een analyse hiervan is van groot belang (zie paragraaf 6.5 en ook [35]).

De analyse van onzekerheden in de effectenmatrix is niet alleen van belang voor de evaluatie van alternatieven. Deze analyse kan ook toegepast worden in het kader van onderzoeksprogrammering. Immers, door bij aanvang van een studie een MKE uit te voeren aan de hand van een geschatte effectenmatrix kan onderzocht worden in welke mate de rangschikking van alternatieven gevoelig is voor onzekerheden met betrekking tot de a priori geschatte scores. Hieruit kan afgeleid worden dat onderzoek naar de meest "kritische" aspecten de voorkeur verdient boven onderzoek naar aspecten die minder invloed hebben op de uiteindelijke rangschikking.

De analyse van trade-offs kurven, zoals dat bij de methode WATERLAB geschiedt, kan tevens een controlefunctie hebben op de volledigheid van het alternatievenbestand (zie subparagraaf 5.5.5).

## 7.2 Technische eisen MKE-procedure

In het onderstaande worden enige eisen genoemd, die idealiter aan een MKE-procedure zouden moeten worden gesteld. Daarbij is een onderscheid gemaakt in procedure in het algemeen, de criteria, de gegevensbehandeling, de standaardisatie, de uiteindelijke rangschikking en de eisen die gesteld moeten worden ten aanzien van het uitvoeren van gevoeligheidsanalyses. De lijst is niet uitputtend en wellicht ook afhankelijk van persoonlijke voorkeuren.

### MKE-procedure - algemeen

#### 1 - overdraagbaarheid van de gebruikte algorithmen en aannamen

Belangrijk voor de acceptatie van een MKE-procedure is dat haar achterliggende basisgedachten en aannamen zonder veel moeite aan de beslissingsnemer uit te leggen zijn; zij moet bovendien aanspreken: de MKE-procedure moet een goede afbeelding zijn van een beslisproces, waar de beslissingsnemer achter kan staan.

#### 2 - expliciteerbaarheid van criteriumgewichten

Een MKE-procedure veronderstelt dat criteriumgewichten beschikbaar zijn. Dit betekent dat de beslisser(s) in staat moet(en) zijn om een rangorde in de belangrijkheid van de criteria te onderkennen en deze over te dragen aan de analist. Naast een rangorde zou de beslisser ook in staat moeten zijn een verschil in belangrijkheid van criteria te kwantificeren (ter bepaling van kardinale gewichten).

### Kriteria

#### 3 - onafhankelijkheid van criteria

De te hanteren criteria moeten onafhankelijk van elkaar zijn. Bij een volledige afhankelijkheid is er immers geen sprake meer van multikriteria evaluatie. Anderzijds komt het vaak voor dat onder criteria verschillende aspecten van de achterliggende doelstelling genoemd worden. Het is in dergelijke gevallen raadzaam een hiërarchie in criteria aan te brengen. Onafhankelijkheid van criteria kan gedetecteerd worden met behulp van correlatie-technieken.



#### 4 - transiviteit van preferenties

De preferenties moeten transitief zijn, dat wil zeggen dat indien criterium 1 belangrijker geacht wordt dan criterium 2 en indien criterium 2 belangrijker geacht wordt dan criterium 3, moet volgen dat criterium 1 belangrijker is dan criterium 3.

### Gegevens

#### 5 - toepassingsmogelijkheden van gemengde data

In veel gevallen is de effectenmatrix een mengsel van kwantitatieve en kwalitatieve data. Het is mogelijk om bijvoorbeeld alle kardinale gegevens in zo'n geval om te vormen naar ordinale of kwalitatieve gegevens en vervolgens met een daarvoor geschikte MKE-methode te verwerken, maar in verband met het verlies aan informatie (bijvoorbeeld de onderlinge afstanden van de alternatieven) is het beter om een MKE-methode toe te passen, die beide typen data op verantwoorde wijze kan verwerken, zoals de in paragraaf 3.8 beschreven Evamix-benadering.

#### 6 - methodologische zuiverheid van de verwerking van gegevenseigenschappen

In aansluiting op het onder 5 gestelde dient aan een MKE-procedure de eis gesteld te worden dat de specifieke eigenschappen van de gegevens op de bij die eigenschappen aansluitende wijze verwerkt worden. Dat betekent bijvoorbeeld dat kwalitatieve gegevens niet geïnterpreteerd worden als gegevens die op een ratioschaal zijn gemeten en ook als zodanig verwerkt worden. Bij het voorkomen van kwalitatieve gegevens moeten dus kwalitatieve MKE-methoden (zoals de kwalitatieve konkordantiemethode en de kwalitatieve schalingsmethoden) toegepast worden.

#### 7 - toelaatbaarheid van negatieve scores

Veel methoden laten slechts positieve scores toe en vereisen daardoor een voorbewerking van scores die nul of negatief zijn. Deze voorbewerking kan echter ernstige gevolgen voor de resultaten hebben. Daarom is het wenselijk haar in voorkomende gevallen te expliciteren als onderdeel van de methode en als zodanig te behandelen. Overigens is het voorkomen van negatieve (of anderszins speciale!) scores mede afhankelijk van het type criterium.

8 - uitwisselbaarheid van ordinale en kardinale gegevens

Als een methode zowel kardinale als ordinale informatie kan verwerken is het wenselijk dat bij een probleem met kardinale gegevens de resultaten van de methode op grond van een ordinale versie van de gegevens (zie subparagraaf 4.2.1) in grote lijnen overeenkomen met die op basis van de oorspronkelijke, kardinale, gegevens. Bij grote verschillen van de kardinale gegevens die door de ordinalisering worden teruggebracht tot relatief kleine verschillen op een ordinale schaal wordt niet aan deze eis voldaan. In feite betreft het hier een standaardisatieprobleem.

Standaardisatie

9 - standaardisatie per criterium

Standaardisatie dient plaats te vinden per criterium op een zodanige wijze dat de gestandaardiseerde score een effectiviteit voorstelt ten opzichte van de middels een criterium weergegeven doelstelling. Vervolgens dienen de rijen van de effectiviteitsmatrix genormeerd te worden zodat de effectiviteiten van verschillende criteria eenduidig onderling te vergelijken zijn.

10 - eenduidigheid van interpretatie

Het kan voorkomen dat het beste alternatief op een bepaald criterium de hoogste score krijgt (en slechtere alternatieven navenant lagere scores), maar ook het omgekeerde is mogelijk: hogere scores betekenen slechtere alternatieven. Voor de effectiviteitsmatrix moet echter eenduidigheid van richting gelden, dus voor alle criteria moet gelden "hoogst = best" of "hoogst = slechtst".

11 - geen onderlinge vervlechting van verschillende typen weging

De standaardisatie mag in principe geen impliciete weging tussen criteria introduceren. Bij gebruik van MKE-methoden die gevoelig zijn voor impliciete weging (zoals gewogen sommering) wordt het voor de beslisser nagenoeg ondoenlijk zich een beeld te vormen van de gewichtenproblematiek op basis waarvan hij/zij op verantwoorde wijze onderscheid in belangrijkheid van criteria kan uitdrukken in de vorm van criteriumgevechten.



## Rangschikking

### 12 - inzicht in de nauwkeurigheid van de eindrangschikking

Het resultaat van een MKE krijgt pas profiel tegen de achtergrond van de nauwkeurigheid waarmee het is vastgesteld; dat wil zeggen dat de gevoeligheid van het eindresultaat voor veranderingen in het uitgangsmateriaal waarvan de grootte-orde in relatie staat met zijn onzekerheid zichtbaar moet kunnen worden gemaakt. Dit is zowel van toepassing op de effectenmatrix (zie de voorbeelden in paragraaf 6.5) als op de gebruikelijke gewichtenvectoren (zie paragraaf 6.4).

### 13 - onderscheidend vermogen van de methode

De belangrijkste taak van een MKE is het aantonen van onderscheid tussen de meespelende alternatieven in hun kwaliteit de gestelde doelstellingen (kriteria) dichtbij te brengen. Als eis moet derhalve worden gesteld dat een methode zulk onderscheid indien aanwezig ook zichtbaar maakt. Omgekeerd is het ongewenst als onderscheid wordt aangebracht tussen alternatieven die even goed zijn. Theoretisch gesproken moet de MKE een zwakke ordening van de alternatieven opleveren die overeenkomt met de volgorde, die op grond van het beeld dat de beslissingsnemer zich van de problematiek gevormd heeft, "juist" moet worden genoemd. In principe kan dan gesteld worden dat de beslisser in dat geval geen behoefte heeft aan MKE, maar indien de problematiek zo complex en ondoorzichtig is dat de beslisser zich geen beeld kan vormen, kan juist MKE helpen de problematiek uiteen te rafelen teneinde dit beeld wel te krijgen. In dat geval wordt niet voldaan aan deze, overigens zeer theoretische, eis.

### 14 - verwerkbaarheid van grote problemen

Hoewel de betrouwbaarheid bij de verwerking van grote effectenmatrices dubieus is, is het van belang dat grote problemen toch door de methode kunnen worden verwerkt. Een methode hiervoor is het stratificeren van het gegevensmateriaal, waardoor het in twee of zelfs drie niveaus hiërarchisch kan worden aangepakt.

15 - gevoeligheid voor eliminatie van alternatieven

De rangschikking van alternatieven mag niet veranderen door weglaten (of toevoegen) van alternatieven. In principe is dit een standaardisatieprobleem in het geval dat gewerkt wordt met intern gedefinieerde extremen (zie paragraaf 4.3).

Gevoeligheidsanalyse

16 - Gevoeligheidsonderzoek behoort een essentieel onderdeel van de MKE-procedure uit te maken. Het verdient dan ook aanbeveling de MKE zodanig in te passen dat op een eenvoudige wijze systematisch gevoeligheidsonderzoek kan worden uitgevoerd. De resultaten hiervan moeten op een eenvoudige manier presenteerbaar zijn. Het gevoeligheidsonderzoek dient de volgende onderdelen te bevatten:

- . onderzoek naar afhankelijkheid van criteria;
- . onderzoek naar onzekerheid in effektscores;
- . onderzoek naar standaardisatie-onzekerheid;
- . onderzoek naar onzekerheid van gewichten;
- . onderzoek naar methode-onzekerheid.

Uit het bovenstaande volgt dat bij een MKE-procedure meerdere methoden en gewichtenvectoren gebruikt dienen te worden.



## APPENDIX - HET PROGRAMMA MKEM

### Algemeen

Ten behoeve van het onderzoek naar de waarde van MKE-methoden is het computer-programma MKEM ontwikkeld. Het is ontstaan in de loop van 1981 en geschreven in Fortran IV. In totaal zijn zestien MKE-methoden geprogrammeerd. Het programma kan zonder wijziging in de code problemen aan bestaande uit maximaal 15 criteria en 32 alternatieven. Programmeur van MKEM is E. Mulder van het Waterloopkundig Laboratorium. Zie voor een verklaring van in deze bijlage of in het hoofdrapport niet-gedefinieerde begrippen de ook beschikbare "Handleiding MKEM".

### Verschillen met rapport

Een aantal MKE-methoden, die om uiteenlopende redenen niet in het rapport voorkomen, is wel opgenomen in MKEM: zo kunnen kwalitatieve resultaten worden verkregen in plaats van de gebruikelijke ordinale resultaten met de kwalitatieve konkordantie- en diskonkordantiemethode (KWALKON).

Van de standaardisatiemethode INDIF (zie subparagraaf 4.2.3) zijn twee varianten opgenomen: INDIFEFF1 standaardiseert tussen 0 en 1, INDIFEFF2 standaardiseert tussen het quotiënt van minimum- en maximumskore 1.

Methode NUTS2 doet hetzelfde als de methode NUTS, met dien verstande dat in plaats van standaardisatie tussen 0 en 1 wordt gestandaardiseerd tussen het quotiënt van minimum- en maximum en 1.

### Mogelijkheden en beperkingen

#### Mogelijkheden

1. Er kan met een deelverzameling van de ingevoerde effectenmatrix worden gerekend. Hiermee kan de gevoeligheid van resultaten voor criteria en/of alternatieve worden onderzocht.

2. Zowel criteria waarvoor geldt dat een hoge score een slechte score is als criteria waarbij het omgekeerde het geval is, kunnen met MKEM worden bewerkt.
3. Het is mogelijk de gevoeligheid van de resultaten van een MKE-methode te onderzoeken voor het elimineren van de slechtst scorende alternatieven (zie ook 1). Deze mogelijkheid is echter (uiteraard) beperkt tot het gebruik van de kardinale methoden.
4. Ook bestaat de mogelijkheid om grafische uitvoer te verkrijgen van gestandaardiseerde scores en/of van het impliciete gewicht dat aan die scores wordt toegerekend bij gebruik van de desbetreffende standaardisatiemethode.
5. De methoden INDIFEFF1, INDIFEFF2, NUTS en NUTS2 maken gebruik van de parameter  $\lambda$ . Deze faktor is o.a. uitgevoerd om het impliciete gewicht van een score te kunnen beïnvloeden. De gebruiker van MKEM kan met behulp van  $\lambda$  de kromming van de standaardisatiefunctie vastleggen.
6. Om de standaardisatie van een criterium verder nog te beïnvloeden is de mogelijkheid ingebouwd externe extremen (zie paragraaf 4.3) te definiëren. Deze extremen moeten liggen buiten de door middel van de alternatieven-effektenmatrix ingevoerde extremen en vlakken de standaardisatie af.
7. Per run kunnen alle of slechts enkele methoden tegelijk worden gebruikt, zij het met één uitzondering. De methoden INDIFEFF1, INDIFEFF2, NUTS en NUTS2 kunnen niet tegelijkertijd worden gebruikt omdat zij konfliktierende eisen stellen ten aanzien van de grootte van  $\lambda$ .

#### Beperking

Het programma kan in principe slechts scores groter dan nul verwerken. Sommige methodes kunnen wel negatieve scores aan, maar aanbevolen wordt niet van deze mogelijkheid gebruik te maken.



## BIBLIOGRAFIE

1. VOOGD, J.H., Het gebruik van multicriteria-evaluatie in de ruimtelijke planning  
Delft, Planologisch Studiecentrum TNO, 1980
2. VOOGD, J.H., MIDDENDORP, C., UDINK, B., VAN SETTEN, A., Multicriteria-methoden voor ruimtelijk evaluatieonderzoek  
Delft, Planologisch Studiecentrum TNO, 1980
3. ANONIEM, Optimalisatietechnieken in kwantitatief en kwalitatief waterbeheer  
Toegepast Onderzoek Waterstaat, R999-4,  
Delft, Waterloopkundig Laboratorium, februari 1978
4. ANONIEM, Optimalisatie op basis van lineair programmeren (LP) en dynamisch programmeren (DP); mogelijkheden en beperkingen  
Toegepast Onderzoek Waterstaat, R 999-5,  
Delft, Waterloopkundig Laboratorium, mei 1982
5. ANONIEM, Optimalisatietechnieken in kwantitatief waterbeheer  
Toegepast Onderzoek Waterstaat, R 999-6,  
Delft, Waterloopkundig Laboratorium, november 1981
6. ANONIEM, Illustrative examples of optimization techniques for quantitative and qualitative water management  
Toegepast Onderzoek Waterstaat, R 999-8,  
Delft, Waterloopkundig Laboratorium, augustus 1983
7. ANONIEM, Interimrapport Stuurgroep Integraal Onderzoek Drinkwatervoorziening Zuid-Holland  
Leidschendam, Rijksinstituut voor Drinkwatervoorziening, augustus 1981
8. JONG, L.C. de, VOOGD, J.H., Planologisch Memorandum 1980-14, Structuurschema burgerluchtbaartterreinen; de evaluatie van de evaluatie  
Delft, Technische Hogeschool Delft, Vakgroep Civieke Planologie, december 1980

BIBLIOGRAFIE (vervolg)

9. COBA, Evaluatiemethoden, een introductie  
's-Gravenhage, Staatsuitgeverij, januari 1982
10. RAND, Protecting and estuary from floods - a policy analysis of the  
Oosterschelde  
Santa Monica (Cal.), Rand Corporation, december 1977
11. HOEN, H. 't, VOOGD, J.H., A note on the use of multicriteria analysis  
for a financial assessment of plans, Planologisch Memorandum 1981-6,  
Delft, Technische Hogeschool Delft, Vakgroep Civiele Planologie, oktober  
1981
12. VOOGD, J.H., Planologisch Memorandum 1981-7, Multicriteria evaluation  
with mixed qualitative-quantitative data  
Delft, Technische Hogeschool Delft, Vakgroep Civiele Planologie, augustus  
1981
13. RIETVELD, P., Multiple objective decision methods and regional planning  
Amsterdam, North-Holland Publishing Company, 1980
14. ANONIEM, Ontwikkeling en toepassing van een methodiek voor de besluit-  
vorming bij grote projecten  
Onderzoeksgroep procesbeheersing en milieubeheer,  
Enschede, Technische Hogeschool Twente, mei 1978
15. NIJKAMP, P., VOOGD, J.H., New multicriteria methods for physical plan-  
ning by means of multidimensional scaling techniques  
Research memorandum 1980-1, Amsterdam, Vrije Universiteit, januari 1980
16. NIJKAMP, P., Soft Econometric Models. An analysis of regional income  
determinants  
Research memorandum 1980-5, Amsterdam, Vrije Universiteit, april 1980



BIBLIOGRAFIE (vervolg)

17. NIJKAMP, P., DELFT, A. van, Multicriteria analysis and regional decision making  
Leiden, M. Nijhoff, 1977
18. LICHFIELD, N., KETTLE, P., WHITBREAD, M., Evaluation in the planning process  
Oxford, Pergamon, 1975
19. MOL, F.J.M. de, Multi-Criteria Analyse. Wat is het en is het wat?  
Beleidsanalyse 79-4, pp. 2-28, 1979
20. MOL, F.J.M. de, POLL, E.H. van de, SCHUT, M., Projektevaluatie economische evaluatie van de civiel-technische projecten  
Delft, Technische Hogeschool, Afdeling Civiele Techniek, 1980
21. NES, Th.J. van de, LOHUIZEN, E.P., Onderzoek naar de toepasbaarheid van multicriteria afwegingsmodellen ten behoeve van de waterhuishouding in Gelderland  
Arnhem, 1978
22. NIJKAMP, P., Theory and application of environmental economics  
Amsterdam, North-Holland Publishing Company, 1977
23. NIJKAMP, P., Naar een prijzenswaardig milieu?!  
Assen, Van Gorcum, 1979
24. VOOGD, J.H., Methoden en technieken betreffende evaluatie  
Delft, Planologisch Studiecentrum TNO, 1976
25. NIJKAMP, P. en VOOGD, J.H., The use of multidimensional scaling in evaluation procedures, Research paper 13,  
Delft, Planologisch Studiecentrum TNO, 1978

BIBLIOGRAFIE (vervolg)

26. BENAYOUN, R., ROY, B., SUSSMAN, B., Manual de référence du programme Electre  
Note de Synthèse et Formation No. 25 de la Direction Scientifique de la SEMA, 1966
27. HOLMES, J.C., An ordinal method of evaluation  
Urban Studies 9, nr. 1, pp. 179-191, 1971
28. JACQUET-LAGREZE, E., L'Aggrégation des opinions Individuelles Informatiques en Sciences Humaines, 4, 1969
29. PAELINCK, J.H.P., Beslissingen op grond van verscheidene criteria bij zuiver kwalitatieve informatie  
Economisch Statistische Berichten 3043, pp. 247-249, 1976
30. SCHIMPELER, C.C., GRECCO, W.L., The community-system evaluation - an approach based on community structure and values  
Highway Research Record 238, 1968
31. SCHLAGER, K., The rank-based expected value method of plan-evaluation  
Highway Research Record 238, pp. 153-158, 1968
32. STANLEY, J.K.A., A cardinal utility approach for project evaluation  
Socio-Economic Planning Science 8, pp. 329-338, 1974
33. HINLOOPEN, E., NIJKAMP, P., RIETVELD, P., Qualitative discrete multiple criteria choice models in regional planning  
Regional Science and Urban Economics 13, pp. 77-102, 1983
34. DORFMAN, R., Conceptual model of a regional water quality authority  
in: DORFMAN, R., JACOBY, H.D., THOMAS, H.A. jr., Models for managing regional water quality  
Cambridge (Mass.), Harvard, 1972



BIBLIOGRAFIE (vervolg)

35. KEENY, R.L., RAIFFA, H., Decisions with Multiple Objectives: Preferences and Value Tradeoffs  
New York, Wiley, 1976
36. FRIEND, J.K., JESSOP, W.N., Local government and strategic choice - An Operational Research Approach to the Processes of Public Planning  
Oxford, Pergamon, 1969 (second edition 1977)
37. KLAASSEN, L.H., VERSTER, A.C.P., Kosten- en batenanalyse in regionaal perspectief  
Groningen, Tjeenk Willink, 1974
38. KOUDSTAAL, R., PENNEKAMP, H.A., Planning for Water Resources Management - framework for analysis (S455-D83/001)  
Delft, Waterloopkundig Laboratorium, 1983
39. DRIEL, G.J. van, RAVENZWAALIJ, C. van, VEENEKLAAS, F.R., Grenzen en mogelijkheden van het economisch stelsel in Nederland  
(Wetenschappelijke Raad voor het Regeringsbeleid, Voorstudies en achtergronden, nr. V4, 1983)  
's-Gravenhage, Staatsuitgeverij, 1983
40. Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer, Meer dan de som der delen; eerste nota over de planning van het milieubeleid,  
's-Gravenhage, Staatsuitgeverij, 1984
41. BRAFF, A.J., Micro-economic Analysis  
New York, Wiley, 1969
42. MISHAN, E.J., Cost-Benefit Analysis  
London, Allen & Unwin, 1971

BIBLIOGRAFIE (vervolg)

43. WALKER, W.E., VEEN, M.A., Policy Analysis of Water Management for the Netherlands, Vol. II, Screening of Technical and Managerial Tactics, N-1800/2-NETH, prepared for The Netherlands Rijkswaterstaat, The Rand Corporation, Santa Monica, California, 1981
44. ANONIEM, Evaluatiemethoden, een introductie  
Rapport van de Afdeling Beleidsanalyse van het Ministerie van Financiën, 2e geheel herziene druk,  
's-Gravenhage, Staatsuitgeverij, oktober 1984
45. JANSSEN, R., Evaluatiemethoden ten behoeve van milieubeleid en -beheer  
Amsterdam, Vrije Universiteit, Economisch en Sociaal Instituut, 1984
46. ANONIEM, Eindrapport Stuurgroep Integraal Onderzoek Drinkwatervoorziening Zuid-Holland  
Uitgebracht door Rijksinstituut voor Drinkwatervoorziening, Leidschendam en Provinciale Waterstaat Zuid-Holland, 's-Gravenhage, augustus 1983



p.o. box 177

2600 mh delft

the netherlands