

MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

AMSTERDAM

STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig

Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S 140

Voorlopig rapport over een vlamstralingsonderzoek.

(Eerste rapport over het onderdeel P.T.5)

door

R.Doornbos

1954

1. Inleiding.

In dit rapport houden we ons bezig met de analyse van de resultaten van de vijfde proevenserie in de reeks experimenten betreffende de warmtestraling van vlammen die te IJmuiden worden uitgevoerd onder auspiciën van het Flame Radiation Research Joint Committee. Deze vijfde proevenserie had ten doel de invloed van een aantal factoren op de intensiteit van de straling na te gaan. In de eerste plaats werden twee soorten brandstof gebruikt, nl. 33% olie en 67% gas naast 67% olie en 33% gas, voortaan aan te duiden met 1, resp. 2. Voorts was de vlam al dan niet symmetrisch, dat wil zeggen, dat het gas en de verstoven olie op dezelfde plaats in de oven werden geblazen, dan wel onder elkaar. Dit geven we aan met c, resp. d.

Met elk van de vier mogelijke combinaties 1c, 1d, 2c en 2d werden vier proeven genomen. Van deze 16 vlammen werd op 6 verschillende punten in de lengterichting van de oven door spleten in de zijwand genummerd 2, 3, 4, 5, 6, 7, de straling gemeten van de vlam alleen, van de vlam plus de daarachterliggende hete zijwand en van deze zijwand alleen (R_1 , R_2 en R_3) in $\text{cal cm}^{-2} \text{sec}^{-1}$. Bovendien werd de straling gemeten op drie punten op de bodem van de oven (floor radiation). De waarde der bodemstraling geven wij verder aan met B. Behalve van de drie soorten straling R_1 , R_2 en R_3 werd gevraagd de invloed van de genoemde factoren op de emissiecoëfficiënt $\epsilon \stackrel{\text{def}}{=} 1 - \frac{R_2 - R_1}{R_3}$ na te gaan.

In eerste instantie werd de analyse uitgevoerd met de methode van de variantieanalyse, afkomstig van R.A.FISHER. Literatuur over dit onderwerp kan men naast het boek van FISHER [1] b.v. vinden in MANN [2] en WILKS [3]. De resultaten van dit onderzoek worden in dit voorlopige rapport vermeld.

Aan de variantieanalyse liggen evenwel onderstellingen ten grondslag, waarvan het niet zeker is, dat ze in ons geval alle zijn vervuld. Daarom ligt het in de bedoeling het onderzoek met andere methoden te repliceren, methoden, waarbij minder onderstellingen omtrent de onbekende waarschijnlijkheidsverdelingen van de waargenomen grootheden behoeven te worden gemaakt. In een definitief rapport zullen dan de resultaten van beide onderzoekingen worden vergeleken en samengevat.

2. Vooronderzoek.

De onderstellingen van de variantieanalyse zijn de volgende:

- a) De waarnemingen zijn onderling onafhankelijk,
- b) De waarnemingsuitkomsten zijn normaal verdeeld,
- c) Al deze normale verdelingen hebben dezelfde spreiding.

Willen we de variantieanalyse dus toe kunnen passen op b.v. alle waarnemingen van R_1 gezamenlijk, waardoor we een drievoudige classificatie met 4 waarnemingen per cel zouden verkrijgen, dan moeten al deze waarnemingen dus onafhankelijk zijn. Daar het zeer goed denkbaar is, dat de waarnemingen van verschillende spleten maar behorende bij één vlam afhankelijk van elkaar zijn, hebben we dit nader onderzocht met behulp van de methode der m rangschikkingen (zie memorandum S 47 (M 14)) ¹⁾. Deze methode werd toegepast op de vier combinaties 1c, 1d, 2c en 2d afzonderlijk, voor R_1 , R_2 , R_3 , e en B.

Voor R_1 , R_2 , R_3 en e vinden we dus 6 rangschikkingen, voor B 3 rangschikkingen van 4. De gevonden overschrijdingskansen k zijn samengevat in tabel 1.

Tabel 1

Overschrijdingskansen bij afhankelijkheidsonderzoek voor de waarnemingen der verschillende spleten (methode der m rangschikkingen)

	R_1	R_2	R_3	e	B
1c	0,0073	0,001	0,00001	0,032	0,0017
1d	0,022	0,00005	0,0000001	0,375	0,017
2c	0,056	0,23	0,0085	0,51	$0,03 < k < 0,2$
2d	0,035	0,22	0,00001	0,609	0,017

De kleine overschrijdingskansen geven aan dat onder de getoetste hypothese, die inhoudt dat voor elk groepje van vier waarnemingen elke permutatie van de rangnummers even waarschijnlijk is, de waarnemingsuitkomsten bijzonder zeldzame gebeurtenissen zijn. We verwerpen daarom deze hypothese en concluderen dat, uitgezonderd misschien voor het geval van de e waarden, de waarnemingen voor de verschillende spleten van dezelfde vlam afhankelijk van elkaar zijn, in die ~~zijn~~ ^{dat} vlam die bij één spleet een hoge waarneming oplevert, ~~dit~~ dit bij de andere spleten ook ^{kan} doen. We mogen dus niet de variatie over de verschillende spleten of bij de bodemstraling over de drie waarnemingspunten als een derde classificatie invoeren.

De tweede onderstelling, de normaliteit betreffende, is aan de hand van de beschikbare waarnemingen vrijwel niet op haar juistheid te toetsen. Het aantal waarnemingen per vakje, dat vier bedraagt, is hiervoor te klein. Wel kunnen we opmerken, dat

1) De memoranda waarnaar in dit rapport wordt verwezen zijn aan het einde van het rapport bijgevoegd.

als R_1 , R_2 en R_3 exact normaal verdeeld zouden zijn, $e = 1 - \frac{R_2 - R_1}{R_3}$ dit zeker niet is, $R_3 - R_2 + R_1$ daarentegen wel. We hebben daarom voor controle een gedeelte van de berekeningen met e herhaald met $R_3 \cdot e = R_3 - R_2 + R_1$, waarbij de conclusies elkaar bleken te dekken.

De gelijkheid van de spreidingen, de derde onderstelling van de variantieanalyse is getoetst met de toets van BARTLETT (zie memorandum S 53 (M 25)). Deze toets is toegepast op alle 4 groepen van 4 waarnemingen per spleet en analoog voor de bodemstraling. We vonden aldus $4 \times 6 + 3 = 27$ overschrijdingskansen, waarvan 3 iets kleiner dan 0,05 waren. Onder de getoetste hypothese, dus als de spreidingen gelijk zijn, is de kans dat we onder 27 drie of meer van dergelijke overschrijdingskansen vinden, gelijk aan 0,15; dit is dus niets bijzonders. We kunnen deze hypothese dus niet verwerpen, zodat er, voor zover wij na kunnen gaan, geen ernstige bezwaren bestaan tegen het toepassen van variantie-analyse op de bij één spleet verkregen waarnemingen.

3. Variantieanalyse.

We toetsen nu per spleet de volgende hypothesen:

- α . De soort brandstof heeft geen invloed op de straling;
- β . De symmetrie of asymmetrie van de vlam beïnvloedt de straling niet;
- γ . Er is geen interactie tussen de twee hiervoor genoemde invloeden. Dit laatste betekent dat het verschil tussen de twee soorten brandstof, zo het bestaat, bij de symmetrische vlam gelijk is aan dat verschil bij de asymmetrische vlam.

De toetsingsgrootheden hebben onder de te toetsen hypothesen een F-verdeling, met in ons geval, 1 en 12 vrijheidsgraden. In tabel 2 zijn de resultaten van deze analyse voor R_1 , R_2 , R_3 en e samengevat. De waarden van F die onder de nulhypothese met kansen 0,05, 0,01 en 0,001 worden overschreden zijn 4,75, 9,33 en 18,64. Is dit het geval, dan hebben we dat in de tabel aangegeven met I, II resp. III. De rijen voor brandstof, symmetrie en interactie zijn aangeduid met α , β , resp. γ .

Voor een enkel geval, nl. de straling R_1 bij spleet 2 hebben we de resultaten van de variantieanalyse weergegeven op de gebruikelijke, meer uitgebreide wijze, in tabel 2a.

Tabel 2
Variantieanalyse met 2 classificaties voor ieder der spleten apart

plot no		2			3			4			5			6		7	
	α	111,34	III	1 < 2	192,66	III	1 < 2	39,89	III	1 < 2	13,31	II	1 < 2	0,39	1 > 2	3,35	1 > 2
R ₁	β	17,12	II	c < d	76,33	III	c > d	108,28	III	c > d	75,44	III	c > d	5,77	I c > d	0,50	c > d
	γ	4,72			2,19			1,29			1,77			0,64		3,88	
	α	91,49	III	1 < 2	72,09	III	1 < 2	28,02	III	1 < 2	6,45	I	1 < 2	4,12	1 > 2	0,44	1 > 2
R ₂	β	21,31	III	c < d	24,24	III	c > d	57,81	III	c > d	16,14	II	c > d	1,87	c > d	0,56	c > d
	γ	2,47			0,12			0,39			0,00			0,04		0,87	
	α	32,10	III	1 < 2	34,82	III	1 < 2	17,27	II	1 < 2	3,10		1 < 2	0,12	1 > 2	2,69	1 > 2
R ₃	β	5,37	I	c < d	4,89	I	c < d	0,75		c < d	0,44		c > d	1,68	c > d	2,91	c > d
	γ	0,27			0,83			0,61			1,29			1,34		1,90	
	α	25,21	III	1 < 2	55,35	III	1 < 2	11,24	II	1 < 2	1,94		1 < 2	0,96	1 < 2	2,54	1 > 2
e	β	0,58		c < d	2,45		c > d	12,35	II	c > d	18,86	III	c > d	2,75	c > d	0,20	c < d
	γ	2,11			0,80			0,84			4,00			1,75		1,46	

$1 < 2$ betekent dat de gemiddelde straling bij gebruik van brandstof 1, gemiddeld dus over c en d, kleiner is dan de gemiddelde straling met brandstof 2. We zien dat er geen aanwijzing bestaat dat ergens interactie aanwezig is tussen de beide factoren.

Tabel 2a
Variantieanalyse voor R_1 , spleet 2

bron van variatie	vrijh. graden	som van kwadraten	gem. som v. kwadraten	F	overschr.kans
brandstof	1	6,5025	6,5025	111,34	$k < 0,001$
symmetrie	1	1,0000	1,0000	17,12	$0,001 < k < 0,01$
interactie	1	0,2756	0,2756	4,72	$k > 0,05$
toeval	12	0,7013	0,0584		
totaal	15	8,4794			

In de figuren 1 t/m 8 is de straling uitgezet voor de verschillende spleten. De overschrijdingskansen kleiner dan 0,05, 0,01 en 0,001 zijn weer met de cijfers I, II en III aangegeven.

Tevens zijn in deze figuren de betrouwbaarheidsintervallen (zie memorandum S 47 (M 18)) met onbetrouwbaarheid 0,05 getekend voor de verschillen van de gemiddelden voor brandstof 2 en brandstof 1 resp. de symmetrische en de asymmetrische vlam.

We moeten hierbij wel in het oog houden dat de uitkomsten van de toetsingen uit tabel 2 en de betrouwbaarheidsintervallen in de figuren voor de verschillende spleten, afhankelijk van elkaar zijn ten gevolge van de in paragraaf 2 gevonden vlaminvloed. De betrouwbaarheidsintervallen zijn dus niet simultaan geldig met dezelfde onbetrouwbaarheid. In het definitieve rapport zullen we trachten hiermee rekening te houden.

Dezelfde analyse is uitgevoerd voor de gemiddelde waarde van de straling over de gehele vlam evenals voor de eerste helft en de tweede helft van de vlam. Dit is zowel voor e als voor $R_1 - R_2 + R_3$ gedaan. In tabel 3 vinden we de waarden van de toetsingsgrootheden.

Tabel 3

Variantieanalyse met 2 classificaties voor gemiddelden over alle of de helft der spleten

		1e helft			2e helft			gehele vlam		
R ₁	α	122,39	III	1 < 2	0,72		1 < 2	44,98	III	1 < 2
	β	62,55	III	c > d	21,94	III	c > d	47,32	III	c > d
	δ	0,40			2,86			1,50		
R ₂	α	68,27	III	1 < 2	0,16		1 < 2	30,11	III	1 < 2
	β	20,37	III	c > d	9,32	I	c > d	20,61	III	c > d
	δ	0,24			0,37			0,42		
R ₃	α	27,57	III	1 < 2	0,03		1 > 2	4,11		1 < 2
	β	3,01		c < d	1,62		c > d	0,00		c > d
	δ	0,57			1,55			1,12		
e	α	57,41	III	1 < 2	0,09		1 < 2	19,94	III	1 < 2
	β	8,55	I	c > d	6,82	I	c > d	10,82	II	c > d
	δ	0,78			4,53			3,37		
R ₁ R ₂ R ₃	α	65,26	III	1 < 2	0,00		1 < 2	13,33	II	1 < 2
	β	4,94	I	c > d	6,08	I	c > d	7,27	I	c > d
	δ	0,86			4,72			3,27		

Bij de bodemstraling doet zich een complicatie voor, daar hier bij de groep 2.c in het eerste en het tweede waarnemingspunt een waarneming ontbreekt. Uit de theorie van de variantieanalyse volgt dat we de juiste toetsingsgrootte krijgen, door voor de ontbrekende waarnemingen het gemiddelde van de andere 3 in te vullen. We rekenen daarna de toetsingsgrootte op de normale manier uit, vermenigvuldigen de uitkomst met $\frac{12}{13}$ en hebben dan een grootte die onder de nulhypothese is verdeeld volgens een F-verdeling met 1 en 11 vrijheidsgraden. De theoretische achtergrond van de hier gevolgde methode kunnen we b.v. vinden in MANN, p. 138 en het gevolg uit stelling 4.3 op p. 37. De waarden van deze F die bij overschrijdingskansen van 0,05, 0,01 en 0,001 horen, zijn voor resp. 4,84, 9,65 en 19,69. We vinden nu de volgende uitkomsten (zie tabel 4).

Tabel 4

Variantieanalyse met 2 classificaties voor ieder der 3 bodempunten afzonderlijk

		1			2			3		
	α	37,97	III	$1 < 2$	1,64		$1 < 2$	5,74	I	$1 < 2$
B .	β	0,67		$c < d$	0,23		$c > d$	0,55		$c > d$
	δ	0,05			1,20			0,83		

In de figuren 9 en 10 vinden we B uitgezet, waarbij in fig. 9 weer over c en d en in fig. 10 over de brandstoffen 1 en 2 is gemiddeld. Bovendien zijn weer analoog aan de figuren 1-8 betrouwbaarheidsintervallen getekend voor de verschillen van de gemiddelde waarden.

Om ten slotte een indruk te krijgen van de gelijkmatigheid van de vlam bij de verschillende groepen is voor iedere vlam, voor R_1 , R_2 , R_3 en e de volgende grootte bepaald. Stel dat de waarnemingen voor de 6 spleten zijn $x_1, x_2, \dots, x_6$. Het gemiddelde noemen we $\bar{x}$. We berekenen nu

$$V^2 = \sum_{i=1}^6 (x_i - \bar{x})^2 = \sum_{i=1}^6 x_i^2 - \frac{(\sum_{i=1}^6 x_i)^2}{6}$$

V^2 is een maat voor de gelijkmatigheid van de vlam in die zin, dat V^2 kleiner is naarmate de vlam gelijkmatiger is. In tabel 5 staan de waarden van V^2 voor R_1 , R_2 , R_3 en e voor de 4 groepen.

Tabel 5

Maat voor gelijkmatigheid (V^2)

	R_1	R_2	R_3	e
1c	7,30	8,84	12,46	0,009
	8,86	10,10	15,11	0,020
	8,08	12,61	17,50	0,035
	8,82	11,34	17,33	0,016
1d	0,53	4,05	13,10	0,010
	0,27	4,99	11,57	0,013
	0,19	5,00	11,34	0,007
	0,27	5,36	15,29	0,024
2c	23,81	13,42	10,75	0,029
	18,91	7,02	11,29	0,074
	17,39	8,10	10,25	0,057
	23,91	14,81	11,79	0,032
2d	10,81	3,66	7,89	0,021
	9,65	4,06	6,30	0,095
	5,49	1,61	7,32	0,046
	6,14	2,57	6,40	0,048

Daar we hier niet aan mogen nemen, dat de voorwaarden van de variantieanalyse vervuld zijn (de gelijkheid van de spreidingen is hier b.v. zeker niet vervuld), passen we hier een verdelingsvrije methode toe om de invloed van brandstof en symmetrie op de gelijkmatigheid van de vlam na te gaan. We passen daarom b.v. voor R_1 de toets van WILCOXON (zie memorandum S 47 (M 7)) toe op de groepen 1c en 2c en ook op de groepen 1d en 2d. Daarop combineren we deze twee toetsen door optelling van de twee toetsingsgrootheden. Op dezelfde manier vergelijken we c tegen d, door de tweetallen 1c-1d en 2c-2d met elkaar te vergelijken en de resultaten te combineren. Hetzelfde doen we voor R_2 , R_3 en e. De overschrijdingskansen zijn samengevat in tabel 6.

Tabel 6

Vergelijking der vlam-gelijkmatigheid met behulp van de toets van Wilcoxon.

	R_1		R_2		R_3		e	
1-2	0,001	$1 < 2$	0,15	$1 > 2$	0,001	$1 > 2$	0,008	$1 < 2$
c-d	0,001	$c > d$	0,001	$c > d$	0,004	$c > d$	0,54	$c > d$

$c > d$ betekent dat $\sqrt{V^2}$ in het eerste geval groter is dan in het tweede, de vlam is dus het gelijkmatigst in het asymmetrische geval.

4. Conclusies.

Gemiddeld over de gehele vlam geeft brandstof 2 grotere straling dan brandstof 1; dit geldt zowel voor R_1 en R_2 als voor R_3 , terwijl ook de emissiecoëfficiënt bij brandstof 2 het grootst is. Voor de eerste helft van de vlam geldt precies hetzelfde, maar voor de tweede helft hebben we geen verschil kunnen aantonen, hoewel voor spleet no. 5 afzonderlijk, brandstof 2 duidelijk hogere waarden geeft dan brandstof 1 wat betreft R_1 en in iets mindere mate voor R_2 . De eerste drie spleten geven hetzelfde beeld als het gemiddelde over deze spleten. Samenvattend kunnen we dus zeggen dat brandstof 2 meer straling levert dan brandstof 1, maar dat dit verschil zich voornamelijk voordoet in de eerste helft van de vlam.

Verder is de straling R_1 gelijkmatiger over de gehele vlam bij brandstof 1 dan bij brandstof 2. Bij R_3 is het andersom en voor de emissiecoëfficiënt geeft brandstof 1 weer de regelmatigste uitkomsten. De bodemstraling is ook weer voor 2 groter dan voor 1, speciaal in het eerste punt, en in geringere mate in het

derde punt.

Het verschil tussen de symmetrische en de asymmetrische vlam, in het volgende weer met c en d aangegeven, is minder overzichtelijk. Over de gehele vlam gemiddeld geeft c hogere uitkomsten dan d, behalve bij R_3 , waar geen verschil is aangetoond. Bij de eerste en de tweede helft bestaat hetzelfde verschil, maar voor de eerste helft zijn de verschillen het grootst. Bezien we nu de spleten afzonderlijk, dan blijkt, dat voor R_1 , R_2 en R_3 de eerste spleet (no. 2) de uitkomsten met c juist lager liggen dan bij d. Bij R_1 en R_2 , waar de verschillen veel groter zijn dan bij e, is bij de spleten 3 en 4 de straling met c alweer veel groter, vandaar dat de eerste helft van de vlam de grootste waarden geeft met c. Bij R_3 bestaat er een aanwijzing dat de spleten 2 en 3 de grootste waarden met d geven. De emissiecoëfficiënt is met c hoger dan met d bij de spleten 4 en 5, de andere spleten geven geen significante uitkomsten. Voor de drie soorten straling R_1 , R_2 en R_3 geeft d gelijkmatiger uitkomsten over de lengte van de vlam dan c. De bodemstraling ten slotte geeft geen aanwijzing voor een verschil tussen c en d.

Literatuur.

- [1] R.A.FISHER, The design of experiments.
- [2] H.B.MANN, Analysis and design of experiments, New York, Dover Publications, inc. (1949)
- [3] S.S.WILKS, Mathematical Statistics, Chapter IX, Princeton University Press, Princeton (1946).

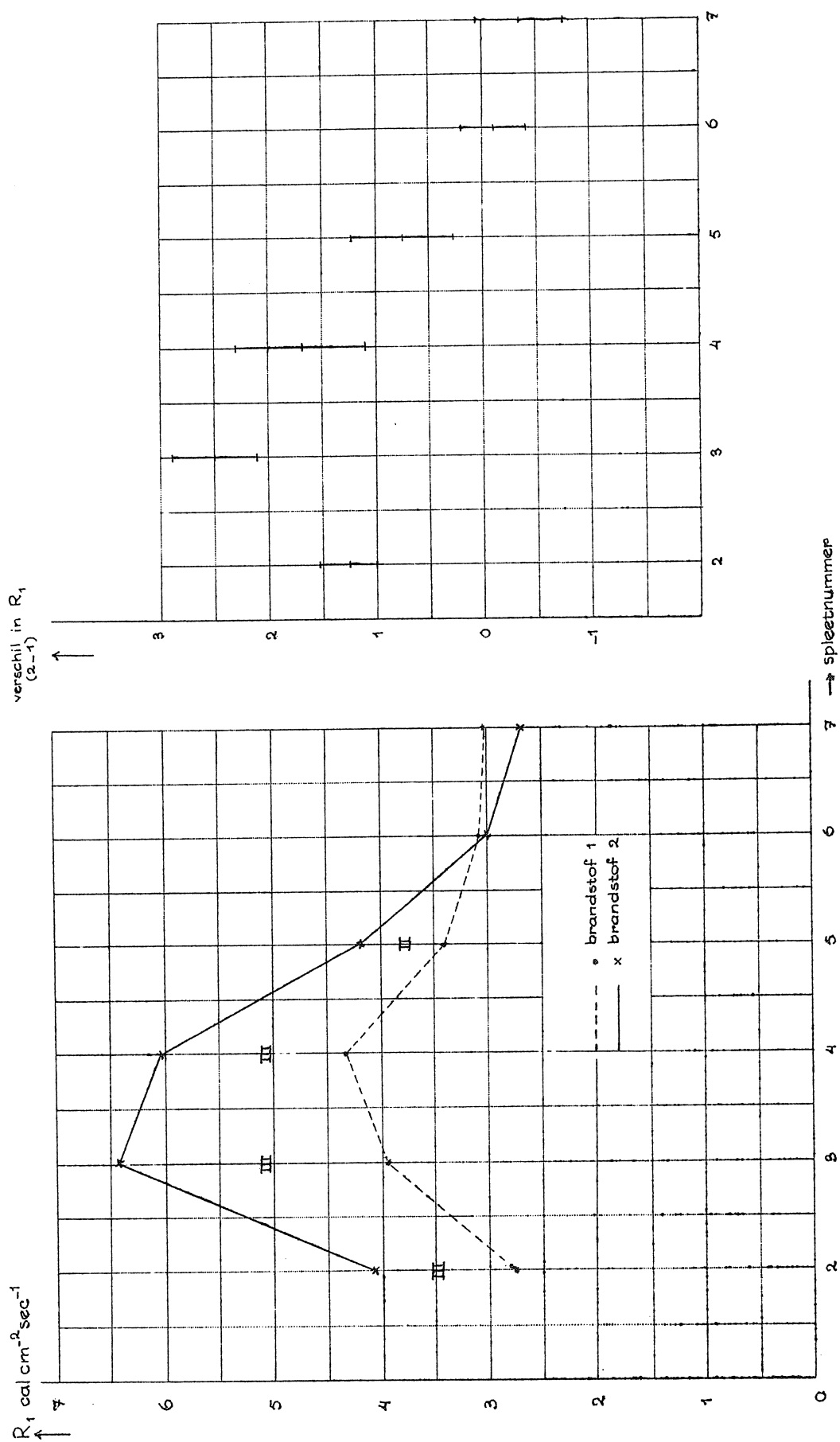
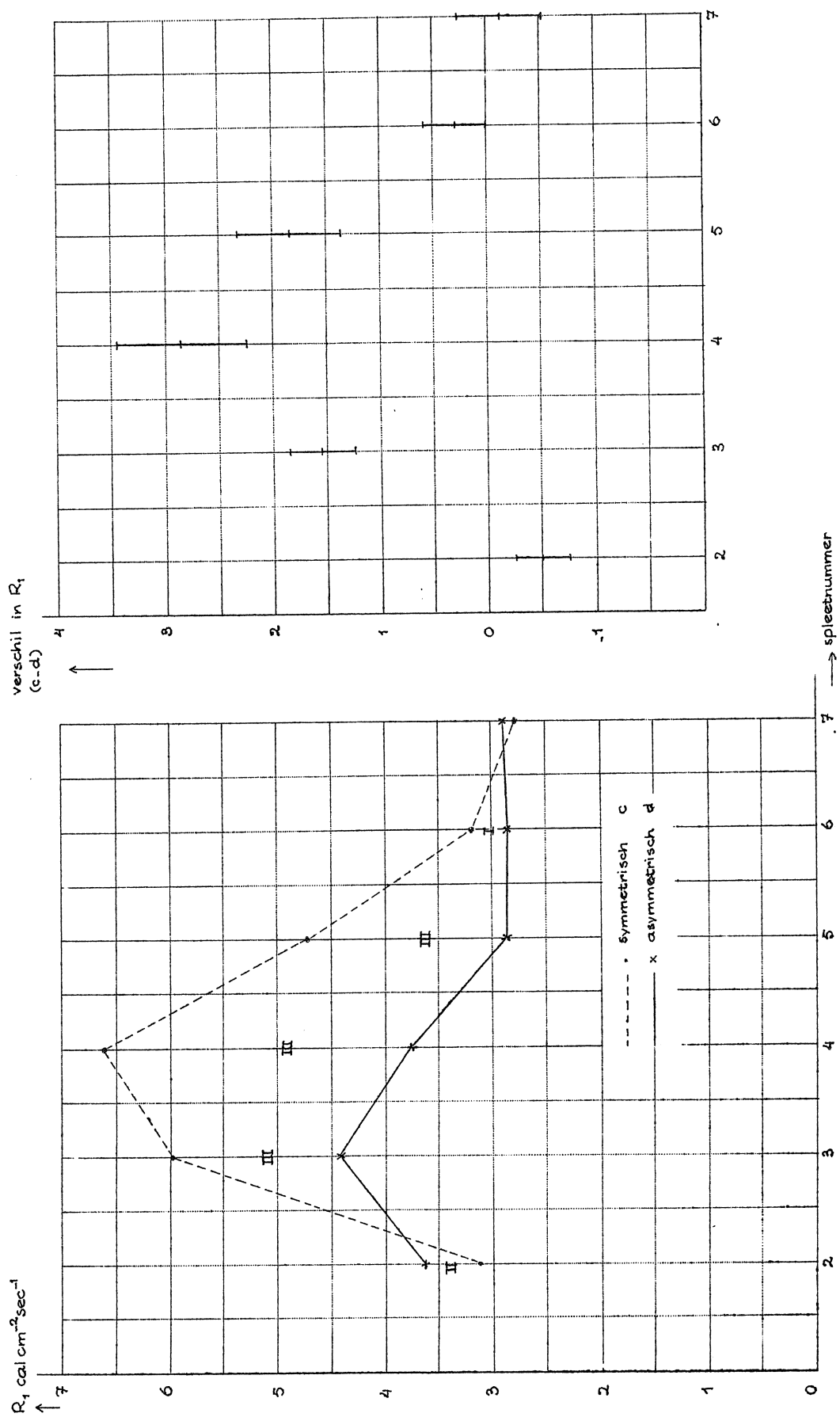


Fig.1



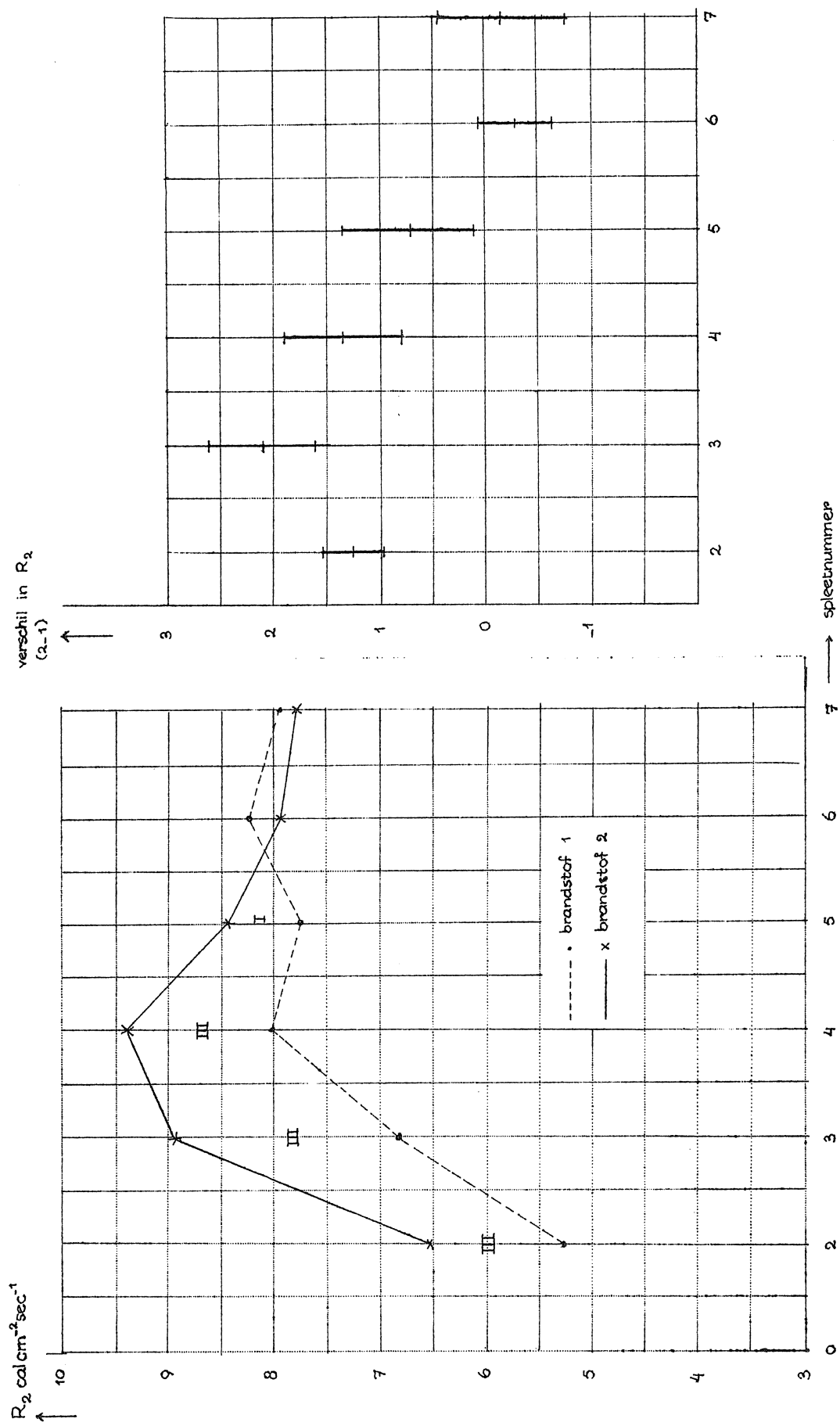


Fig. 3

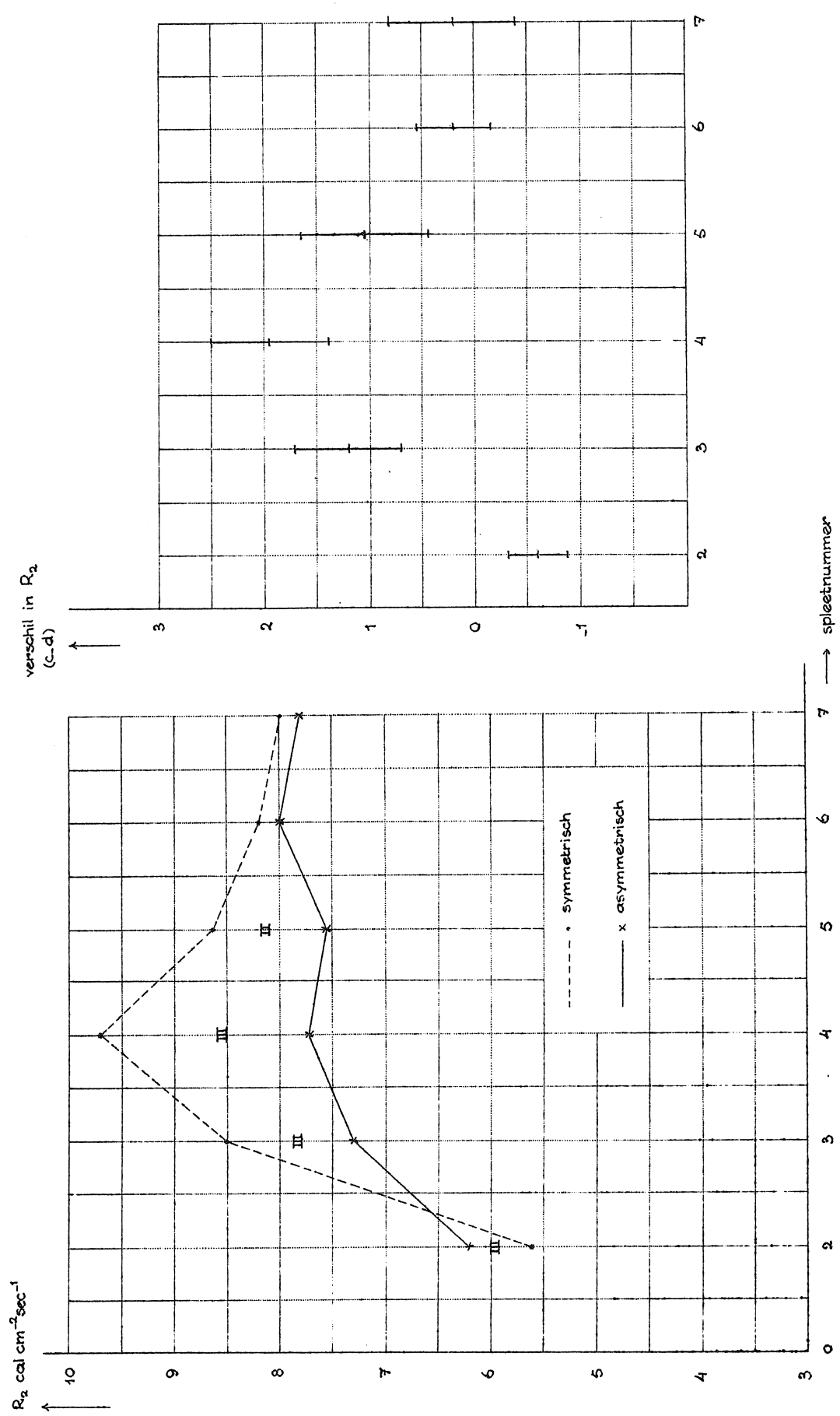


Fig. 4

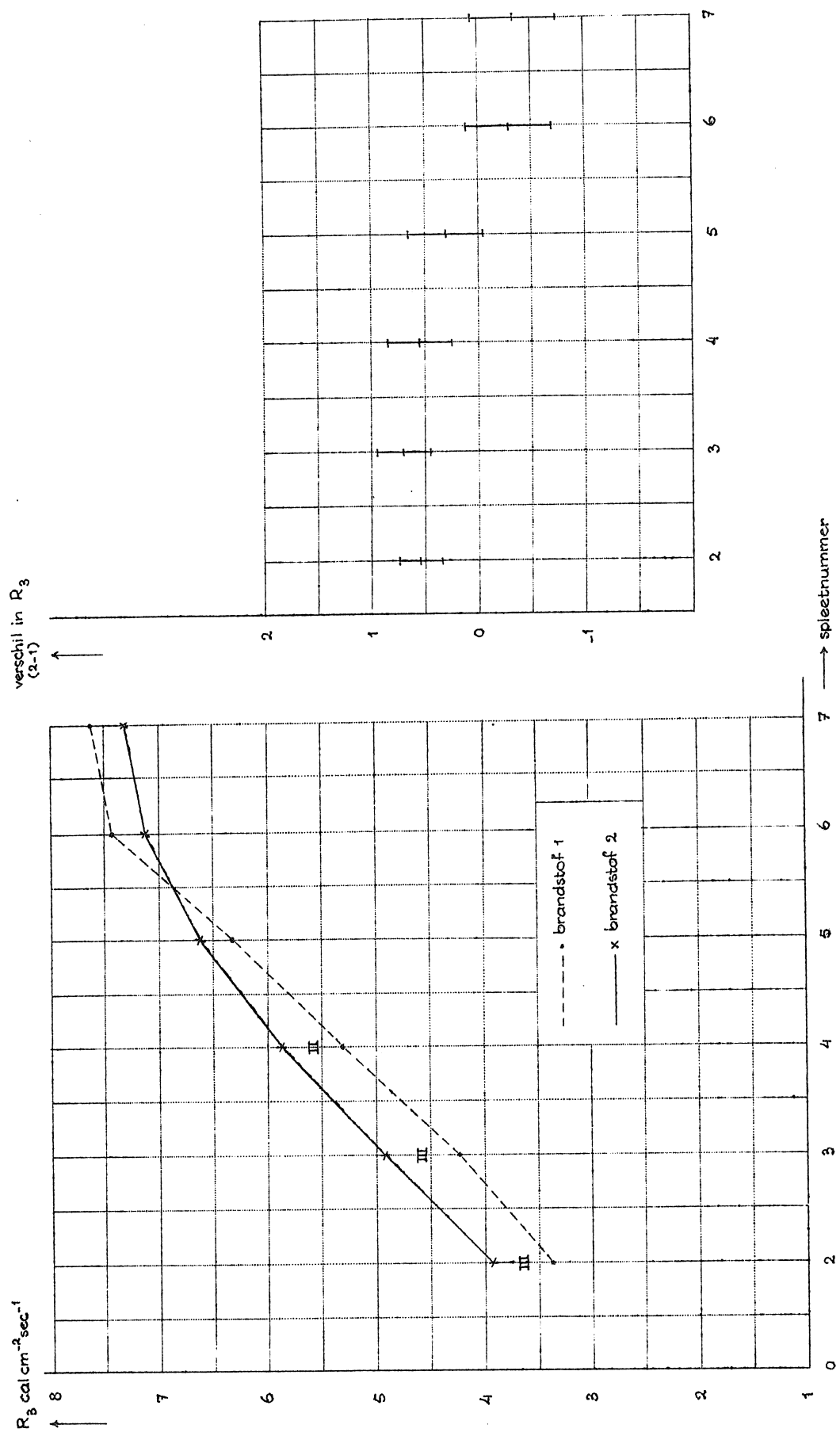
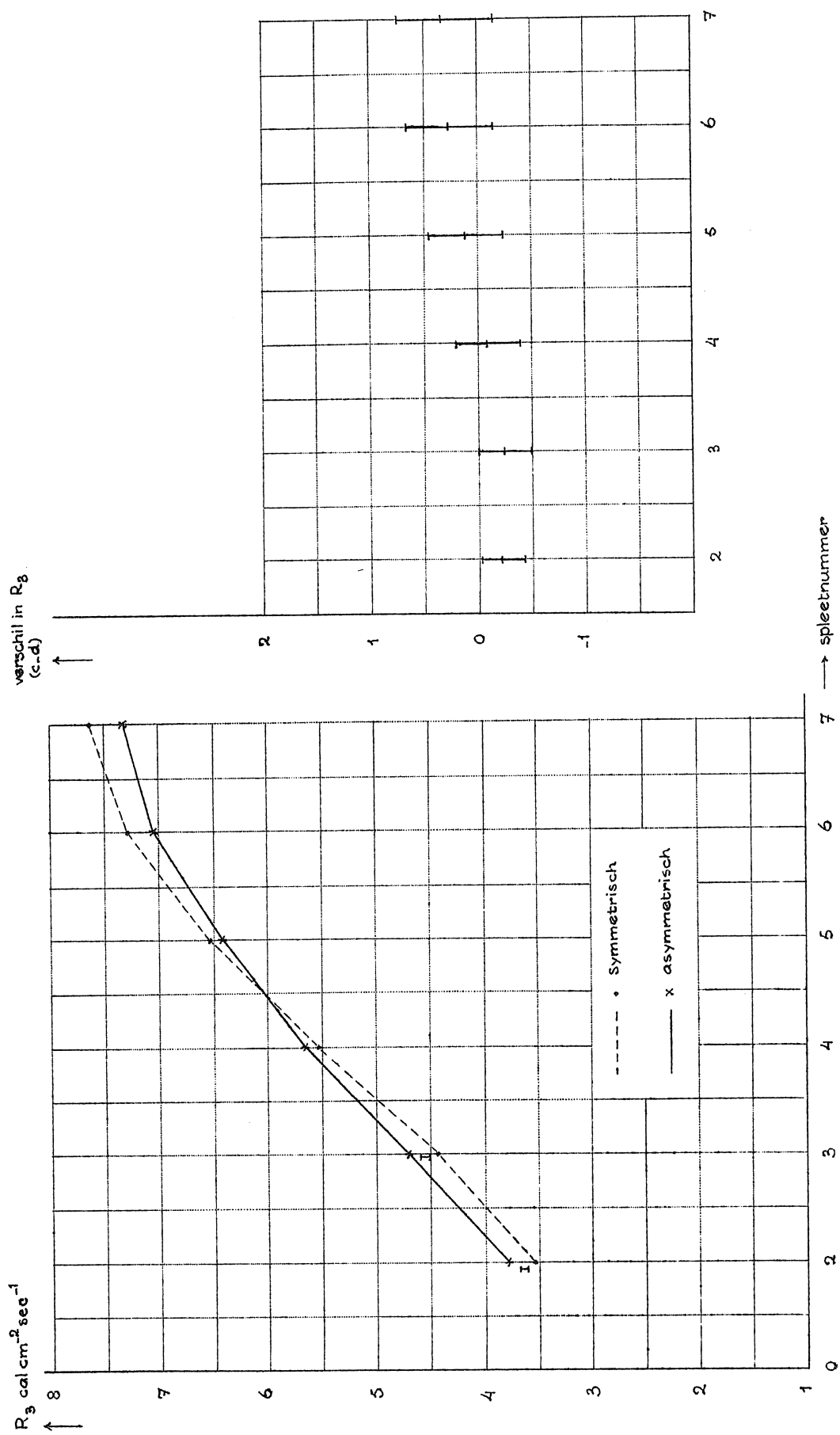


Fig. 5

Fig. 6



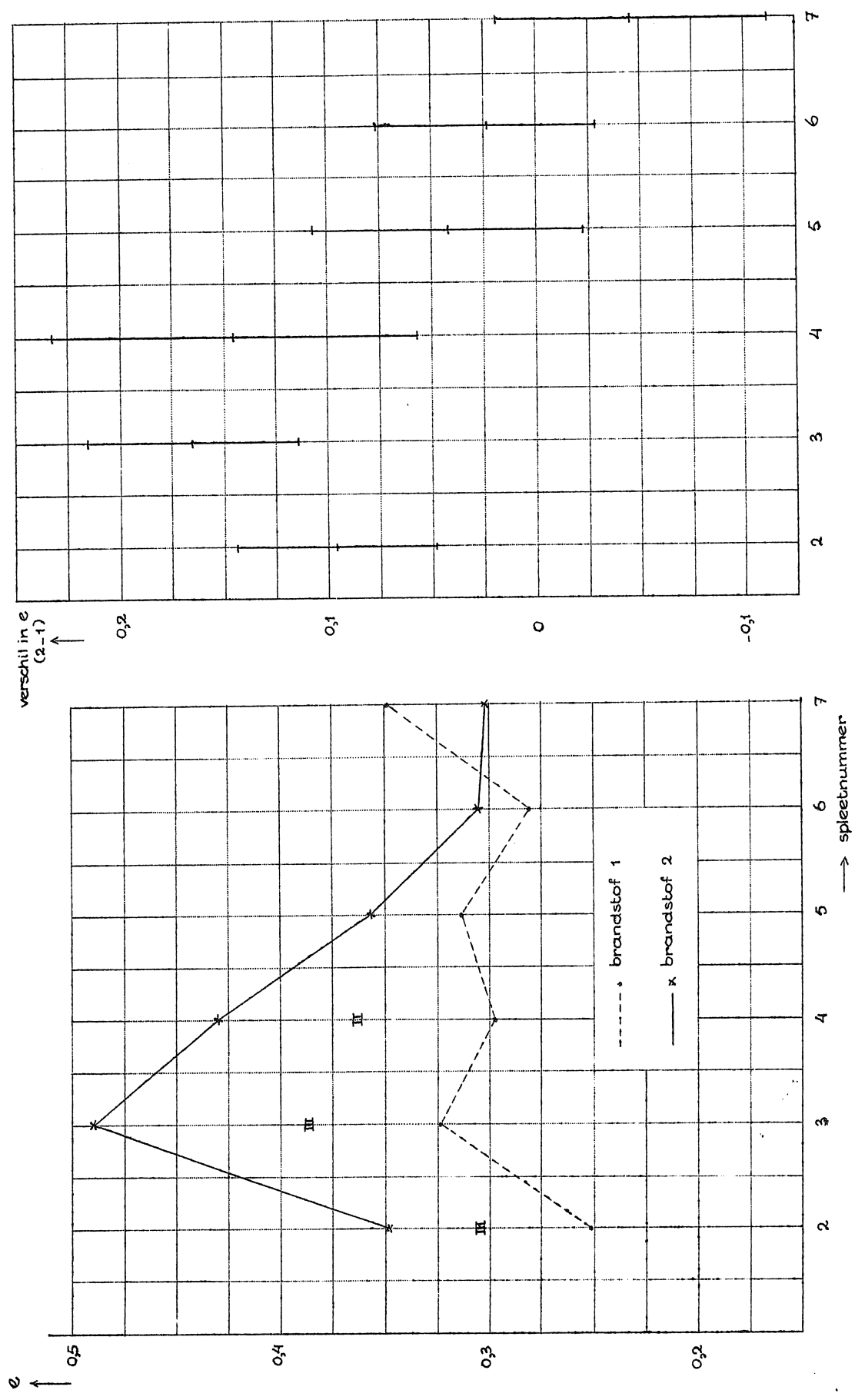


Fig.7

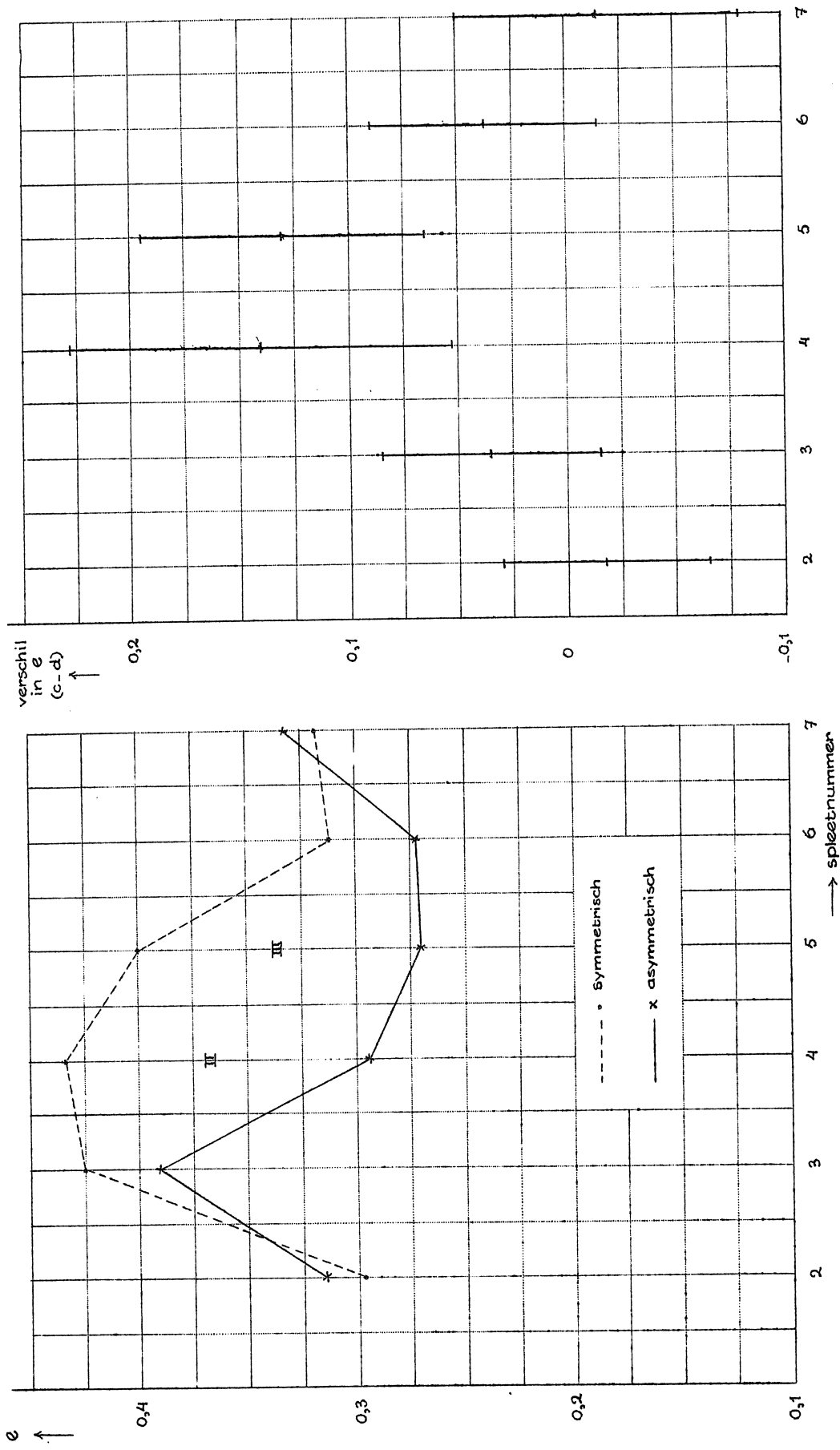


Fig. 8

Fig. 9

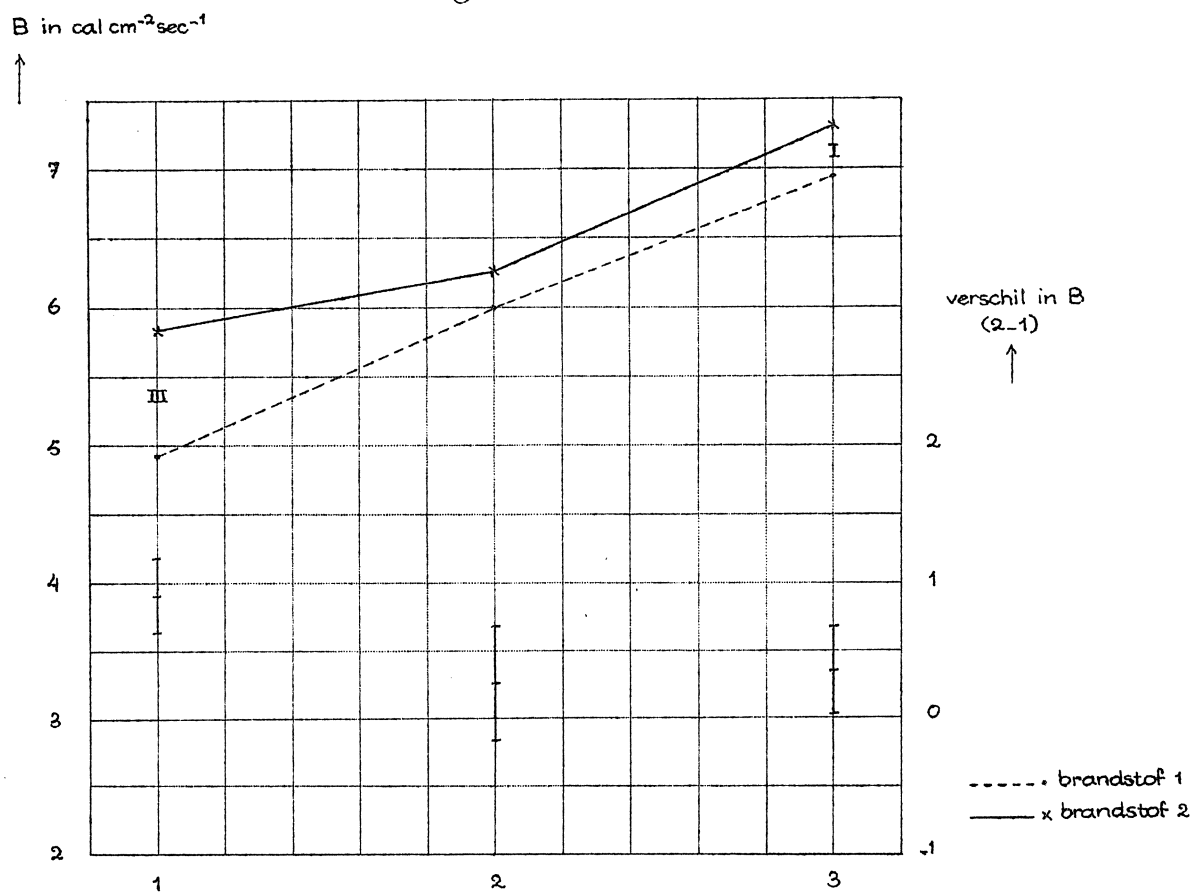
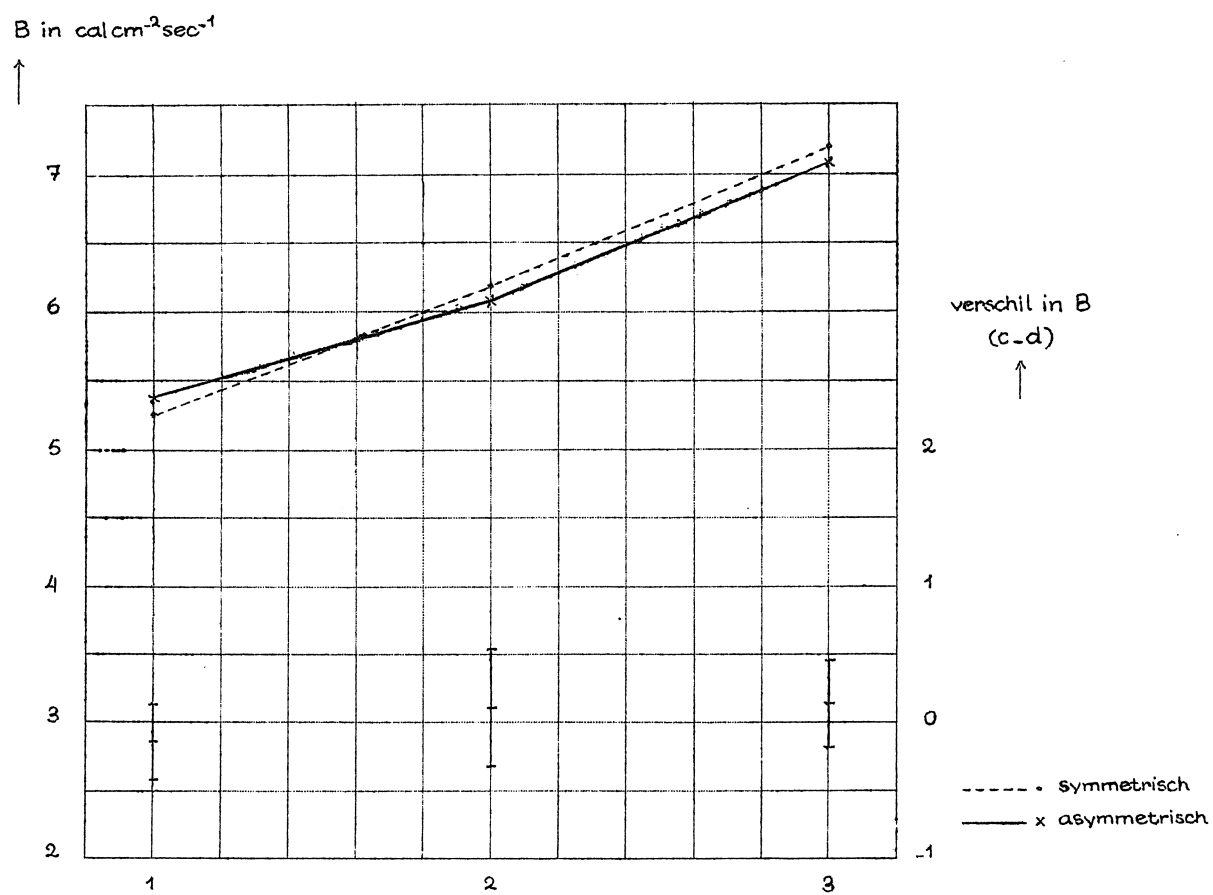


Fig. 10



Methode der m rangschikkingen ¹⁾

Een duidelijk voorstelling van deze toetsingsmethode verkrijgt men door n elementen te beschouwen, die een bepaald kenmerk, eventueel in verschillende mate, bezitten. dit kenmerk wordt door m waarnemers beoordeeld en ieder van deze waarnemers rangschikt deze n elementen volgens zijn beoordeeling naar opklimmende waardering. Op deze wijze ontstaan m rijen van rangschikkingen. We willen nu een maat aangeven voor de overeenstemming tussen deze rangschikkingen, m.a.w. een maat voor de overeenstemming tussen de m beoordeelingen. De hypothese $\mathcal{H}_0$, die met deze methode getoetst kan worden, houdt in dat er geen overeenstemming tussen de waarnemers bestaat; precieser gezegd, dat alle rangschikkingen onafhankelijk van elkaar op toevallige wijze zijn ontstaan. Dit is b.v. het geval, als het betrokken kenmerk in werkelijkheid voor alle elementen dezelfde waarde bezit.

We kunnen de afleiding voor de maat van overeenstemming het eenvoudigst geven aan de hand van een voorbeeld.

elementen		A	B	C	D	E	F
rangnummers toegekend door waarnemer	a	5	4	1	6	3	2
	b	2	3	1	5	6	4
	c	4	1	6	3	2	5
	d	4	3	2	5	1	6
		15	11	10	19	12	17

De som van alle rangnummers is $\frac{1}{2} n m (n+1)$. Onder de hypothese $\mathcal{H}_0$ is het theoretische gemiddelde van iedere kolom:
 $\frac{1}{2} m (n+1)$

We beschouwen nu de afwijkingen van dit gemiddelde. In ons voorbeeld is het theoretisch kolomgemiddelde gelijk aan 14. De afwijkingen daarvan zijn

1 -3 -4 5 -2 3

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid

De som der kwadraten van deze afwijkingen noemen wij S .

In ons voorbeeld is $S = 64$.

Als alle m rangschikkingen gelijk zijn wordt het maximum van S bereikt.

Dit maximum is $\frac{1}{12} m^2 (n^3 - n)$.

We definiëren nu als coëfficiënt van overeenstemming

$$W = \frac{12 S}{m^2 (n^3 - n)}$$

In ons voorbeeld is $W = \frac{12 \times 64}{16 \times 210} = 0,229$.

W varieert dus tussen 0 en 1.

De verdeling van $\underline{S}$ onder de hypothese H_0 is exact berekend voor een aantal waarden van n en m [1], terwijl voor grote m en n benaderingen bekend zijn.

De meeste gebruikelijke benaderingen zijn de volgende.

1°. De χ^2 -benadering:

$\chi_r^2 = m(n-1)\underline{W} = \frac{12 S}{mn(n+1)}$ heeft voor $m \rightarrow \infty$ een χ^2 -verdeling met $n-1$ vrijheidsgraden ([1] pg. 84 [2] pg. 36-37).

2°. De z -benadering:

$\underline{V} = (m-1) \frac{\underline{W}}{1-\underline{W}}$ is bij benadering verdeeld als $\underline{F} = e^{2\underline{z}}$

($\underline{F}$ is de $\underline{F}$ van Snedecor, $\underline{z}$ de $\underline{z}$ van Fisher) met

$$\underline{V}_1 = n-1-\frac{2}{m}$$

$$\underline{V}_2 = (m-1) \underline{V}_1 \quad \text{vrijheidsgraden ([1] pg. 84 [2] pg. 33-36)}.$$

Met behulp van de verdelingen van $\underline{S}$ of $\underline{W}$ onder de hypothese H_0 , kan deze hypothese getoetst worden, waarbij H_0 verworpen wordt als $\underline{W}$ waarden dichtbij 1 (resp. $\underline{S}$ dichtbij $\frac{1}{12} m^2 (n^3 - n)$) aanneemt, de kritieke z -one is dus van de vorm $W \geq W_0$ (resp. $S \geq S_0$).

Het kan voorkomen dat de waarnemers geen onderscheid ontdekken in de mate waarin verschillende elementen het kenmerk bezitten. Ze geven deze elementen dan gelijke rangnummers.

Veronderstel, dat door een waarnemer geen onderscheid wordt gemaakt tussen de elementen, die de rangnummers 3 t/m 6 moeten dragen. Dan wordt als rangnummer van ieder van deze elementen het gemiddelde van de rangnummers $\frac{1}{4} (3 + 4 + 5 + 6) = 4\frac{1}{2}$ gebruikt.

Daar het maximum van $\underline{S}$ nu verandert, moeten wij een correctie op de formule voor $\underline{W}$ toepassen. Deze vindt men in [1] (pg. 82) en [2] (pg. 28-30). Eveneens veranderen dan de formules voor de χ^2 -benadering ([1] pg. 86, [2] pg. 37) en voor de z -benadering ([1] pg. 86 [2] pg. 34), doch deze correcties zijn van weinig betekenis, tenzij het aantal gelijken groot is.

Literatuur: [1]

M.G.Kendall, Rank correlation methods, London 1948, Hoofdstuk 6, pag. 80.

Tabel van de verdelingsfunctie van $\underline{S}$ voor:

$$n = 3 \quad m = 2 \text{ t/m } 10$$

$$n = 4 \quad m = 2 \text{ t/m } 6$$

$$n = 5 \quad m = 3$$

op pag. 146-149.

Tabel van de waarden van S , waarvan de overschrijdingskansen onder de hypothese H_0 gelijk zijn aan 0,05 of 0,01, berekend met behulp van de z -benadering voor:

$$n = 3 \quad m = 8, 9, 10, 12, 14, 15, 16, 18, 20$$

$$n = 4 \quad m = 4, 5, 6, 8, 10, 15, 20$$

$$n = 5 \text{ t/m } 7 \quad m = 3, 4, 5, 6, 8, 10, 15, 20$$

op pag. 150.

[2]

Ph.van Elteren, Methode der m rangschikkingen, Cursus "Parametervrije Methoden", Hoofdstuk II, Rapport S 59, Mathematisch Centrum (1951), Blz. 18-45.

Toets van Bartlett¹⁾

Stel we hebben N waarnemingen:

$$\begin{array}{ccccccc} x_{11} & x_{21} & \cdot & \cdot & \cdot & x_{n_1 1} \\ x_{21} & x_{22} & \cdot & \cdot & \cdot & x_{n_2 2} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ x_{1p} & x_{2p} & \cdot & \cdot & \cdot & x_{n_p p} \end{array}$$

die verdeeld zijn in p subgroepen, terwijl het aantal waarnemingen in de j^e subgroep n_j is.

Veronderstel verder dat ieder der subgroepen een steekproef is uit een normale verdeling. De te toetsen hypothese H_0 is dan dat de spreidingen der subgroepen gelijk zijn.

Is $\bar{x}_j$ het gemiddelde van de waarnemingen in de j^e subgroep, $\nu_j = n_j - 1$, $N' = \sum_{j=1}^p \nu_j = N - p$

$$s_j^2 = \frac{1}{\nu_j} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$s^2 = \sum_{j=1}^p \frac{\nu_j}{N'} s_j^2$$

dan definiëren wij:

$$L_1^* = \prod_{j=1}^p \left(\frac{s_j^2}{s^2} \right)^{\frac{\nu_j}{N'}}$$

L_1^* is dus de verhouding van het geometrisch en het rekenkundig gemiddelde van de varianties der subgroepen en is dus ≤ 1 .

Als toetsingsgrootte wordt nu genomen:

¹⁾ Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

$$M = -N' \lg L_1^* = N' \lg \sum_{j=1}^p \frac{\nu_j}{N'} s_j^2 - \sum_{j=1}^p \nu_j \lg s_j^2$$

en M is dus ≥ 0 .

Als H_0 juist is zal M in het algemeen klein zijn en de kritieke zone bestaat dus uit grote waarden van M .

Deze toets van Bartlett is analoog aan een enkele jaren eerder door J. Neymann en E.S. Pearson afgeleide toets, waarbij in de toetsingsgrootte n_j en N in plaats van ν_j en N' gebruikt worden. De met L_1^* overeenkomende grootte noemden zij L_1 . De toets wordt, in deze oudere vorm, L_1 -toets genoemd.

Litteratuur: H.O. Hartley: Testing the homogeneity of a set of estimated variances. Biometrika 31 (1940) p. 249.

Tabellen: H.O. Hartley en E.S. Pearson: Tables for testing the homogeneity of a set of estimated variances. Biometrika 33 (1946) p. 296.

Betrouwbaarheidsintervallen (algemeen).¹⁾

Zij x een stochastische grootheid, die een verdelingsfunctie bezit die, op een onbekende parameter θ na, geheel bekend is (θ kan bv. het gemiddelde van x zijn, of de spreiding of iets dergelijks), dan kan men de vraag stellen uit een aantal waarnemingen van x een schatting voor θ af te leiden.

Een betrouwbaarheidsinterval $\mathcal{I}$ voor θ is een interval, waarvan de grenzen afhankelijk zijn van de waarnemingen $x_1, \dots, x_n$ van x , en dat de eigenschap bezit, behoudens een zekere gegeven onbetrouwbaarheid α , de juiste waarde van θ te bevatten. Dit betekent, dat bij een serie bepalingen van betrouwbaarheidsintervallen slechts in ongeveer een fractie α van deze gevallen het interval $\mathcal{I}$ zo zal uitvallen, dat het θ niet bevat. Hierbij is dus θ constant en het interval $\mathcal{I}$ veranderlijk (en wel stochastisch). Hierin ligt het grote verschil met een zgn. voorspellingsinterval, d.i. een gegeven vast interval, waar een stochastisch punt met een zekere waarschijnlijkheid in valt.

Het algemene principe ter bepaling van een betrouwbaarheidsinterval is het volgende: zij $\mathcal{T}$ een toets voor de hypothese $\theta = \theta_0$ (vgl. S47(M6)), dan is $\mathcal{I}$ de verzameling van die waarden θ_0 die bij toepassing van $\mathcal{T}$ op grond van de gevonden waarnemingen $x_1, \dots, x_n$ niet voor verwerping in aanmerking komen. Is $\mathcal{T}$ toegepast met een onbetrouwbaarheidsdrempel α , dan is dit ook de onbetrouwbaarheidsdrempel van het betrouwbaarheidsinterval.

Litteratuur:

- M.G. Kendall, The Advanced Theory of Statistics, London 1946, deel II, p.62-84.
A.M. Mood, Introduction to the theory of Statistics, London 1950, p.220.
J. Neyman, First course in probability and statistics, N.Y.1950.

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

De toets van Wilcoxon.¹⁾

Deze methode dient tot het toetsen van de hypothese H_0 , inhoudende, dat twee steekproeven $x_1, \dots, x_n$ en $y_1, \dots, y_m$ afkomstig zijn uit één collectie (ook wel populatie of universum genaamd).

Voor het toetsen van de hypothese H_0 wordt gebruik gemaakt van een toetsingsgrootte $\underline{U}$ ²⁾, die als volgt uit de waarnemingen berekend wordt. Onderstellen we, dat de waarnemingen $x_1, \dots, x_n$ en $y_1, \dots, y_m$ naar opklimmende grootte gerangschikt zijn, dan bepalen we eerst het aantal waarnemingen uit de tweede steekproef, dat kleiner is dan de kleinste waarneming x_1 uit de eerste steekproef (bij gelijkheid tellen wij $\frac{1}{2}$ in plaats van 1). Noem dit aantal V_1 . Vervolgens wordt het aantal waarnemingen uit de tweede steekproef bepaald, dat kleiner is dan de op één na kleinste waarneming x_2 uit de eerste steekproef (bij gelijkheid wordt weer $\frac{1}{2}$ in plaats van 1 geteld). Dit aantal noemen we V_2 . Evenzo worden met betrekking tot $x_3, x_4, \dots, x_n$ de aantallen $V_3, V_4, \dots, V_n$ bepaald. De waarde $\underline{U}$ van de toetsingsgrootte $\underline{U}$ wordt voor de twee steekproeven dan gegeven door

$$\underline{U} = V_1 + V_2 + \dots + V_n.$$

Wanneer onder de waarnemingen niet te veel gelijken voorkomen, kan bewezen worden, dat de toetsingsgrootte $\underline{U}$ onder de hypothese H_0 voor grote waarden van n en m (beide ≥ 10) bij benadering een normale verdeling bezit. De waarnemingen $x_1, \dots, x_n$ en $y_1, \dots, y_m$ tezamen genomen vallen uiteen in een aantal groepen van gelijke waarnemingen. Noem het aantal van deze groepen k , dan is k minstens 1 (als alle waarnemingen gelijk zijn) en hoogstens $m+n$ (als alle waarnemingen verschillend zijn).

¹⁾ Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

²⁾ Stochastische grootheden worden door onderstreeping aangeduid.

Zijn $t_1, \dots, t_k$ de aantallen waarnemingen in deze groepen van gelijken, dan worden het gemiddelde μ en de variantie σ^2 van de toetsingsgrootte $\underline{U}$ gegeven door

$$\mu(\underline{U}) = \frac{1}{2}nm,$$

en

$$\sigma^2 = \text{Var}(\underline{U}) = \frac{1}{12} \frac{nm}{(n+m)(n+m-1)} \left\{ (n+m)^3 - (t_1^3 + t_2^3 + \dots + t_k^3) \right\} \quad 1)$$

De grootte $\mu(\underline{U})$ is dus onafhankelijk van de waarden vast. Indien de hypothese H_0 niet vervuld is, zal de grootte $\underline{U}$ grote of kleine waarden bezitten, al naar gelang $\underline{y}$ systematisch kleiner of groter is dan $\underline{x}$.

De (tweezijdige) toets bestaat nu daarin, dat men H_0 verworpt indien de gevonden waarde $\underline{U}$ van $\underline{U}$ te sterk van μ afwijkt, d.w.z. als

$$\frac{|\underline{U} - \mu|}{\sigma} > \frac{z_{\alpha}}{2}, \quad 2)$$

waarin α de onbetrouwbaarheidsdrempel is en z_{α} volgt uit

$$\frac{1}{\sqrt{2\pi}} \int_{\frac{z_{\alpha}}{2}}^{\infty} e^{-\frac{1}{2}x^2} dx = \frac{1}{2} \alpha,$$

en in een tabel van de normale verdeling kan worden opgezocht.

De (tweezijdige) overschrijdingskans k , behorende bij T , is gedefinieerd als

$$k = \frac{2}{\sqrt{2\pi}} \int_{\left| \frac{\underline{U} - \mu}{\sigma} \right|}^{\infty} e^{-\frac{1}{2}x^2} dx \quad 2)$$

en kan ook in een tabel van de normale verdeling worden gevonden.

Bij eenzijdige toetsing wordt α door 2α vervangen, resp. k gehalveerd.

Een bijzonder geval van het bovenstaande is, dat onder de waarnemingen voor $\underline{x}$ en $\underline{y}$ in 't geheel geen gelijken voorkomen. In dat geval kan de uitdrukking voor de variantie herleid worden tot

$$\sigma^2 = \frac{1}{12} nm(n+m+1).$$

1) Deze formule is een door T.J.Terpstra gegeven vereenvoudiging van de door J.Hemelrijk ([5] en [7]) afgeleide formule. De afleiding van deze vereenvoudigde formule zal nog gepubliceerd worden.

2) Deze formules berusten op de normale benadering van de verdeling van $\underline{U}$.

Indien n en m kleiner zijn dan 10, zijn tabellen beschikbaar voor het berekenen van de overschrijdingskans k voor de uit de steekproef bepaalde waarde U van $\underline{U}$ (zie [2] en [4]). Dergelijke tabellen bestaan echter niet voor het geval, dat gelijke waarnemingen optreden.

Opmerking. Men kan gemakkelijk bewijzen, dat de variantie van $\underline{U}$ door het optreden van gelijke waarnemingen vermindert. Het verschil, dat door deze gelijken optreedt, is echter in het algemeen gering. Men kan daarom in eerste instantie deze correctie op σ^2 verwaarlozen. De overschrijdingskansen, die men dan vindt, zijn iets te groot.

Litteratuur:

1. F.Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945), p.80-83.
- 2 H.B.Mann and D.R.Whitney On a test of whether one of two random variables is stochastically larger than the other, Amer.Math.Stat. 18 (1947), p. 50-60.
- 3 H.R.van der Vaart Some remarks on the power function of Wilcoxon's test for the problem of two samples, Proceedings van de Kon. Ned.Ak.v.Wet., 53 (1950), p. 494-520.
- 4 H.R.van der Vaart Gebruiksaanwijzing voor de toets van Wilcoxon, met tabellen voor n en $m \leq 10$, Rapport S32 (M4) (1950).
- 5 H.R.van der Vaart De toets van Wilcoxon voor het probleem van twee steekproeven. (Cursus "Parametervrije Methoden", 1951-'52).
- 6 D.van Dantzig Kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam (1947-'50), hoofdst. 6, § 3.
- 7 J.Hemelrijk Note on Wilcoxon's two sample test, when ties are present, Ann.Math.Stat. 23 (1952) no. 2.