

PROEFSTATION VOOR DE AKKER- EN WEIDEBOUW
WAGENINGEN

EEN GRAFISCHE METHODE VAN BEWERKING, TOEGEPAST
OP GEGEVENS VAN KOSTPRIJGELEDRIJVEN

Drs. C. Postma (P.A.W.)

en

Ir. D. van der Schaaf (I.B.V.L.)

MEDEDELING NR. 34

FEBRUARI 1960

Gezamenlijke uitgave van het
Proefstation voor de Akker- en Weidebouw en het
Instituut voor Bewaring en Verwerking van Landbouwprodukten

I N H O U D S O P G A V E

	Blz.
Woord vooraf	3
<u>I. Grafische of rekenkundige methode</u>	4
1. De vereffening van een rechte lijn	4
2. De vereffening bij afhankelijkheid van twee factoren	6
2.1. Geen correlatie tussen x_1 en x_2	7
2.1.1. Twee-dimensionale werkwijze	7
2.1.2. Drie-dimensionale werkwijze	12
2.2. Correlatie tussen x_1 en x_2	14
<u>II. De gegevens der kostprijsbedrijven</u>	16
Samenvatting	19

WOORD VOORAF

De aanleiding tot deze publikatie voert terug tot een lezing, die de tweede schrijver dezes hield voor onderzoekers van de instituten P.A.W., I.B.S. en I.B.V.L. in de winter van 1957 op 1958.

Aan de hand van kostprijscijfers van zuivere weidebedrijven werd toen een aantal factoren aangewezen dat de winst van die bedrijven bepaalde. Het voor de aankoop van voeder bestede bedrag bleek één van die factoren te zijn. Hoe hoger de voederkosten waren, des te lager was de winst. Hieruit kwam ook in dit geval de betekenis van de winning van voldoende en goed wintervoer aan het licht.

De analyse van de cijfers der kostprijsbedrijven was gedaan door ze grafisch te bewerken. Tijdens de bespreking rees er twijfel of de grafische bewerkingsmethode wel tot goede uitkomsten zou leiden. Daarom is naderhand contact opgenomen met de eerste auteur die het materiaal aan een regressieberekening onderwierp.

Gezien het feit, dat de resultaten van de beide bewerkingen een grote overeenstemming vertoonden, werd het nuttig gevonden de hier volgende beschouwingen en voorbeelden in bredere kring onder de aandacht te brengen. De gebruikte grafische methode wordt aan de hand van enige voorbeelden toegelicht.

I. GRAFISCHE OF REKENKUNDIGE METHODE

De resultaten van een grafische bewerking worden door velen met enige achterdocht beschouwd, terwijl anderen daarentegen de waarde van deze methoden overschatten. De oorzaak van eerstgenoemde houding spruit dikwijls voort uit onbekendheid met de mogelijkheden van de grafische methoden, welke uiteraard niet de grote exactheid van de rekenkundige methodieken bezitten. Deze uiterste nauwkeurigheid is echter voor vele problemen niet noodzakelijk; in het bijzonder niet bij oriënterende onderzoekingen. Bij grafisch werken is geen indruk te verkrijgen van de nauwkeurigheid der verkregen uitkomsten. Voor exacte uitkomsten is natuurlijk de berekening de aangewezen methode. Een bezwaar van de grafische methodiek is, dat hierbij de waarde van de verschillen gemakkelijk wordt overschat.

I.1. De vereffening van een rechte lijn

Als voorbeelden zullen allereerst enige gevallen van toenemende ingewikkeldheid worden behandeld. Deze betreffen steeds gevallen van lineaire relaties. De vergelijking van een rechte lijn luidt, zoals algemeen bekend is :

$$y = m x + q$$

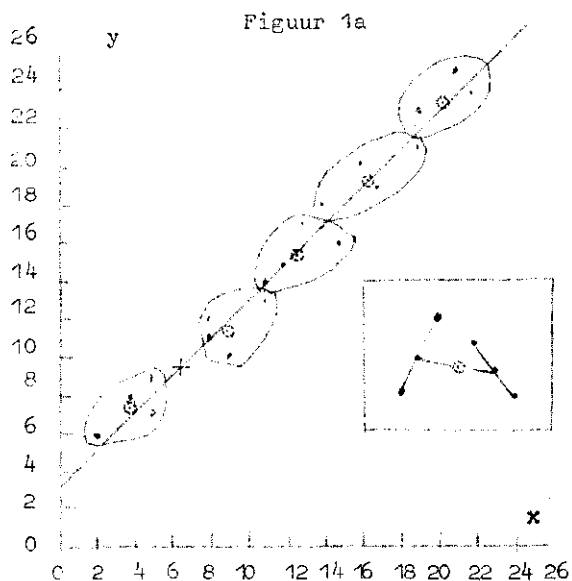
Indien men bij onderzoek te maken krijgt met twee factoren, die lineair met elkaar samenhangen, meet men een aantal waarden van y en de bijbehorende waarden van x . De vereffening heeft dan tot doel de (gemiddelde) waarden van m en q te vinden. In tabel 1 is een voorbeeld gegeven van een aantal waarnemingen van y bij verschillende waarden van x .

Tabel 1.

Eerste voorbeeld

y	x
20	16
16	15
10	9
22	19
8	4
23	19
9	5
17	13
6	2
12	8
24	22
11	8
7	5
19	17
14	11
18	14
15	12
21	19
13	11
25	21

De grafische manier van behandelen is nu, dat op millimeterpapier de x - en y -waarden in een assenstelsel tegen elkaar worden uitgezet. Zo ontstaat de puntenzwerm van figuur 1a. Hierna zijn de 20 punten verdeeld in 5 groepen van vier punten. Van elke groep wordt met behulp van een liniaal het zwaartepunt bepaald door eerst het midden tussen



twee van de vier punten met een stip aan te duiden en dan het midden tussen de twee overgebleven punten met een stip te merken. Vervolgens is het midden tussen de twee stippen het zwaartepunt van de groep *). Dit is rechts in figuur 1a weergegeven. Figuur 1a geeft de vijf zwaartepunten der groepen als cirkeltjes met een stip erin. Omdat de vijf zwaartepunten niet helemaal op een rechte lijn liggen, is het midden van de twee zwaartepunten links beneden in de figuur bovendien bepaald en gemerkt met een kruisje. Het kruisje en de overige drie zwaartepunten liggen nagenoeg op de rechte lijn, die is getekend.

Deze lijn wordt voorgesteld door de formule $y = m x + q$.

Berekening van m vindt plaats door de helling van de lijn te meten.

$$\text{Dus : } m = \frac{26,9 - 3,3}{24} = \frac{23,6}{24} = 0,98$$

De grootte van q leest men af op de lijn voor $x = 0$, dus daar waar de lijn de y -as snijdt, hier is $q = 3,3$.

De regressierekening- de rekenkundige manier dus om m en q te weten te komen - levert in de twee gevallen, namelijk dat x zeker en y onzeker en dat x onzeker en y even onzeker is, uitkomsten zoals in onderstaande tabel 2 is weergegeven.

Tabel 2. Grafisch en rekenkundig bepaalde waarden van m en q

Grafisch	Rekenkundig		
	x zeker y onzeker	x en y even onzeker	
m 0,98	0,97 \pm 0,042	0,99 \pm 0,011	
q 3,3	3,4 \pm 0,58	3,2 \pm 0,47	

Beide methoden geven in dit geval vrijwel hetzelfde resultaat. Het voordeel van de rekenkundige manier is echter, dat de standaardafwijkingen tevens bekend zijn en wij dus getallen vinden, die de nauwkeurigheid van de uitkomsten aangeven. In beide gevallen is evenwel de correlatie-coëfficiënt tussen de volgens de genoemde formule berekende en de waargenomen y te bepalen.

Een tweede soortgelijk voorbeeld vormt de cijferreeks van tabel 3. Te werk gaande als bij het eerste voorbeeld, vindt men de zwaartepunten van de vijf groepen van vier punten. Deze zwaartepunten blijken vrijwel op de rechte lijn te liggen, die in figuur 1b door de puntenzwerm is getekend. Uit deze figuur zijn weer een m en q te berekenen van de vergelijking :

$$y = m x + q \quad m = \frac{27,3 - 7,6}{40} = \frac{19,7}{40} = 0,49$$

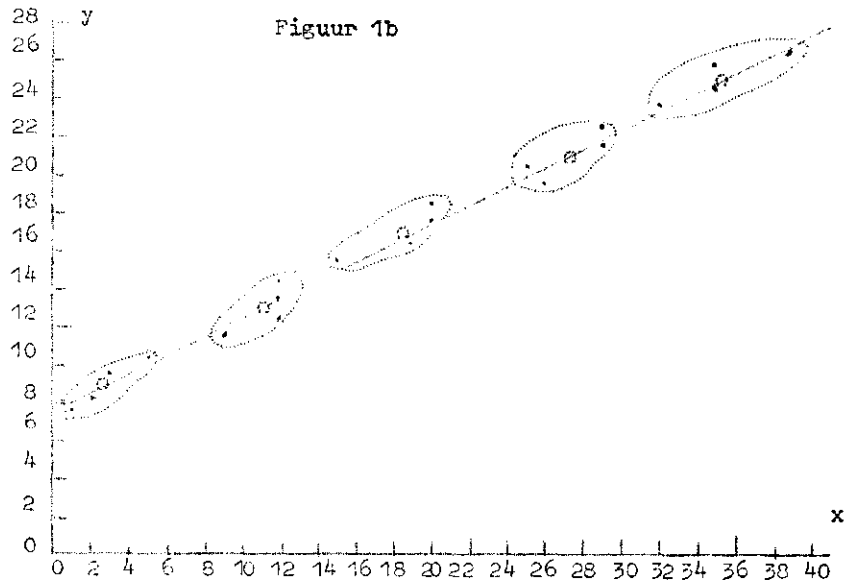
$$q = 7,6$$

*) Het is onverschillig op welke wijze wij de vier punten in groepen verdelen. Immers indien de vier punten worden weergegeven door de vier vectoren $\bar{a}, \bar{b}, \bar{c},$ en \bar{d} , dan is het midden van \bar{a} en \bar{b} het punt $\frac{1}{2}(\bar{a} + \bar{b})$ en dat van \bar{c} en \bar{d} het punt $\frac{1}{2}(\bar{c} + \bar{d})$. Het midden van de verbindingslijn van deze middens is dus $\frac{1}{4}(\bar{a} + \bar{b} + \bar{c} + \bar{d})$, waarin $\bar{a}, \bar{b}, \bar{c},$ en \bar{d} willekeurig in paren zijn te groeperen.

Tabel 3.

Tweede voorbeeld

q	x
21,5	29
17,5	20
11,5	9
23,5	32
9,5	3
24,5	35
10,5	5
18,5	20
7,5	1
13,5	12
25,5	35
12,5	12
8,5	2
20,5	25
15,5	15
19,5	26
16,5	19
22,5	29
14,5	12
26,5	39



De regressieberekening levert uitkomsten zoals vermeld in tabel 4.

Tabel 4. Grafisch en rekenkundig bepaalde waarden van m en q

Grafisch		Rekenkundig	
		x zeker y onzeker	x en y even onzeker
m	0,49	0,49 ± 0,012	0,50 ± 0,012
q	7,6	7,6 ± 0,27	7,6 ± 0,27

Ook het tweede voorbeeld demonstreert duidelijk een goede overeenstemming tussen de uitkomsten met beide methoden verkregen. Dit is echter in beide voorbeelden mede veroorzaakt door de grote overeenstemming tussen de waargenomen y en die volgens de gevonden lijn, zoals uit de figuren blijkt.

1.2. De vereffening bij afhankelijkheid van twee factoren

De vergelijking in het geval dat een afhankelijke grootte lineair wordt bepaald door twee factoren is :

$$y = m_1 x_1 + m_2 x_2 + q.$$

De vereffening dient de waarden van m_1 , m_2 en q te geven. Er kunnen verder twee mogelijkheden worden onderscheiden :

- 1e : x_1 en x_2 zijn onderling vrijwel niet gecorreleerd
- 2e : x_1 en x_2 zijn onderling gecorreleerd.

I.2.1. Geen correlatie tussen x_1 en x_2

De grafische behandeling van dit vraagstuk is op twee manieren te doen. De ene manier is het werken met twee dimensionale figuren, net als bij de vorige twee voorbeelden. De andere manier maakt gebruik van methoden zoals deze in de beschrijvende meetkunde worden toegepast (orthogonale projectie). Evenals bij de eerder gegeven voorbeelden, komt de twee dimensionale grafische werkwijze eerst in behandeling. De cijferreeksen van tabel 5 dienen als voorbeeld. De reeksen van tabel 5 zijn ontstaan door een willekeurige samenvatting van die der beide vorige voorbeelden.

Tabel 5

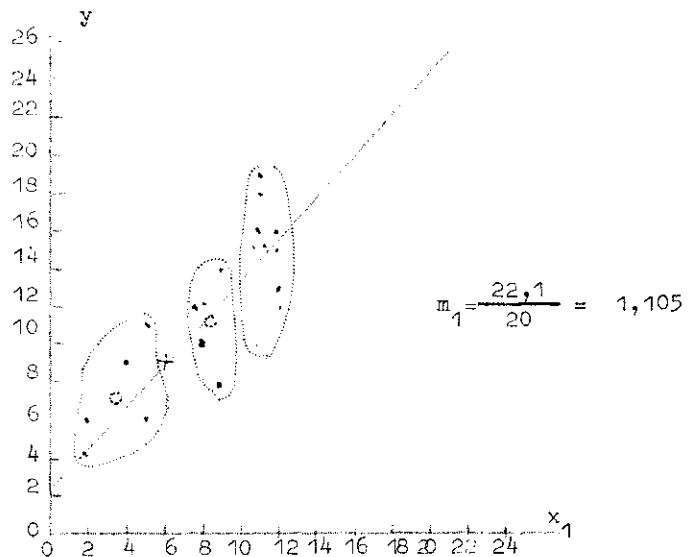
Derde voorbeeld

y	x_1	x_2
4	2	3
8	9	3
19	11	19
6	5	5
10	8	3
18	11	15
9	4	12
16	12	12
14	9	15
13	12	3
12	8	9
12	12	2
15	12	9
11	5	12
16	11	12
15	11	12
10	11	1
6	2	9
15	11	12
12	8	12

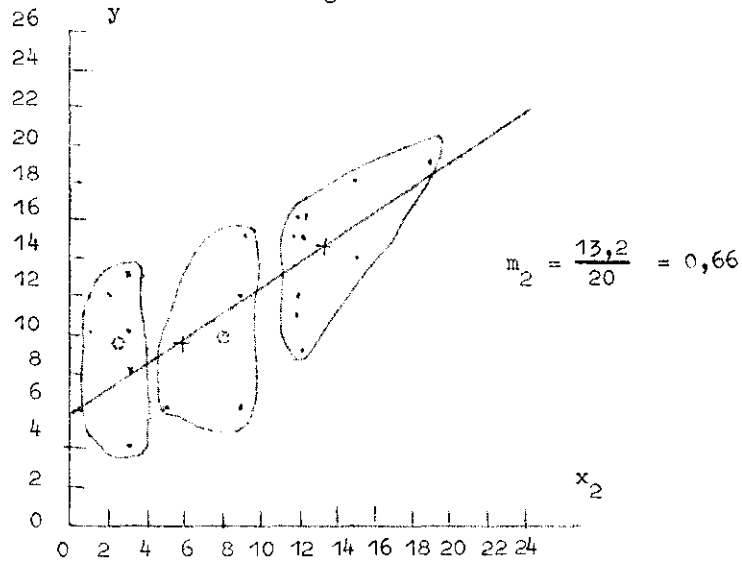
I.2.1.1. Twee-dimensionale werkwijze

Bij deze methode wordt y afzonderlijk uitgezet tegen x_1 (fig. 2a) en tegen x_2 (fig. 2b). Tevens zet men x_1 en x_2 tegen elkaar uit, om de onderlinge correlatie vast te stellen (fig. 2c). Figuur 2c laat zien, dat tussen x_1 en x_2 zeer weinig correlatie bestaat. Verder is de correlatie tussen y en x_1 het duidelijkst. Daarom wordt in figuur 2a de "gemiddelde" lijn geconstrueerd. De punten zijn hiervoor ingedeeld in twee groepen van vijf punten en één groep van tien. De groep van tien punten levert een zwaartepunt op en de zwaartepunten van de twee groepen van vijf punten geven eveneens een gemiddeld zwaartepunt van tien (het kruisje in de figuur).

Figuur 2a

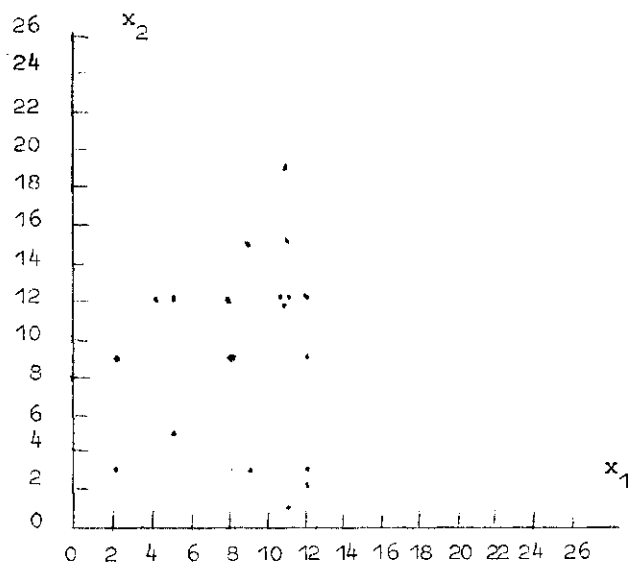


Figuur 2b



De gemiddelde lijn is nu te trekken. Deze lijn van figuur 2a is een eerste benadering voor het meten van de juiste waarde van m_1 , die hier 1,105 is. Ter verkrijging van m_2 wordt nu de variatie van y door x_1 veroorzaakt, zoveel mogelijk geëlimineerd.

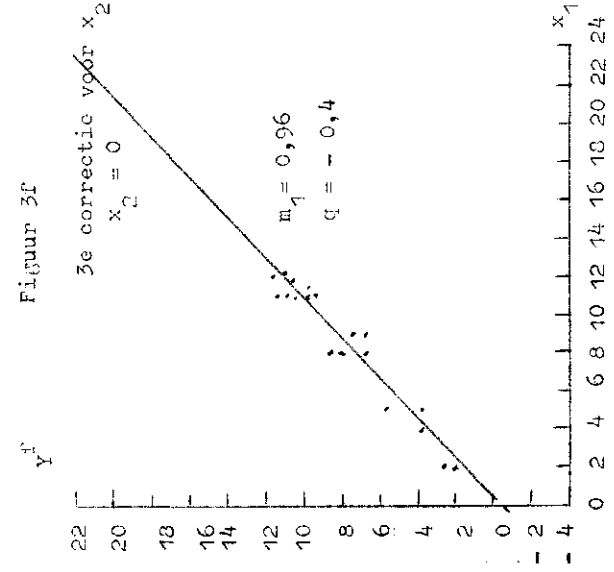
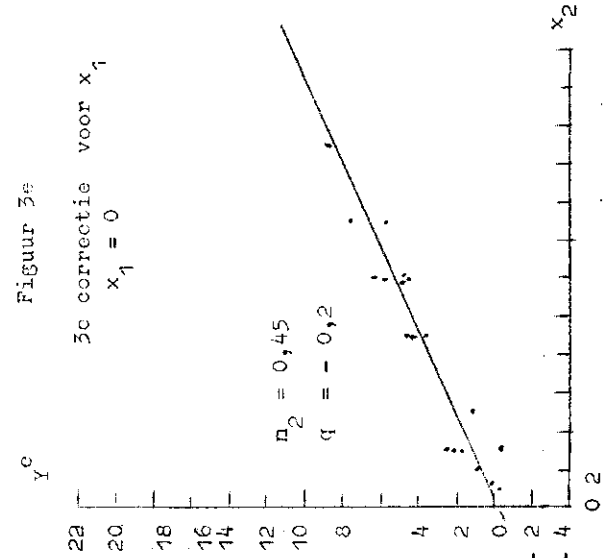
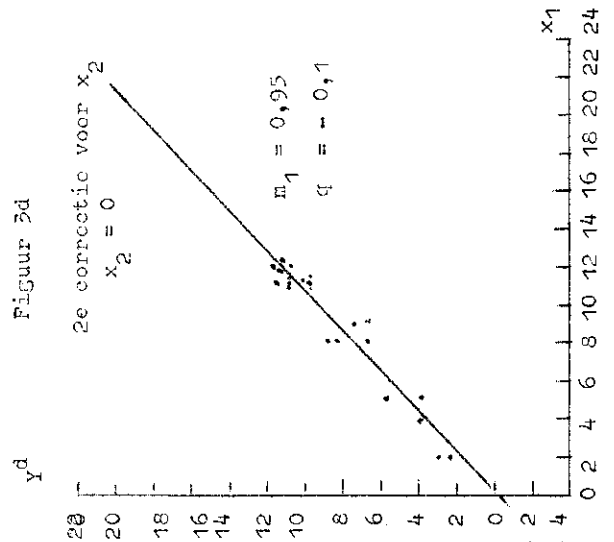
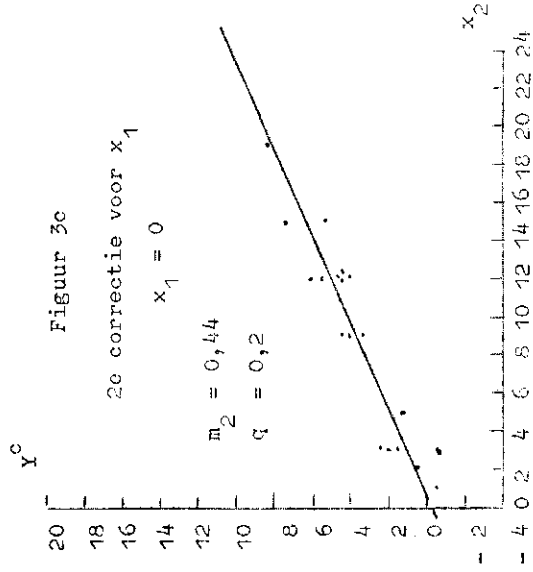
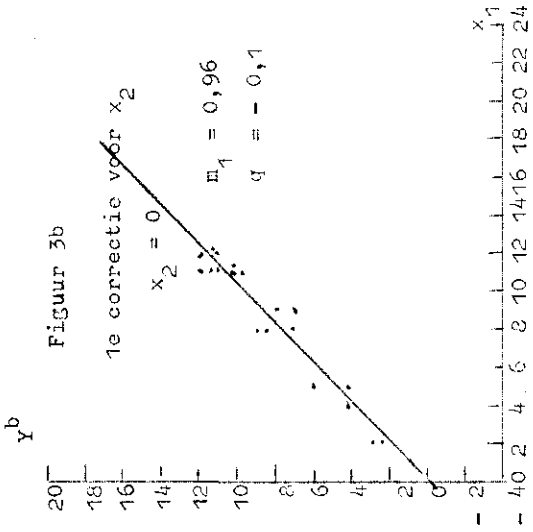
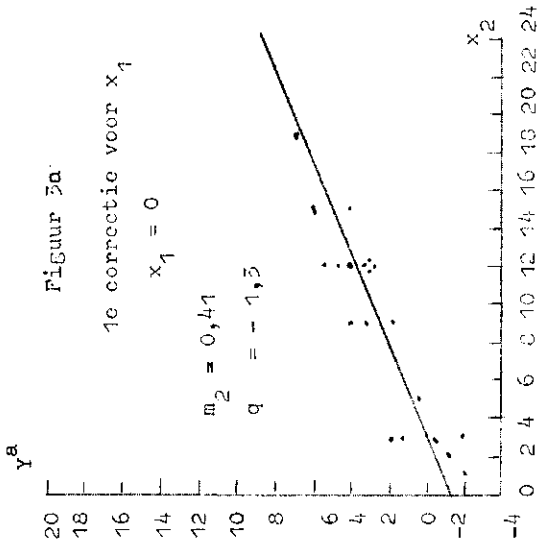
Figuur 2c



Tabel 6

Eliminatie van de invloed van x_1 op y en van x_2 op y

Rijnum-	y	x_1	x_2	x_1 op y correctie	de x_1 correctie	x_2 op y correctie	de x_2 correctie	x_1 op x_2 correctie	de x_1 correctie	x_2 op x_1 correctie	de x_2 correctie	x_1 op x_2 correctie	de x_1 correctie	x_2 op x_1 correctie	de x_2 correctie
1	4	2	3	2,1	1,9	1,2	2,8	2,0	2,0	1,3	2,7	1,8	2,2	1,4	2,6
2	8	9	3	9,9	-1,9	1,2	6,8	-0,7	8,7	1,3	6,7	8,5	-0,5	1,4	6,6
3	19	11	19	12,1	6,9	7,8	11,2	8,4	10,6	8,3	10,7	10,3	8,7	8,6	10,4
4	6	5	5	5,5	0,5	2,0	4,0	1,1	4,9	2,2	3,8	4,7	1,3	2,3	3,7
5	10	8	3	8,8	1,2	1,2	8,8	2,2	7,8	1,3	8,7	7,5	2,5	1,4	8,6
6	13	11	15	12,1	5,9	6,2	11,8	7,4	10,6	6,6	11,4	10,3	7,7	6,8	11,2
7	9	4	12	4,4	4,6	5,0	4,0	5,1	3,9	5,2 ⁵	3,7 ⁵	3,7	5,3	5,4	3,6
8	16	12	12	13,3	2,7	5,0	11,0	4,5	11,5	5,2 ⁵	10,7 ⁵	11,3	4,7	5,4	10,6
9	11	9	15	9,9	4,1	6,2	7,8	5,3	8,7	6,6	7,4	8,5	5,5	6,8	7,2
10	13	12	3	13,3	-0,3	1,2	11,8	1,5	11,5	1,3	11,7	11,3	1,7	1,4	11,6
11	12	8	9	8,8	3,2	3,7	8,3	4,2	7,8	3,9	8,1	7,5	4,5	4,1	7,9
12	12	12	2	13,3	-1,3	0,8	11,2	0,5	11,5	0,8	11,2	11,3	0,7	0,9	11,1
13	15	12	9	13,3	1,7	3,7	11,3	3,5	11,5	3,9	11,1	11,3	3,7	4,1	10,9
14	11	5	12	5,5	5,5	5,0	6,0	6,1	4,9	5,2 ⁵	5,7 ⁵	4,7	6,3	5,4	5,6
15	16	11	12	12,1	3,9	5,0	11,0	5,4	10,6	5,2 ⁵	10,7 ⁵	10,3	5,7	5,4	10,6
16	15	11	12	12,1	2,9	5,0	10,0	4,4	10,6	5,2 ⁵	9,7 ⁵	10,3	4,7	5,4	9,6
17	10	11	1	12,1	-2,1	0,4	9,6	-0,6	10,6	0,4	9,6	10,3	-0,3	0,5	9,5
18	6	2	9	2,1	3,9	3,7	2,3	4,0	2,0	3,9	2,1	1,8	4,2	4,1	1,9
19	15	11	12	12,1	2,9	5,0	10,0	4,4	10,6	5,2 ⁵	9,7 ⁵	10,3	4,7	5,4	9,6
20	12	8	12	8,8	3,2	5,0	7,0	4,2	7,8	5,2 ⁵	6,7 ⁵	7,5	4,5	5,4	6,6



De invloed van x wordt daartoe afgelezen uit figuur 2a. Deze wordt geëlimineerd door de gevonden waarden van y te verminderen met de waarde van deze invloed. De uitkomsten hiervan zijn vermeld in kolom 5 van tabel 6. Aangezien x_1 en x_2 niet gecorreleerd zijn verondersteld, moet de invloed van x_2 nog² volledig aanwezig zijn in de voor x_1 gecorrigeerde y^a .

Zetten wij dus de resten y^a uit tegen x_2 en construeren de gemiddelde lijn zoals hiervoor uiteengezet. dan geeft deze de invloed op y van x_2 weer (fig. 3a). Deze wordt aangegeven door de waarden $m_2 x_2$, hetgeen in dit geval - zie figuur 3a - $0,41 x_2$ blijkt te zijn.

De op overeenkomstige wijze als voor x_1 , voor x_2 gecorrigeerde waarden van y zijn weergegeven in kolom 7 van tabel 6. De voor x_2 gecorrigeerde waarden y^b van kolom 7 zetten wij nu weer uit tegen x_1 (fig. 3b) en lezen dan een nieuwe waarde voor m_1 af $0,96$. Dit geeft aanleiding tot een nieuwe (verbeterde) correctie van de oorspronkelijke y voor x_1 ,^c die in kolom 8 is vermeld, terwijl de gecorrigeerde waarden y^c in kolom 9 van tabel 6 staan. Deze y^c zetten we weer uit tegen x_2 (fig. 3c) en vinden een $m_2 = 0,44$. De hiermede voor x_2 gecorrigeerde y is als y^d weergegeven in kolom 11. Deze wordt weer uitgezet tegen x_1 (fig. 3d). Zo voortgaande vinden wij op denduur waarden voor m_1 en m_2 en constanten q - van de laatstgenoemde in iedere figuur één - die nauwelijks meer veranderen. Uit tabel 7 blijkt dat dit na de tweede correctie reeds het geval is.

Tabel 7. De grootte van de coëfficiënten m_1 , m_2 en van q in de diverse stadia van de grafische bewerking

Stadium	m_1	m_2	q
Zonder correctie	1,105	0,66	
Na 1e correctie	0,96	0,41	-1,3 -0,1) gem.-0,7
na 2e correctie	0,95	0,44	-0,2 -0,1) gem.-0,2
na 3e correctie	0,96	0,45	-0,2 -0,4) gem.-0,3

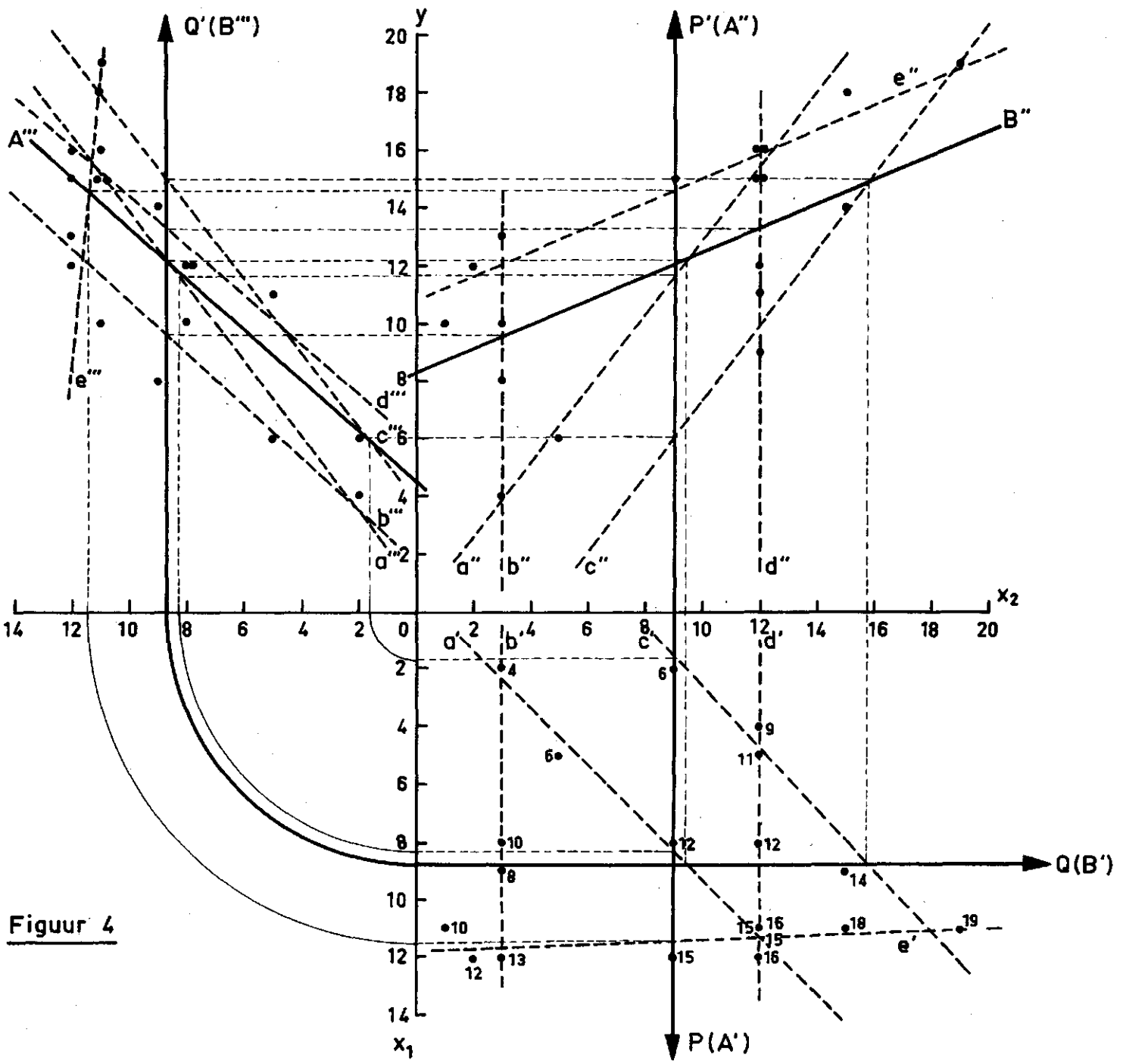
Op grond van de multipele regressieberekening hebben de coëfficiënten de volgende waarden :

$$m_1 = 0,89 \pm 0,06$$

$$m_2 = 0,46 \pm 0,04$$

$$q = 0,2 \pm 0,6$$

waaruit blijkt, dat bij de 2e correctie de grafische manier van bewerken al coëfficiënten van voldoende nauwkeurigheid gaf.



Figuur 4

I. 2.1.2. Drie-dimensionale werkwijze

Terwijl tot nu toe figuren in gebruik waren met twee assen, nl. de x -as en de y -as, heeft de drie-dimensionale figuur drie assen; te weten de x_1 -as, de x_2 -as en de y -as. In het grondvlak van deze figuur worden x_1 en x_2 tegen elkander uitgezet. De stippen in dit grondvlak zijn de projecties van de punten op hoogte y boven het grondvlak. Deze punten worden ook geprojecteerd op de verticaal staande vlakken (y, x_1) en (y, x_2) . Door deze projecties zijn de plaatsen der punten in de ruimtelijke figuur vastgelegd. De punten kan men zich in de ruimte gegroepeerd denken rond een (gemiddeld) vlak, net zoals in de vorige figuren de punten gerangschikt lagen rond een (gemiddelde) lijn. Toen moest een gemiddelde lijn worden geconstrueerd, nu moet een gemiddeld vlak worden bepaald.

De vlakken (x_1, x_2) , (y, x_1) en (y, x_2) van de ruimtelijke figuur worden daartoe heergeslagen in een plat vlak op de wijze als bij de orthogonale projectie gebruikelijk is.

Figuur 4 laat de drie in het platte vlak neergeslagen vlakken zien. Het zou hier te ver voeren de methode van de orthogonale projectie uiteen te zetten. Daarvoor wordt verwezen naar de elementaire leerboeken over beschrijvende meetkunde.

De punten zijn genummerd. Deze nummers zijn alleen bij de horizontale projecties vermeld.

Verschillende groepen van punten in het grondvlak liggen bijna op een rechte lijn. Het is daardoor in dit geval mogelijk een vijftal streeplijnen in dit vlak te trekken. Deze zijn aangegeven met de letters a tot en met e. De streeplijnen stellen de projecties voor van lijnen bij benadering in het gezochte vlak gelegen die bij de overeenkomstige punten behoren en op verschillende hoogten y boven het vlak (x_1, x_2) liggen. Volgens de methode der beschrijvende meetkunde vindt men de projecties van deze lijnen op de vlakken (y, x_1) en (y, x_2) . Ze zijn in deze vlakken met dezelfde geaccentueerde letters aangegeven. De projecties van de lijnen, dus de streeplijnen in de drie vlakken (x_1, x_2) , (y, x_1) en (y, x_2) leggen het gemiddelde vlak bij benadering vast.

De snijlijn van dit vlak met het (y, x_1) vlak of een daaraan evenwijdig vlak levert de waarden op van m_1 en q . Evenzo vindt men m_2 en q , door de constructie van de snijlijn van dat gemiddelde vlak met het (y, x_2) vlak of een daaraan evenwijdig vlak. Om nu de waarden van m_1, m_2 en q te bepalen, zoeken wij de snijlijn van het gemiddelde vlak met twee andere vlakken P en Q, ieder evenwijdig aan één van de verticale projectievlakken en gaande door het gemiddelde punt van de gehele puntenwolk. Zij P het vlak evenwijdig aan het projectievlak (y, x_1) en Q dat evenwijdig aan (y, x_2) .

We snijden nu deze vlakken P en Q met de lijnen a t/m e. De projecties van de snijpunten met P op het (y, x_1) vlak liggen bij benadering op de 3e projectie van de snijlijn A van het gezochte vlak. Wij trekken dus door de projecties van die snijpunten een lijn op de wijze als in figuur 1a is geschied. Door deze projectie A'' en het vlak P is de snijlijn A vastgelegd. Bij deze constructie gebruiken wij de snijpunten van de lijnen a t/m e met de vlakken P en Q en niet die met de projectievlakken - hetgeen

in principe ook mogelijk is - omdat de laatstgenoemde geheel buiten het waarnemingsgebied liggen en de gevonden snijpunten verder van de snijlijn zullen afwijken dan in het midden van de puntenwolk is te verwachten. De vlakken P en Q hadden eveneens als projectievlakken kunnen dienen. Dit zou de figuren echter zeer onoverzichtelijk maken.

Op de bovenomschreven wijze bepalen wij evenzo de snijlijn B van het gezochte vlak en het vlak Q. Hier vinden wij de projectie B'' op het (y, x_2) vlak, welke te zamen met het vlak Q de stand van de lijn in de ruimte vastlegt. De lijnen A en B zullen elkander snijden op de snijlijn van P en Q in het zwaartepunt van de puntenwolk. De hoeken van de 2e en 3e doorgangen van het gemiddelde vlak met respectievelijk de x_2 -as en de x_1 -as geven de coëfficiënten m_2 , m_1 en het snijpunt met de y-as levert de waarde van q. Omdat de 2e en 3e doorgangen evenwijdig zijn met B en A, kunnen m_2 en m_1 onmiddellijk uit de hellingen van B'' en A''' worden afgelezen, men vindt zodoende $m_1 = 0,88$ en $m_2 = 0,42$.

De grootte van q wordt verder berekend door eerst de waarde van y af te lezen op de y-as, waar lijn B'' deze snijdt, y is daar gelijk aan 8,3. Deze hoogte geldt voor $x_1 = \bar{x}$, maar q wordt berekend voor $x_1 = 0$ en $x_2 = 0$. Derhalve leest men met behulp van lijn A''' af hoeveel y is voor $x_1 = 0$.

Men vindt dus : $x_1 = \bar{x}_1 = 8,7$ geeft $y = 12,1$
 en $x_1 = 0$ geeft $y = 4,5$

Verskil 7,6

Het resultaat is dientengevolge : $q = 8,3 - 7,6 = +0,7$

Ter controle is q op dezelfde wijze te berekenen door na te gaan waar lijn A''' de y-as snijdt, men leest af :

$y = 4,5$ geldend voor $x_1 = 0$ en $x_2 = \bar{x}_2$

Om q te vinden moet ook $x_2 = 0$. De aftrekwaarde is af te leiden uit lijn B'';

voor $x_2 = \bar{x}_2 = 9,0$ heeft y een waarde van 12,1

voor $x_2 = 0$ heeft y een waarde van 8,3

Verskil 3,8

Bijgevolg : $q = 4,5 - 3,8 = +0,7$

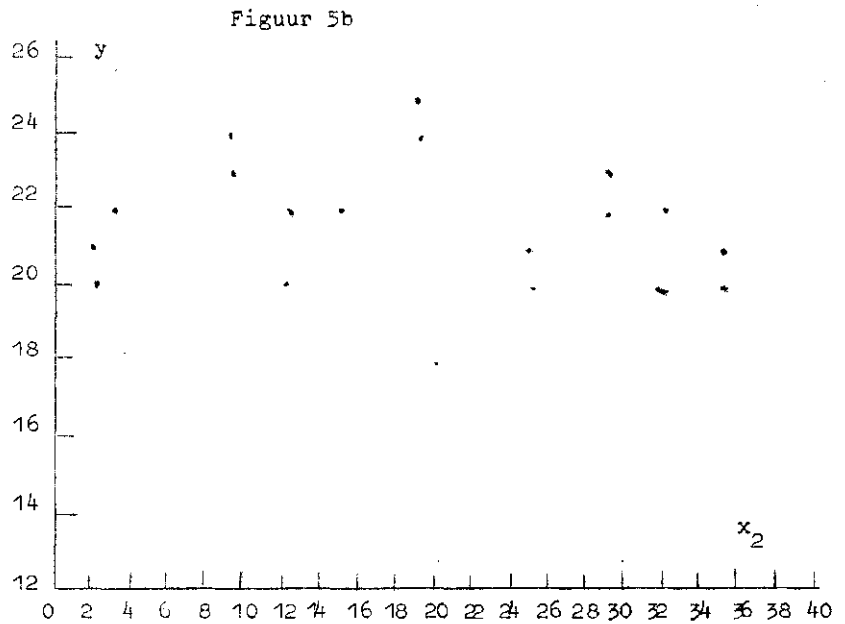
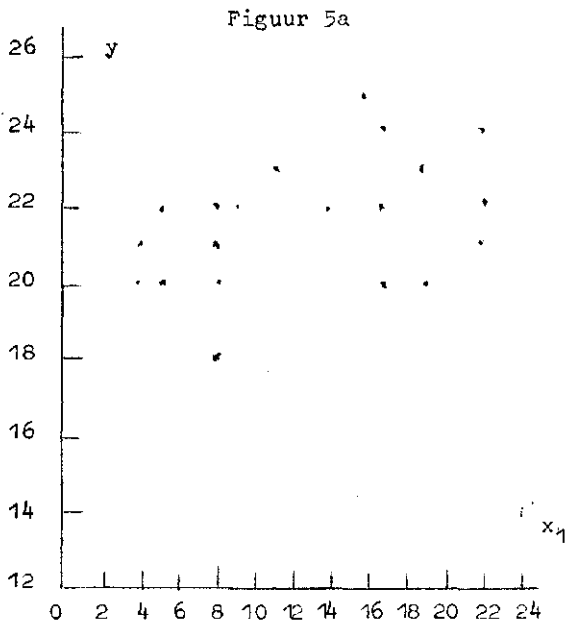
De uitkomsten van deze methode kloppen weer goed met die van de regressieberekening. In tabel 8 zijn de resultaten bij dit voorbeeld verkregen nog eens samengevat.

Tabel 8. De grootte van de coëfficiënten m_1 , m_2 en van q met grafische en rekenkundige bewerking verkregen

Methode	m_1	m_2	q
Grafisch 2-dimensionaal 2e stadium	0,95	0,44	-0,2
3e stadium	0,96	0,45	-0,3
Grafisch orthogonale projectie	0,88	0,42	0,7
Regressiemethode	$0,89 \pm 0,06$	$0,46 \pm 0,04$	$0,2 \pm 0,6$

I.2.2. Correlatie tussen x_1 en x_2

In de praktijk is het dikwijls zo, dat de factoren die de grootte van y bepalen, onderling een correlatie vertonen. De oplossing van de vraagstukken wordt dan moeilijker, omdat de correlatie van de factoren in het "grondvlak", zoals men dat noemt, de werkelijke invloed van elke factor op de grootte van y maskeert. De grafische twee-dimensionale methode is in zo'n geval dikwijls misleidend, want correlatie in het grondvlak laat in de twee-dimensionale figuren doorgaans geen samenhang zien tyssen y en de afzonderlijke factoren van het grondvlak. Als voorbeeld zij gegeven de cijferreeks van tabel 9, die wederom is samengesteld uit de gegevens van tabel 1 en 3. Twee-dimensionaal behandeld ontstaan de figuren 5a, 5b en 5c.



Tabel 9.
Vierde voorbeeld

y	x_1	x_2
22	5	39
20	5	35
22	9	32
20	4	32
21	4	35
24	22	9
20	19	2
21	22	2
23	19	9
22	22	3
22	8	29
20	8	25
23	11	29
21	8	25
18	8	20
24	17	19
20	17	12
22	17	12
22	14	15
25	16	19

Volgens figuur 5a bestaat er vrijwel geen invloed van x_1 op y en figuur 5b zou doen vermoeden, dat x_2 helemaal geen invloed heeft op y . Figuur 5c toont echter de vrij nauwe samenhang tussen x_1 en x_2 .

Voor de grafische methode van bewerken is de manier met orthogonale projectie in dit geval beter.

Figuur 6 geeft deze wijze van bewerken weer. Ter oriëntering aan de eerste orthogonale bewerking (fig.4) zijn de symbolen, die de snijlijnen, waar het om gaat, aanduiden, gelijk gehouden.

Lijn A''' doet zien, dat er wel degelijk een invloed van x_1 op y is en lijn B''' geeft de betrekking tussen x_2 en y . De met deze bewerking gevonden uitkomsten voor de formule:

$$y = m_1 x_1 + m_2 x_2 + q \text{ zijn :}$$

$$m_1 = 0,63$$

$$m_2 = 0,30$$

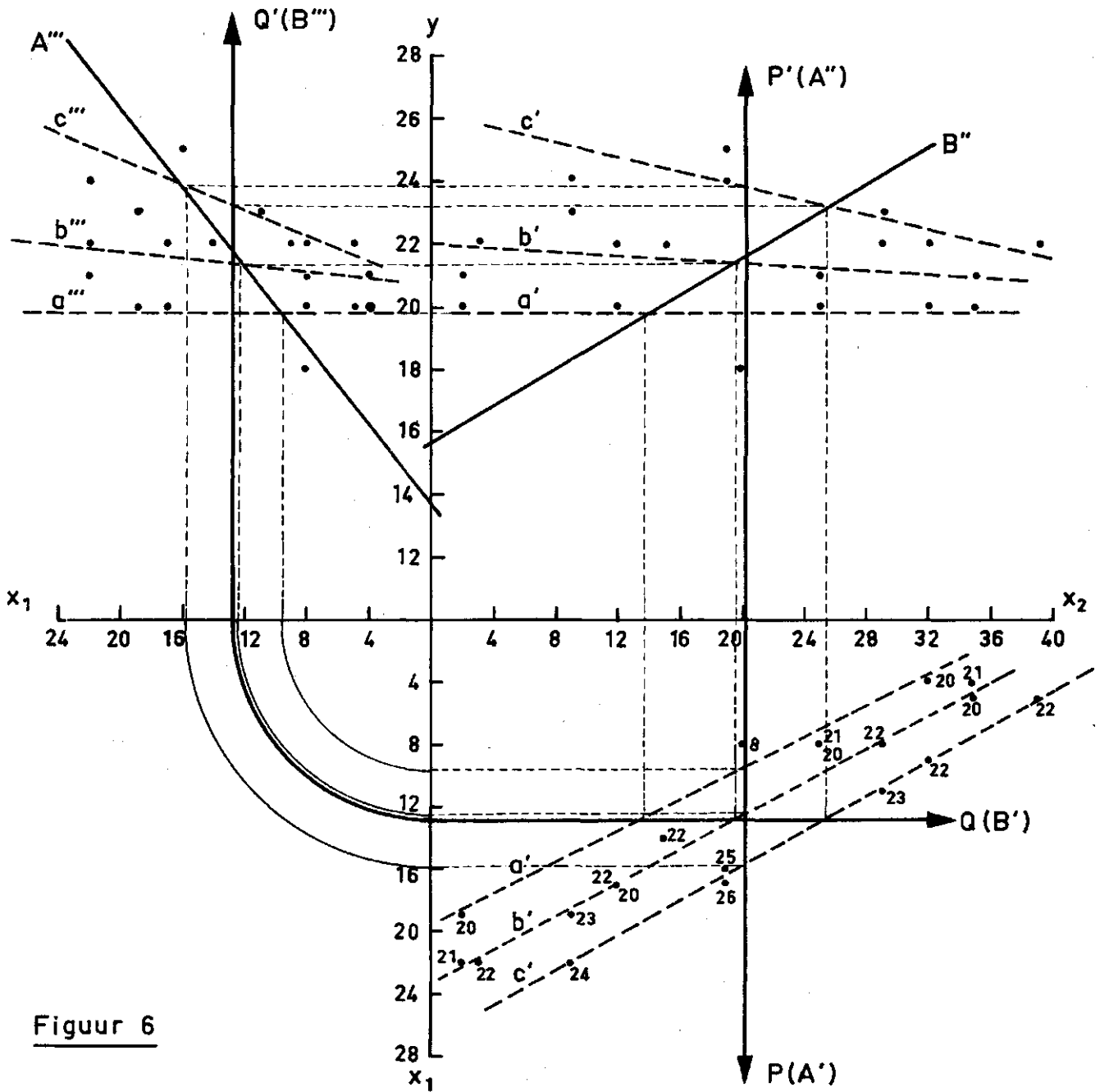
$$q = 7,7$$

De regressieberekening levert als resultaat op:

$$m_1 = 0,63 \pm 0,09$$

$$m_2 = 0,30 \pm 0,05$$

$$q = 7,5 \pm 2,0$$



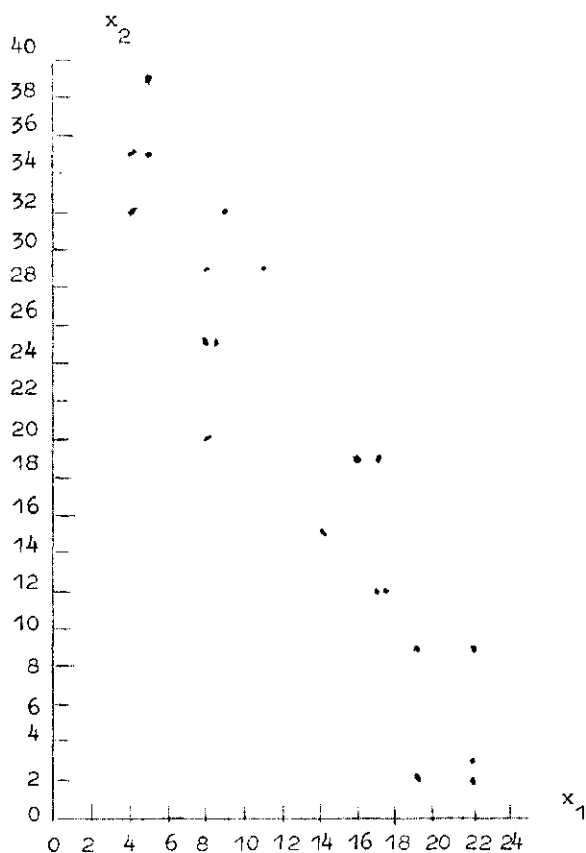
Figur 6

De overeenstemming der uitkomsten van beide bewerkingsmethoden is weer zeer goed. Opgemerkt zij, dat de overeenstemming alleen geldt voor de gemiddelde waarden van x_1 en x_2 . De meetkundige figuur 5c geeft namelijk aanwijzingen, dat er een interactie aanwezig is tussen x_1 en x_2 , wat de waarde van y betreft.

In figuur 6 is dit af te leiden uit de projectie der lijnen in de vlakken (y, x_1) en (y, x_2) . Het aantal waarnemingen van de serie is echter te beperkt om aan de interactie meer aandacht te besteden.

Uit deze aanwijzing blijkt evenwel, dat de grafische methode de mogelijkheid tot informatie verschaft, die van belang kunnen zijn voor de verdere rekenkundige bewerking van het cijfermateriaal.

Figuur 5c



II. DE GEGEVENS DER KOSTPRIJSBEDRIJVEN

Na de vier voorbeelden, die ter toelichting van de grafische manier van werken dienen, passen wij deze thans toe op de cijferreeksen van een aantal kostprijsbedrijven. Het betreft bedrijven, voorkomend in Overzicht no. 231 van het Landbouw Economisch Instituut. Dit overzicht geeft cijfers van 21 zuivere weidebedrijven over het boekjaar 1949/1950. De grafische bewerking bracht naar voren, dat de winst in guldens per ha lineair samenhang met de volgende factoren :

- | | |
|---|-------------|
| 1. Stuks grootvee per 100 ha | symbool = G |
| 2. Geldopbrengst per stuk grootvee | " = O |
| 3. Voeraankoop in guldens per stuk grootvee | " = V |
| 4. Loon per ha | " = L |

De algemene formule voor de relatie tussen de cijferreeksen van dit overzicht is dus :

$$\text{Winst} = m_1 G + m_2 O + m_3 L + m_4 V + q$$

De meetkundige aanpak der reeksen gebeurde met orthogonale projectie. De winst werd uitgezet tegen G en L. Dit gaf een eerste oriëntering omtrent de invloed van G en L afzonderlijk op de winst. De invloed van G en L is toen aan de hand van de gevonden lijnen globaal geëlimineerd op overeenkomstige wijze als in I.2.1. (tabel 6). Alleen zijn hierbij de invloeden van G en L tegelijk geëlimineerd, zodat we het verband tussen de gecorrigeerde W en O en V volgens de grafische methode van I.2.1.2. (blz.12) kon worden bepaald. De gecorrigeerde winstcijfers zijn daartoe uitgezet tegen O en V; waardoor de betrekking tussen de winst en dit paar factoren werd verkregen. De invloed van O en V is vervolgens globaal geëlimineerd op de oorspronkelijke winst om nauwkeuriger de relatie tussen de winst en G en L te weten te komen, enzovoort.

Er is dus op analoge manier gewerkt als beschreven onder I.2.1.1., alleen met dit verschil, dat men nu een drie-assenstelsel in plaats van een twee-assenstelsel heeft gebruikt. Ook werden nu steeds twee factoren gelijktijdig geëlimineerd, terwijl dit bij het genoemde voorbeeld van I.2.1.1. slechts voor één factor werd gedaan.

De coëfficiënten gevonden bij de eerste bewerking met nog niet gecorrigeerde winstcijfers en de factoren G en L en die verkregen bij de eerste bewerking met gecorrigeerde cijfers, maar dan met de factoren O en V, zijn gebruikt voor het opstellen van de eerste benaderende vergelijking. De tweede vergelijking werd samengesteld met de coëfficiënten opgebouwd door de volgende twee bewerkingen, enz. In totaal zijn 7 opeenvolgende formules bepaald. Eigenlijk had de eerste vergelijking uit de bewerking van beide paren factoren met ongecorrigeerde winstcijfers moeten worden opgesteld, maar dat is iets minder doelmatig.

De op de omschreven wijze gevonden uitkomsten zijn vermeld in tabel 10.

Tabel 10. De grootte van de coëfficiënten m_1 , m_2 , m_3 en m_4 benevens de waarde van q der vergelijking $W' = m_1 G + m_2 O + m_3 L + m_4 V + q$, verkregen door grafische bewerking

Vergelijkingen	m_1	m_2	m_3	m_4	q
1e vergelijking	5,92	0,65	-0,44	-2,05	-535
2e "	5,64	1,02	-0,90	-1,59	-772
3e "	4,70	1,12	-0,99	-1,75	-669
4e "	4,49	1,22	-1,21	-1,87	-636
5e "	4,67	1,42	-1,25	-1,95	-810
6e "	4,24	1,47	-1,29	-2,20	-732
7e "	4,64	1,27	-1,15	-1,95	-704

De vergelijkingen zijn gebruikt om de winsten van de bedrijven te berekenen. Vervolgens is nagegaan, hoe hoog de correlatie was tussen de aldus berekende winsten en de werkelijke winsten. De voor de 7 benaderingen gevonden correlatiecoëfficiënten zijn in tabel 11 medegedeeld.

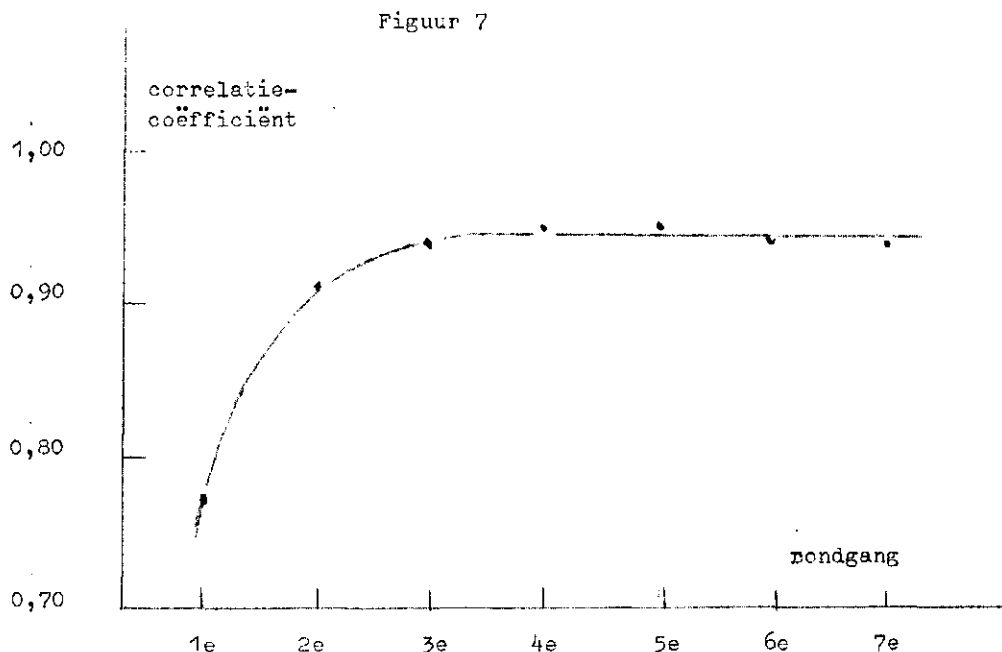
Tabel 11. Correlatiecoëfficiënten tussen werkelijke winst en berekende winst

Rangorde van de vergelijking	1	2	3	4	5	6	7
Correlatiecoëfficiënt	0,77	0,91	0,94	0,95	0,95	0,94	0,94

Bovendien is in figuur 7 het in tabel 11 gegeven verloop nog eens grafisch weergegeven. Duidelijk ziet men, dat na de 3e vergelijking geen verbetering in de correlatie meer voorkomt.

De regressieberekening toegepast op de cijfers van de kostprijsbedrijven, leverde als coëfficiënt op :

$$\begin{aligned}
 m_1 &= 4,64 \pm 0,75 \\
 m_2 &= 1,36 \pm 0,20 \\
 m_3 &= -1,34 \pm 0,11 \\
 m_4 &= -1,96 \pm 0,20 \\
 q &= -729 \pm 132
 \end{aligned}$$



Opnieuw blijkt een goede overeenstemming in uitkomsten der beide methoden.

De vergelijking samen te stellen met de bij de regressieberekening gevonden coëfficiënten, levert berekende winsten, die een correlatiecoëfficiënt van 0,94 met de werkelijke winst hebben net als bij de meetkundige methode.

De grafische bewerking van cijfermateriaal heeft dus in alle behandelde gevallen dezelfde resultaten geboekt als de regressieberekening. Het voorbeeld onder I.2.2. liet zien, dat de meetkundige methode bovendien inlichtingen verschaft omtrent de aard van de wederzijdse betrekkingen der onderzochte grootheden. De meetkundige manier kan bij voorbeeld opheldering geven of interactie aanwezig is, zoals in het genoemde voorbeeld was te zien. Tevens ziet men of de betrekkingen rechtlijnig dan wel kromlijnig zijn. Het karakter van de relaties is dus uit de figuren af te lezen. Dit karakter bepaalt nu welke vergelijking bij de rekenkundige behandelingswijze moet worden gebruikt. De rekenkunde vormt dan de afsluiting van de bewerking van het cijfermateriaal en de meetkunde geeft vooraf inzicht in het karakter van het materiaal, waarmee bij de berekeningen rekening kan worden gehouden. In de literatuur vindt men hier voorbeelden van in van Uven : *Mathematical Treatment* 2e druk blz. 135 t/m 143 en ook Tinbergen noemt in *Econometrie* blz. 118 deze wijze van werken.

Het voordeel van de vergelijking is verder hierin gelegen, dat het vraagstuk toegankelijk geworden is voor eventuele nadere analytische bewerkingen.

Uit het voorgaande blijkt dat de grafische bewerking en rekenkundige methode elkaar aanvullen.

SAMENVATTING

Uit een viertal voorbeelden, die in hoofdstuk I ter demonstratie van de methoden uitvoerig zijn besproken, blijkt dat er een goede overeenstemming is te bereiken tussen de uitkomsten van de grafische en de rekenkundige methode van bewerken. Bovendien gaven beide methoden dezelfde uitkomst bij de bewerking van gegevens uit de boekhoudingen van 21 zuivere weidebedrijven.

De grafische methoden geven dikwijls op eenvoudige wijze een goed inzicht in de samenhang tussen de statistische gegevens. Ook zijn in eenvoudige gevallen op snelle wijze uitkomsten te verkrijgen, die meestentijds voldoende nauwkeurig zijn. In gevallen, waarbij een groot aantal factoren een rol spelen, kan het benodigde tekenwerk zeer omvangrijk worden, zoals wel uit de gegeven drie-dimensionale voorbeelden blijkt en het in hoofdstuk II behandelde geval gaf eveneens aanleiding tot veel tekenwerk. Indien het tekenwerk te groot wordt, kunnen evenwel enkele eenvoudige twee-dimensionale figuren dikwijls verhoeden, dat onderling sterk gecorreleerde factoren in een regressieberekening worden opgenomen. De grafische en rekenkundige methode vullen elkander aan. Daarnaast kan de rekenkundige bewerking zo nodig een veel grotere nauwkeurigheid bereiken en een in getallen uitgedrukte nauwkeurigheid der uitkomsten aangeven. Verder is daarbij het mogelijke aantal factoren in principe onbeperkt.

S 926

400 ex

Ko/LH

17-2-1960