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 215

Statistische analyse van experimenten, uitgevoerd met betrekking tot de zuivering van enkele sera.

door

Ir A.R. Bloemena

April 1957

Inhoud

blz.

1. Inleiding

1.1.

1.2. De onderzochte factoren 1

1.3. De waargenomen grootheden 2

2. De gevolgde werkwijze

2.0. Inleiding 2

2.1. Het mathematische model 2

2.2. Eigenschappen van kleinste kwadratenschattingen 6

2.3. Extrapolatie naar punten met hogere responsie 7

3. Een roteerbaar proefschema 7

4. Zuivering van schapenantidifterieserum

4.1. Eerste experiment 13

4.2. Het tweede experiment 21

4.3. De waarnemingen verricht in O,...,O 23

4.4. De invloed van de factoren op de O en de RF 26

4.5. Het aangepaste oppervlak 27

4.6. De mate van aanpassing 28

4.7. Extrapolatie naar punten met hogere RF 30

4.8. Het aantal herhalingen van de controleproeven 31

4.9. Gewenste nauwkeurigheid, waarmee de factoren dienen te worden aangehouden 32

4.10. De controle proeven 33

4.11. De ammoniumsulfaatconcentratie bij de tweede uitzouting 34

4.12. Conclusies 34

5. Zuivering van paardenantidifterieserum

5.1. Inleiding 35

5.2. De proeven verricht in O,...,O 36

5.3. De invloed van de factoren 37

5.4. Het aangepaste oppervlak 37

5.5. Extrapolatie naar punten met hogere RF 39

5.6. De proeven, uitgevoerd ter bevestiging van de conclusies van 5.5 41

5.7. Conclusies 44

6. Zuivering van paardenantitetanusserum

6.1. Inleiding 45

6.2. Waarnemingen verricht in O,...,O 45

6.3. Het aangepaste oppervlak 47

6.4. Extrapolatie naar punten met hogere O 49

6.5. Invloed van de dagverschillen 52

6.6. Proeven, uitgevoerd ter bevestiging van de conclusies van 6.4 en 6.5 53

6.7. Conclusies 54

1. Inleiding

1.1.

Bij de voorbereiding van een onderzoek betreffende de experimentele bepaling van de optimale condities voor de zuivering van schapenantidifterieserum werd het door het Rijksinstituut voor de Volksgezondheid, alwaar dit onderzoek zou worden verricht, wenselijk geacht zich te laten adviseren over de wijze, waarop men de experimenten zo goed mogelijk kon opzetten. Hiervoor wendde men zich tot T.N.O.-Afdeling Bewerking Waarnemingsuitkomsten, welk instituut op haar beurt zich tot het Mathematisch Centrum wendde. Op deze wijze werd door het Mathematisch Centrum een aantal adviezen verstrekt, terwijl de waarnemingen die verkregen werden bij de volgens deze adviezen uitgevoerde experimenten voor een aanzienlijk deel op het Mathematisch Centrum werden bewerkt. Tijdens het onderzoek betreffende schapenantidifterieserum werd het wenselijk geacht ook enkele reeds in productie zijnde sera, t.w. paardenantitetanus- en paardenantidifterieserum te onderzoeken, teneinde na te gaan in hoeverre de bij de huidige productie toegepaste condities als optimaal zijn te beschouwen.

Alle stadia van het door het Mathematisch Centrum verrichte deel van het onderzoek zijn beschreven in korte rapporten, o.a. [1], [2], [3] en [4]. In dit rapport zal een algeheel overzicht worden gegeven van de opzet van de experimenten en de verwerking van de waarnemingsresultaten.

Bij de berekeningen is op uitgebreide schaal gebruik gemaakt van electronische rekenapparatuur. Zo zijn de matrices A, B en C (zie (4.5;4), (4.4;5) en (6.3;4)), de translaties (4.2;4), (5.4;3) en de schattingen in Tabel II alle door de ARMAC van het Mathematisch Centrum berekend. De programma's hiervoor zijn gemaakt door A. BENARD, E.W. DIJKSTRA en Mevrouw R. ZONNEVELD-MULDER.

1.2. De onderzochte factoren

Op verzoek van het R.I.V. werden de volgende factoren in het onderzoek betrokken:

a) Stadium van de enzyminwerking:

1	zuurgraad	z_1	
2	temperatuur	T_1	(^{o}C)
3	tijdsduur	t_1	(minuten)
4	hoeveelheid pepsine	p	(mg)

b) Stadium van eerste uitzouting en de thermocoagulatie:

5	hoeveelheid $(NH_4)_2SO_4$	z_1	(%)
6	zuurgraad	z_2	
7	temperatuur	T_2	(°C)
8	tijdsduur	t_2	(minuten).

In dit rapport zullen wij de factoren aangeven met $z_1, T_1, \ldots, t_2$ maar ook indien geen verwarring dreigt met 1,...,8.

1.3. De waargenomen grootheden

Bij elke der uitgevoerde proeven werd een waarneming verricht van de opbrengst aan gezuiverd serum (aangegeven door O) en een reinigingsfactor (aangegeven door RF). Bij het opsporen van optima zal men in de regel de keus hebben:

1) de RF zo groot mogelijk te maken, onder de voorwaarde dat O een bepaalde waarde heeft, ofwel
2) de O zo groot mogelijk te maken, onder de voorwaarde dat de RF een bepaalde waarde heeft.

Bij elke type serum zal aangegeven worden welke van deze beide mogelijkheden gevolgd is.

2. De gevolgde werkwijze

2.0.

Het onderzoek werd geleid volgens een experimenteertechniek, die vooral door BOX in de literatuur is beschreven (zie [6] - [11]). Een overzicht over deze werkwijze kan ook worden gevonden in [5] , Enkele belangrijke punten zullen hier in beknopte vorm aangehaald worden.

2.1. Het mathematische model

Bij een experiment, waarbij k continu te variëren factoren $x_1, \ldots, x_k$ zijn betrokken, meet men een responsie $\underline{y}$, die men zich voorstelt als een functie van deze k factoren en een responsiefluctuatie onder constante experimentele condities, $\underline{e}$:

$$\underline{y} = \Phi(x_1, \ldots, x_k) + \underline{e} \qquad 1) \qquad (2.1;1)$$

---

1) Onderstreepte grootheden zijn stochastisch, d.w.z. zij bezitten een kansverdeling. Een waargenomen waarde van een dergelijke grootheid wordt weergegeven door hetzelfde - niet onderstreepte - symbool.

Stel dat het doel van het onderzoek is het vinden van die combinatie van waarden van de factoren, $x_1^{(m)}, \ldots, x_k^{(m)}$, waarvoor $\Phi(x_1, \ldots, x_k)$ maximaal is.

Betreffende de waarnemingsfluctuatie $\underline{e}$ maken wij de volgende onderstellingen:

1. $E\,\underline{e} = o$.
2. de verdeling van $\underline{e}$ is dezelfde voor alle combinaties van waarden van de k factoren $(var\{\underline{e}\} = \sigma^2)$.

Het stochastische karakter van de $\underline{e}$ komt bijvoorbeeld tot uiting in de verschillen die optreden tussen duplowaarnemingen. Vandaar dat in dit rapport de term "waarnemingsonnauwkeurigheid" ook gebruikt wordt om de $\underline{e}$ aan te duiden.

Bij een experiment worden nu N proeven uitgevoerd, en dus N metingen van de responsie verkregen. De u-de proef van het experiment wordt zo uitgevoerd, dat daarbij de k factoren de waarden $x_{1u}, x_{2u}, \ldots, x_{ku}$ aannemen. Men kan op de volgende wijze elke proef opvatten als een punt in de k-dimensionale ruimte. Elke coördinaatas van een k-dimensionaal assenstelsel komt overeen met een van de factoren. De u-de proef kan men dan voorstellen als het punt met coördinaten $x_{1u}, x_{2u}, \ldots, x_{ku}$. In het volgende zal herhaaldelijk van deze wijze van afbeelden gebruik gemaakt worden.

Is de bij de u-de proef waargenomen responsie y_u, dan geldt dus:

$$E\,\underline{y}_u = \Phi(x_{1u}, \ldots, x_{ku}). \qquad (2.1;2)$$

In aansluiting op de hierboven gemaakte onderstellingen, zal een derde onderstelling gemaakt worden:

3. de waarnemingen zijn onderling onafhankelijk.

De twee laatste onderstellingen kunnen in matrixnotatie geschreven worden als

$$E(\underline{y} - E\,\underline{y})(\underline{y} - E\,\underline{y}) = \sigma^2 I \qquad (2.1;3)$$

Nemen wij aan, dat de functie Φ in het beschouwde deel van de k-dimensionale ruimte continue afgeleiden bezit, dan kan (2.1;2) worden ontwikkeld in een reeks van TAYLOR. Hiertoe voeren wij de volgende notatie in:

voor de waarde van de functie Φ in de oorsprong 0:

$$[\Phi]_{x_1=\cdots=x_k=0} = \varphi_0 ,$$

voor de partieel afgeleiden naar x_i in 0:

$$\left[\frac{\partial \Phi}{\partial x_i}\right]_{x_1=\cdots=x_k=0} = \varphi_i , \qquad (2.1;4)$$

voor de tweede partieel afgeleide in 0:

$$\left[\frac{\partial^2 \Phi}{\partial x_i \partial x_j}\right]_{x_1=\cdots=x_k=0} = \varphi_{ij} .$$

Voorts:

$$x_{0u} \overset{def}{\equiv} +1 \qquad (2.1;5)$$

Nu kan, onder de genoemde voorwaarden, (2.1;2) worden geschreven als:

$$\begin{aligned} E(\underline{y}_u) = {} & \varphi_0 x_{0u} + \varphi_1 x_{1u} + \cdots + \varphi_k x_{ku} + \tfrac{1}{2}\varphi_{11} x_{1u}^2 + \cdots + \tfrac{1}{2}\varphi_{kk} x_{ku}^2 + \\ & + \varphi_{12} x_{1u} x_{2u} + \cdots + \varphi_{k-1,k} x_{k-1,u} x_{ku} + \cdots + \\ & + \tfrac{1}{6}\varphi_{111} x_{2u}^3 + \cdots + \tfrac{1}{2}\varphi_{122} x_{1u} x_{2u}^2 + \cdots \end{aligned} \qquad (2.1;6)$$

Teneinde bij de toepassing vergissingen ten gevolge van de coëfficiënten $\frac{1}{2}$, $\frac{1}{6}$, etc. uit te sluiten definiëren wij

$$\left.\begin{array}{ll} \varphi_i = \beta_i & \tfrac{1}{6}\varphi_{iii} = \beta_{iii} \\ \tfrac{1}{2}\varphi_{ii} = \beta_{ii} & \tfrac{1}{2}\varphi_{iij} = \beta_{iij} \\ \varphi_{ij} = \beta_{ij} & \text{etc.} \end{array}\right\} \qquad (2.1;7)$$

Hierbij noemen wij

β_i = het i-de lineaire effect,

β_{ii} = het i-de quadratische effect en

β_{ij} = het (i,j) interactie-effect.

Op grond van bepaalde onderstellingen van de functie Φ kan men veelal aannemen, dat termen van de graad d en hoger in x verwaarloosbaar klein zijn. Zo zal, indien Φ plaatselijk goed te benaderen is door een plat vlak gelden:

$$E\,\underline{y}_u = \beta_0 x_{0u} + \beta_1 x_{1u} + \cdots + \beta_k x_{ku} . \qquad (2.1;8)$$

Men noemt dit het <u>model</u>, en wel het lineaire model.

Is Φ plaatselijk bij benadering een tweede-graads-oppervlak, dan komt men tot het kwadratische model:

$$\begin{aligned} E\,\underline{y}_u = {} & \beta_0 x_{0u} + \beta_1 x_{1u} + \cdots + \beta_k x_{ku} + \beta_{11} x_{1u}^2 + \cdots + \beta_{kk} x_{ku}^2 + \\ & + \beta_{12} x_{1u} x_{2u} + \cdots + \beta_{k-1,k} x_{k-1,u} x_{ku} . \end{aligned} \qquad (2.1;9)$$

Deze modellen stellen alleen exact het oppervlak voor, als Φ een plat vlak, resp. een tweede-graads-oppervlak is. Evenwel zullen, indien Φ locaal slechts weinig van een plat vlak, resp. tweede-graads-oppervlak afwijkt, (2.1;8) en (2.1;9) redelijke benaderingen zijn.

Uitgaande van een bepaald model zullen uit de waargenomen responsies $y_u (u = 1, \ldots, N)$ schattingen van de β_0 , β_i , en eventueel β_{ii} worden berekend. Hierdoor wordt een indruk verkregen van het gedrag van de functie Φ in het voor ons interessante gebied.

Indien gebruik wordt gemaakt van matrixnotatie kan zowel (2.1;8) als (2.1;9) worden geschreven als

$$E\underline{y} = X\beta \qquad (2.1;10)$$

waarbij X de matrix der x_{iu} (u = rijnummer, i = kolomnummer) en $\underline{y}$, resp, β de kolomvector der $\underline{y}_u$ resp. van de β_i voorstelt. Immers (2.1;8) kan geschreven worden als:

$$E\begin{pmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_N \end{pmatrix} = \begin{array}{c} \begin{array}{cccc} (0) & (1) & & (k) \end{array} \\ \begin{pmatrix} x_{01} & x_{11} & \cdots & x_{k1} \\ \vdots & & & \vdots \\ x_{0N} & x_{1N} & \cdots & x_{kN} \end{pmatrix} \end{array} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} . \qquad (2.1;11)$$

Op analoge wijze kan (2.1;9) geschreven worden als:

$$E\begin{pmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_N \end{pmatrix} = \begin{array}{c} \begin{array}{ccccccccc} (0) & (1) & & (k) & (11) & & (kk) & (12) & & (k-1,k) \end{array} \\ \begin{pmatrix} x_{01} & x_{11} & \cdots & x_{k1} & x_{11}^2 & \cdots & x_{k1}^2 & (x_{11}x_{21}) & \cdots & (x_{k-1,1}x_{k1}) \\ \vdots & \vdots & & & & & & & & \\ x_{0N} & x_{1N} & \cdots & x_{kN} & x_{1N}^2 & \cdots & x_{kN}^2 & (x_{1N}x_{2N}) & \cdots & (x_{k-1,N}x_{kN}) \end{pmatrix} \end{array} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \\ \beta_{11} \\ \vdots \\ \beta_{kk} \\ \beta_{12} \\ \vdots \\ \beta_{k-1,k} \end{pmatrix} . \qquad (2.1;12)$$

2.2. Eigenschappen van kleinste kwadraten-schattingen

Wij nemen aan dat een bepaald model gekozen is en zullen nu aangeven, hoe schattingen van de β 's kunnen worden berekend uit de vector y van de waarnemingsresultaten.

De volgende eisen kunnen worden gesteld aan de schattingen van de β 's:

a. zij moeten zuiver zijn

en

b. van alle mogelijke zuivere schattingen de kleinste variantie hebben.

De methode der kleinste kwadraten geeft hiervoor de oplossing die vooral in verband met (2.1;3) eenvoudig van vorm is (zie [15]). De volgende resultaten worden hierbij verkregen (uitgaande van (2.1;10)).

a. voor de schattingen b van β

$$b = (X'X)^{-1} X'y \,, \qquad (2.2;1)$$

b. voor de covariantie-matrix van deze schattingen:

$$E\,(\underline{b}-\beta)(\underline{b}-\beta)' = (X'X)^{-1}\sigma^2 \,, \qquad (2.2;2)$$

c. voor een zuivere schatting van σ^2:

$$s^2 = (N-L)^{-1}(y'y - b'X'Xb) \,, \qquad (2.2;3)$$

waarin N = het aantal experimenten
L = het aantal elementen van de vector β .

Er zijn evenwel twee voorwaarden, waaraan voldaan moet zijn:

<u>a</u>. $N \geq L$, dus het aantal experimenten moet minstens even groot zijn als het aantal te schatten coëfficiënten,

en

<u>b</u>. $(X'X)$ mag niet singulier zijn (d.i. determinant 0 hebben). Dan bestaat dus de inverse matrix $(X'X)^{-1}$.

Een belangrijk feit is dat als een bepaalde proefopzet uitgevoerd zal worden, de matrices $(X'X)$, $(X'X)^{-1}$ en $(X'X)^{-1}X'$ reeds van tevoren kunnen worden uitgerekend. Wij weten dan, afgezien van de σ^2 , reeds hoe groot de varianties en covarianties van de b 's zullen worden.

2.3. Extrapolatie naar punten met hogere responsie

De berekende schattingen $\mathcal{B}$ kunnen nu worden ingevuld in de modelvergelijking (2.1;8), resp. (2.1;9). Bij een lineair model kan hieruit dan een richting bepaald worden, waarin de toename van de verwachte responsie maximaal is (de methode van de steilste weg). Bij een kwadratisch model is het mogelijk de coördinaten te bepalen van het extreem van het oppervlak, en na te gaan of dit een maximum, minimum, dan wel een zadelpunt is. Vaak zal evenwel blijken, dat dit extreem zo ver verwijderd is van het deel van de k -dimensionale ruimte, waar de proeven zijn uitgevoerd, dat aan het extreem geen physische betekenis kan worden toegekend. Men kan dan echter uit de tweedegraadsvergelijking de coördinaten vinden van het punt, waarvan men kan verwachten dat de responsie een bepaalde (gunstige) waarde aanneemt en welk punt zo dicht mogelijk bij het centrum van de uitgevoerde proefopzet gelegen is. In de volgende paragrafen zal dit nog herhaaldelijk ter sprake komen.

3. Een roteerbaar proefschema

Bij ieder van de drie onderzochte serazuiveringsprocessen is een proefschema toegepast, waarvan de bouw en de opzet terwille van de overzichtelijkheid in deze paragraaf zal worden behandeld.

Het proefschema werd steeds uitgevoerd met het doel schattingen te verkrijgen van de coëfficiënten $\beta_0, \beta_1, \ldots, \beta_8, \beta_{11}, \ldots, \beta_{88}, \beta_{12}, \ldots, \beta_{78}$ van (2.1;9). Het aantal te schatten coëfficiënten is $\binom{8+2}{8} = 45$, zodat (zie 2.2;3) het aantal uit te voeren proeven minstens 45 moet bedragen. Het roteerbare schema (op de naam en de verklaring hiervan komen wij nader terug), dat een zo klein mogelijk aantal proeven omvat, en nog aan de genoemde eis voldoet, bestaat uit:

a) 64 proeven volgens een $\frac{1}{4}$ deel van een 2^8 factorieel schema,

b) 16 proeven volgens een oktaëder in acht dimensies om hetzelfde centrum, en

c) 20 proeven in dat centrum.

Voor de partiële factoriële proefschema's en hun eigenschappen verwijzen wij naar [8] en [13].

Men heeft ten aanzien van de bouw van het roteerbare schema nog de keuze over:

1) welk $\frac{1}{4}$ deel van het 2^8 factoriële schema men gebruiken wil; en

2) de verhouding van afstand van de punten van de oktaëder tot het centrum, en die van het factoriële schema tot het centrum.

Wij zullen deze keuze niet in detail motiveren (zie hiervoor [12]), maar slechts aangeven dat met de door ons gekozen waarden het proefschema een aantal gunstige eigenschappen krijgt.

Als kenmerkende relatie van het partiële factoriële schema is gekozen:

$$\text{eenheid} = 11 = \ldots = 88 = 13578 = 12346 = 245678, \qquad (3;1)$$

terwijl voor de onder 2) aangegeven verhoudingen de waarde 1 werd genomen.

De proefopzet in de gestandaardiseerde waarden van de niveaux van de factoren wordt dan als aangegeven in Tabel I, linker gedeelte. De waarde $\sqrt{8}$ volgt daarbij uit:

$$1 = \frac{\text{afstand punten oktaëder tot centrum}}{\text{afstand punten van het factoriële deel tot centrum}} = \frac{\text{afstand punten oktaëder tot centrum}}{\sqrt{8}}.$$

Wij zullen nu nagaan welke effecten geschat kunnen worden, indien het responsieoppervlak ter plaatse exact een derdegraadsoppervlak is met als vergelijking:

$$\begin{aligned} E\,y = {} & \beta_0 + \beta_1 x_1 + \cdots + \beta_8 x_8 + \beta_{11} x_1^2 + \cdots + \beta_{88} x_8^2 + \beta_{12} x_1 x_2 + \cdots + \beta_{78} x_7 x_8 + \\ & + \beta_{111} x_1^3 + \cdots + \beta_{888} x_8^3 + \beta_{112} x_1^2 x_2 + \cdots + \beta_{788} x_7 x_8^2 + \\ & + \beta_{123} x_1 x_2 x_3 + \cdots + \beta_{678} x_6 x_7 x_8 . \qquad (3;2) \end{aligned}$$

Uit (3;1) volgt dat voor elke proef van het beschouwde schema geldt, dat bijv.

$$x_1 x_2 = x_3 x_4 x_6 ,$$

Tabel I. Het roteerbare proefschema en de waarnemingen gedaan bij de uitvoering hiervan t.a.v. drie sera

	factor	1	2	3	4	5	6	7	8	D_1		D_2		D_3	
no proef	no R.I.V.	z_1	T_1	t_1	p	a	z_2	T_2	t_2	O	R.F.	O	R.F.	O	R.F.
1	8	+1	+1	+1	+1	+1	+1	+1	+1	22,1	1,28	50,5	1,99	58,1	1,48
2	55	+1	+1	+1	+1	+1	+1	-1	-1	37,4	1,26	54,6	1,72	47,7	1,96
3	63	+1	+1	+1	+1	-1	+1	+1	-1	37,4	1,26	58,1	2,06	53,0	1,74
4	35	+1	+1	+1	+1	-1	+1	-1	+1	57,3	1,39	50,5	1,62	59,3	1,39
5	18	+1	+1	+1	-1	+1	-1	+1	+1	13,3	1,30	42,5	2,56	44,4	2,31
6	49	+1	+1	+1	-1	+1	-1	-1	-1	25,0	1,52	44,4	2,15	49,5	2,11
7	28	+1	+1	+1	-1	-1	-1	+1	-1	21,7	1,07	48,3	2,04	47,7	2,08
8	15	+1	+1	+1	-1	-1	-1	-1	+1	35,4	1,22	54,8	1,95	53,3	1,72
9	21	+1	+1	-1	+1	+1	-1	+1	-1	13,3	0,94	44,4	2,53	49,9	2,49
10	12	+1	+1	-1	+1	+1	-1	-1	+1	45,8	1,36	44,4	2,05	47,7	1,97
11	36	+1	+1	-1	+1	-1	-1	+1	+1	26,3	1,15	48,3	2,30	46,2	2,22
12	58	+1	+1	-1	+1	-1	-1	-1	+1	37,4	1,21	52,7	1,98	51,6	1,88
13	4	+1	+1	-1	-1	+1	+1	+1	-1	25,0	1,11	46,0	2,36	51,0	2,13
14	33	+1	+1	-1	-1	+1	+1	-1	-1	31,5	1,15	48,3	1,96	52,5	1,53
15	3	+1	+1	-1	-1	-1	+1	+1	+1	34,0	1,31	51,7	2,21	53,6	1,92
16	37	+1	+1	-1	-1	-1	+1	-1	-1	52,1	1,28	57,2	1,54	36,5	2,01
17	75	+1	-1	+1	+1	+1	-1	+1	+1	22,1	1,08	44,4	2,68	45,7	2,29
18	26	+1	-1	+1	+1	+1	-1	-1	-1	23,3	0,96	44,4	2,35	51,4	1,93
19	7	+1	-1	+1	+1	-1	-1	+1	-1	26,3	1,11	48,6	2,03	51,4	2,18
20	66	+1	-1	+1	+1	-1	-1	-1	+1	37,4	1,22	55,6	2,14	57,1	1,85
21	74	+1	-1	+1	-1	+1	+1	+1	+1	26,3	0,97	49,2	2,02	47,7	2,04
22	61	+1	-1	+1	-1	+1	+1	-1	-1	29,2	1,04	49,2	1,99	50,8	1,30
23	32	+1	-1	+1	-1	-1	+1	+1	-1	37,4	1,12	45,1	1,89	54,9	1,90
24	78	+1	-1	+1	-1	-1	+1	-1	+1	41.7	1,12	52,7	1,27	57,7	1,45
25	45	+1	-1	-1	+1	+1	+1	+1	-1	32,8	1,40	50,8	1,92	54,9	1,96
26	24	+1	-1	-1	+1	+1	+1	-1	+1	32,8	1,04	52,0	2,01	54,9	1,53
27	68	+1	-1	-1	+1	-1	+1	+1	+1	41,7	1,39	55,6	2,05	49,2	2,53
28	44	+1	-1	-1	+1	-1	+1	-1	-1	46,9	1,30	62,2	1,59	60,0	1,47
29	56	+1	-1	-1	-1	+1	-1	+1	-1	18,8	1,06	41,6	2,59	44,0	2,49
30	20	+1	-1	-1	-1	+1	-1	-1	+1	17,7	0,88	42,5	2,13	47,7	1,91
31	38	+1	-1	-1	-1	-1	-1	+1	+1	25,0	1,11	44,4	2,23	47,7	1,99
32	31	+1	-1	-1	-1	-1	-1	-1	-1	37,4	1,21	52,7	1,89	51,6	1,61
33	2	-1	+1	+1	+1	+1	-1	+1	-1	10,0	0,98	44,4	2,90	52,5	2,38
34	41	-1	+1	+1	+1	+1	-1	-1	+1	23,3	1,30	48,4	1,96	51,0	1,90
35	48	-1	+1	+1	+1	-1	-1	+1	+1	31,7	1,36	49,2	2,08	49,9	2,15
36	57	-1	+1	+1	+1	-1	-1	-1	-1	34,0	1,37	53,3	1,75	53,0	1,72
37	47	-1	+1	+1	-1	+1	+1	+1	-1	29,2	1,32	46,3	2,05	52,7	1,86
38	6	-1	+1	+1	-1	+1	+1	-1	+1	33,3	1,41	55,6	1,94	56,0	2,16
39	17	-1	+1	+1	-1	-1	+1	+1	+1	29,6	2,02	55,3	1,81	54,9	1,64
40	73	-1	+1	+1	-1	-1	+1	-1	-1	37,4	1,03	53,6	1,49	59,4	1,61
41	30	-1	+1	-1	+1	+1	+1	+1	+1	25,0	1,14	48,3	2,08	49,9	2,02
42	14	-1	+1	-1	+1	+1	+1	-1	-1	32,8	1,15	49,2	1,80	54,9	1,53
43	46	-1	+1	-1	+1	-1	+1	+1	-1	37,4	1,51	55,1	2,04	55,8	1,86
44	11	-1	+1	-1	+1	-1	+1	-1	+1	46,9	1,65	61,1	1,59	57,7	1,50
45	70	-1	+1	-1	-1	+1	-1	+1	+1	15,0	1,02	48,0	2,58	47,7	2,45
46	43	-1	+1	-1	-1	+1	-1	-1	-1	23,3	1,18	46,3	2,05	52,3	2,16
47	23	-1	+1	-1	-1	-1	-1	+1	-1	16,7	0,87	49,2	2,26	47,7	2,04
48	42	-1	+1	-1	-1	-1	-1	-1	+1	37,5	1,45	52,7	1,83	52,6	1,85
49	72	-1	-1	+1	+1	+1	+1	+1	-1	33,3	1,29	46,7	1,88	52,7	2,00
50	1	-1	-1	+1	+1	+1	+1	-1	+1	33,3	1,48	51,4	1,96	58,5	1,86
51	67	-1	-1	+1	+1	-1	+1	+1	+1	34,1	1,28	49,2	1,75	50,5	2,34
52	27	-1	-1	+1	+1	-1	+1	-1	-1	52,1	1,38	61,6	1,63	58,1	1,40
53	69	-1	-1	+1	-1	+1	-1	+1	-1	26,3	1,45	44,4	2,64	47,7	2,45
54	71	-1	-1	+1	-1	+1	-1	-1	+1	26,3	1,16	46,7	2,36	49,5	2,14
55	51	-1	-1	+1	-1	-1	-1	+1	+1	26,0	1,34	48,3	2,28	49,9	2,20
56	52	-1	-1	+1	-1	-1	-1	-1	-1	37,4	1,37	58,1	2,13	54,3	1,86
57	39	-1	-1	-1	+1	+1	-1	+1	+1	13,3	0,82	46,7	2,39	47,4	2,08
58	40	-1	-1	-1	+1	+1	-1	-1	-1	26,3	1,21	42,4	2,22	52,2	1,91
59	34	-1	-1	-1	+1	-1	-1	+1	-1	23,3	1,17	50,5	2,36	47,7	2,18
60	59	-1	-1	-1	+1	-1	-1	-1	+1	37,4	1,31	53,3	1,99	52,7	1,89
61	62	-1	-1	-1	-1	+1	+1	+1	+1	29,6	1,26	46,7	2,20	49,9	2,04
62	29	-1	-1	-1	-1	+1	+1	-1	-1	39,6	1,32	51,7	1,90	50,3	1,52
63	10	-1	-1	-1	-1	-1	+1	+1	-1	37,4	1,06	52,7	1,73	54,4	1,86
64	50	-1	-1	-1	-1	-1	+1	-1	+1	47,4	1,53	49,8	1,73	57.1	1,47

65	99	$+\sqrt{8}$	0	0	0	0	0	0	0	34,0	1,26	42,5	1,88	44,8	2,00
66	97	$-\sqrt{8}$	0	0	0	0	0	0	0	32,8	1,31	42,5	2,08	49,5	1,86
67	91	0	$+\sqrt{8}$	0	0	0	0	0	0	37,5	1,29	49,5	1,79	50,0	1,82
68	85	0	$-\sqrt{8}$	0	0	0	0	0	0	37,4	1,31	51,7	1,97	52,6	1,98
69	82	0	0	$+\sqrt{8}$	0	0	0	0	0	38,5	1,34	52,1	1,87	50,8	1,82
70	83	0	0	$-\sqrt{8}$	0	0	0	0	0	41,7	1,48	51,7	2,06	51,1	1,98
71	94	0	0	0	$+\sqrt{8}$	0	0	0	0	41,7	1,58	49,5	2,08	52,6	1,98
72	92	0	0	0	$-\sqrt{8}$	0	0	0	0	31,7	1,26	50,0	2,16	48,2	2,00
73	93	0	0	0	0	$+\sqrt{8}$	0	0	0	16,3	0,90	46,7	2,10	45,0	2,24
74	84	0	0	0	0	$-\sqrt{8}$	0	0	0	53,3	1,66	57,2	1,77	53,4	1,71
75	79	0	0	0	0	0	$+\sqrt{8}$	0	0	52,2	1,48	60,7	1,51	56,0	1,37
76	88	0	0	0	0	0	$-\sqrt{8}$	0	0	16,3	0,86	40,0	2,31	41,8	2,38
77	80	0	0	0	0	0	0	$+\sqrt{8}$	0	18,8	1,09	45,1	2,26	42,3	2,22
78	86	0	0	0	0	0	0	$-\sqrt{8}$	0	46,9	1,39	58,2	1,69	54,8	1,54
79	96	0	0	0	0	0	0	0	$+\sqrt{8}$	37,5	1,33	51,7	2,02	53,4	1,95
80	98	0	0	0	0	0	0	0	$-\sqrt{8}$	41,7	1,43	51,7	1,91	53,4	1,82
81	5	0	0	0	0	0	0	0	0	37,5	1,42	54,6	1,97	54,9	1,74
82	9	0	0	0	0	0	0	0	0	33,3	1,27	53,7	1,74	52,3	1,82
83	13	0	0	0	0	0	0	0	0	37,5	1,74	55,9	1,82	52,5	1,96
84	16	0	0	0	0	0	0	0	0	35,4	1,27	50,5	1,83	49,9	1,74
85	19	0	0	0	0	0	0	0	0	37,4	1,51	52,7	2,14	29,9	1,74
86	22	0	0	0	0	0	0	0	0	37,5	1,24	50,8	2,07	49,9	1,69
87	25	0	0	0	0	0	0	0	0	37,5	1,26	54,6	2,06	52,3	1,79
88	53	0	0	0	0	0	0	0	0	37,5	1,30	50,2	2,06	52,7	1,91
89	54	0	0	0	0	0	0	0	0	34,0	1,21	50,5	1,97	54,9	1,96
90	60	0	0	0	0	0	0	0	0	34,1	1,21	55,6	2,14	52,7	1,85
91	64	0	0	0	0	0	0	0	0	33,3	1,22	48,3	2,04	52,3	1,91
92	65	0	0	0	0	0	0	0	0	37,4	1,51	49,2	1,99	52,7	1,86
93	76	0	0	0	0	0	0	0	0	33,3	1,15	52,2	1,88	54,9	1,74
94	77	0	0	0	0	0	0	0	0	29,6	1,08	46,7	2,08	54,9	1,88
95	81	0	0	0	0	0	0	0	0	37,4	1,39	52,4	2,00	50,0	1,82
96	87	0	0	0	0	0	0	0	0	37,5	1,33	51,7	1,90	50,8	1,84
97	89	0	0	0	0	0	0	0	0	39,7	1,30	51,7	2,00	51,1	2,00
98	90	0	0	0	0	0	0	0	0	37,4	1,34	52,2	1,92	51,5	1,92
99	95	0	0	0	0	0	0	0	0	41,7	1,51	52,2	1,91	52,2	1,86
100	100	0	0	0	0	0	0	0	0	44,3	1,40	50,5	1,70	52,0	1,86

Tabel II. Verwachtingen en berekende waarden van de schattingen van de coëfficiënten

effect	verwachting	D_1 0	D_1 RF	D_2 0	D_2 RF	D_3 0	D_3 RF
b_0	β_0	36.665	1.3330	51.81	1.961	52.22	1.845
b_1	$\beta_1 + \beta_{111} + \beta_{122} + \dots + \beta_{188}$	+ 0.3624	-0.0314	- 0.6037	+0.0034	- 0.8187	-0.0058
b_2	$\beta_2 + \beta_{112} + \beta_{222} + \dots + \beta_{288}$	- 0.5197	+0.0134	+ 0.1910	-0.0201	- 0.2119	-0.0014
b_3	$\beta_3 + \beta_{113} + \beta_{223} + \dots + \beta_{388}$	- 0.3231	+0.0064	+ 0.1004	-0.0270	+ 0.6269	-0.0107
b_4	$\beta_4 + \beta_{114} + \beta_{224} + \dots + \beta_{488}$	+ 1.2411	+0.0285	+ 0.6311	-0.0148	+ 0.8506	-0.0059
b_5	$\beta_5 + \beta_{115} + \beta_{225} + \dots + \beta_{588}$	- 5.3756	-0.0560	- 2.6100	+0.1127	+ 1.0895	+0.0710
b_6	$\beta_6 + \beta_{116} + \beta_{226} + \dots + \beta_{688}$	+ 5.3942	+0.0531	+ 2.3902	-0.1658	+ 1.9970	-0.1505
b_7	$\beta_7 + \beta_{117} + \beta_{227} + \dots + \beta_{788}$	- 4.9335	-0.0470	- 1.7494	+0.1379	- 1.5444	+0.1665
b_8	$\beta_8 + \beta_{118} + \beta_{228} + \dots + \beta_{888}$	- 0.1210	+0.0088	- 0.0963	+0.0119	+ 0.0800	+0.0101
b_{11}	β_{11}	- 0.9006	-0.0149	- 0.7702	+0.0097	- 0.4605	+0.0121
b_{22}	β_{22}	- 0.3943	-0.0130	- 0.1952	-0.0028	+ 0.0582	+0.0083
b_{33}	β_{33}	- 0.0631	+0.0008	- 0.0327	+0.0078	- 0.0145	+0.0083
b_{44}	β_{44}	- 0.4881	-0.0005	- 0.3014	+0.0272	- 0.0543	+0.0196
b_{55}	β_{55}	- 0.7256	-0.0155	- 0.0264	+0.0041	- 0.2043	+0.0177
b_{66}	β_{66}	- 0.7943	-0.0292	- 0.2139	+0.0010	- 0.2418	+0.0052
b_{77}	β_{77}	- 0.9694	-0.0205	- 0.0639	+0.0091	- 0.2855	+0.0058
b_{88}	β_{88}	- 0.1256	-0.0030	- 0.0577	+0.0078	+ 0.3207	+0.0064
b_{12}	$\beta_{12} + \beta_{346}$	+ 1.2223	+0.0386	- 0.1578	+0.0278	- 0.6221	+0.0106
b_{13}	$\beta_{13} + \beta_{246} + \beta_{578}$	- 0.5252	-0.0135	- 0.1672	-0.0053	+ 0.1626	-0.0535
b_{14}	$\beta_{14} + \beta_{236}$	+ 1.0410	+0.0042	+ 0.6422	+0.0194	+ 0.6158	+0.0181
b_{15}	$\beta_{15} + \beta_{378}$	- 0.5095	+0.0014	+ 0.0078	+0.0025	- 0.0375	-0.0188
b_{16}	$\beta_{16} + \beta_{234}$	- 0.1751	+0.0070	+ 0.4266	+0.0063	- 0.1532	-0.0031
b_{17}	$\beta_{17} + \beta_{358}$	- 0.2251	+0.0298	+ 0.0859	+0.0159	+ 0.4439	+0.0134
b_{18}	$\beta_{18} + \beta_{357}$	+ 0.2470	-0.0117	- 0.2828	+0.0069	+ 0.4283	-0.0416
b_{23}	$\beta_{23} + \beta_{146}$	- 0.4220	+0.0048	+ 0.1078	-0.0244	+ 0.2907	-0.0297
b_{24}	$\beta_{24} + \beta_{136}$	+ 0.7065	+0.0095	- 0.4266	-0.0109	- 0.0438	-0.0431
b_{25}	β_{25}	- 0.1500	+0.0136	+ 0.0016	-0.0128	+ 0.5408	+0.0256
b_{26}	$\beta_{26} + \beta_{134}$	- 0.1969	-0.0086	+ 0.1266	+0.0403	- 0.1188	-0.0094
b_{27}	β_{27}	- 1.4098	-0.0261	+ 0.3109	+0.0394	+ 0.7534	-0.0541
b_{28}	β_{28}	+ 1.1504	+0.0298	+ 0.3234	-0.0041	+ 0.5127	-0.0428
b_{34}	$\beta_{34} + \beta_{126}$	+ 0.1282	-0.0064	- 0.2984	-0.0059	- 0.2595	-0.0078
b_{35}	$\beta_{35} + \beta_{178}$	- 0.0156	+0.0408	+ 0.3234	+0.0184	- 0.5252	+0.0203
b_{36}	$\beta_{36} + \beta_{124}$	- 0.4188	-0.0439	- 0.3828	-0.0259	+ 0.1219	-0.0109
b_{37}	$\beta_{37} + \beta_{158}$	+ 0.6440	+0.0142	- 0.4047	-0.0269	- 0.2751	-0.0256
b_{38}	$\beta_{38} + \beta_{157}$	- 0.1657	-0.0030	+ 0.2203	-0.0047	+ 0.1094	+0.0131
b_{45}	β_{45}	- 0.5627	-0.0377	- 0.3859	-0.0363	+ 0.2470	-0.0344
b_{46}	$\beta_{46} + \beta_{123}$	+ 0.2219	+0.0439	+ 0.6203	+0.0025	+ 0.2501	+0.0044
b_{47}	β_{47}	- 0.5220	+0.0008	+ 0.1609	-0.0134	- 0.2782	+0.0159
b_{48}	β_{48}	+ 0.7940	+0.0080	- 0.1953	+0.0019	- 0.4439	+0.0059
b_{56}	β_{56}	- 0.4658	+0.0048	+ 0.4547	-0.0081	+ 0.0656	-0.0456
b_{57}	$\beta_{57} + \beta_{138}$	+ 1.0034	+0.0142	+ 0.6516	-0.0091	+ 0.4189	-0.0203
b_{58}	$\beta_{58} + \beta_{137}$	- 0.5002	-0.0323	+ 0.7078	+0.0000	- 0.2845	-0.0216
b_{67}	β_{67}	+ 0.5689	+0.0202	- 0.0609	-0.0128	+ 0.4970	+0.0041
b_{68}	β_{68}	- 0.9972	+0.0030	- 0.2609	+0.0056	+ 0.5377	+0.0241
b_{78}	$\beta_{78} + \beta_{135}$	- 0.3848	-0.0120	+ 0.3109	+0.0003	- 0.8909	+0.0025
b_{125}	β_{125}	+ 0.2313	+0.0233	- 0.1172	-0.0147	+ 0.4626	-0.0084

Tabel II (vervolg) Verwachtingen en berekende waarden van de schattingen van de coëfficiënten

		D_1 O	D_1 RF	D_2 O	D_2 RF	D_3 O	D_3 RF
b_{127}	β_{127}	- 1.4911	-0.0189	- 0.1422	-0.0100	+ 0.5439	-0.0250
b_{128}	β_{128}	- 0.4064	-0.0017	- 0.8391	+0.0003	+ 0.7221	-0.0388
b_{145}	β_{145}	+ 1.0878	+0.0283	+ 0.2953	-0.0063	- 0.3126	+0.0097
b_{147}	β_{147}	- 0.3470	+0.0067	+ 0.3797	-0.0122	- 0.1188	-0.0256
b_{148}	β_{148}	+ 0.8440	+0.0048	- 0.3078	+0.0256	- 0.1907	+0.0069
b_{156}	β_{156}	- 0.9034	-0.0273	+ 0.3047	-0.0118	+ 0.5502	-0.0309
b_{167}	β_{167}	+ 0.0688	+0.0117	+ 0.3517	+0.0247	+ 0.6252	-0.0044
b_{168}	β_{168}	+ 0.0406	-0.0142	- 0.1297	-0.0219	+ 0.4470	-0.0269
b_{235}	β_{235}	- 0.4314	+0.0227	+ 0.1984	-0.0088	- 0.1876	+0.0081
b_{237}	β_{237}	+ 0.1844	+0.0004	+ 0.4141	-0.0091	- 0.1000	-0.0191
b_{238}	β_{238}	- 0.1500	-0.0130	- 0.1859	+0.0113	- 0.0094	-0.0091
b_{245}	$\beta_{245} + \beta_{678}$	- 0.3720	-0.0327	+ 0.4266	-0.0078	- 0.7221	+0.0228
b_{247}	$\beta_{247} + \beta_{568}$	- 0.1250	+0.0039	+ 0.1422	+0.0150	+ 0.4283	+0.0156
b_{248}	$\beta_{248} + \beta_{567}$	+ 0.4408	+0.0061	- 0.7016	-0.0491	- 0.0813	-0.0281
b_{256}	$\beta_{256} + \beta_{478}$	- 0.2876	-0.0214	- 0.3953	-0.0091	- 0.0782	+0.0047
b_{257}	$\beta_{257} + \beta_{468}$	- 0.8878	-0.0202	- 0.6547	+0.0050	- 0.1313	+0.0088
b_{258}	$\beta_{258} + \beta_{467}$	+ 0.1469	+0.0183	- 0.2609	-0.0053	- 0.5283	+0.0075
b_{267}	$\beta_{267} + \beta_{458}$	+ 0.2032	+0.0283	+ 0.1828	+0.0100	+ 0.4408	-0.0556
b_{268}	$\beta_{268} + \beta_{457}$	- 0.7502	-0.0214	+ 0.1328	-0.0059	+ 0.7878	-0.0419
b_{278}	$\beta_{278} + \beta_{456}$	- 0.4064	-0.0008	- 0.3891	-0.0175	- 0.0719	+0.0216
b_{345}	β_{345}	- 0.9315	-0.0130	+ 0.2797	+0.0028	+ 0.3501	+0.0131
b_{347}	β_{347}	- 0.2657	-0.0145	+ 0.0578	+0.0081	+ 0.5377	-0.0091
b_{348}	β_{348}	- 0.1938	+0.0283	- 0.6859	+0.0184	+ 0.8784	-0.0291
b_{356}	β_{356}	+ 0.3845	+0.0189	+ 0.6078	-0.0097	- 0.0250	+0.0219
b_{367}	β_{367}	- 0.8065	-0.0245	+ 0.2734	+0.0069	- 0.2813	-0.0272
b_{368}	β_{368}	+ 0.1469	+0.0002	+ 0.1109	-0.0359	+ 0.0907	+0.0053
Schatting van spreiding van							
$\underline{b}_0$		+ 0.723	+0.0340	+ 0.528	+0.0276	+0.375	+0.0191
$\underline{b}_i$		+ 0.361	+0.0170	+ 0.264	+0.0138	+0.189	+0.0096
$\underline{b}_{ii}$		+ 0.286	+0.0134	+ 0.209	+0.0109	+0.150	+0.0076
$\underline{b}_{ij}$		+ 0.404	+0.0190	+ 0.295	+0.0154	+0.212	+0.0107
$\underline{b}_{ijk}$		+ 0.404	+0.0190	+ 0.295	+0.0154	+0.212	+0.0107

zodat (3;2) wat de proefopzet betreft een term $(\beta_{12}+\beta_{346})x_1x_2$ bevat in plaats van de twee termen $\beta_{12}\,x_1x_2$ en $\beta_{346}\,x_3x_4x_6$. Dit betekent dus tevens dat men niet β_{12} en β_{346} apart zal schatten, maar slechts hun som $\beta_{12}+\beta_{346}$. Op deze wijze kan men uit (3;1) vinden dat slechts 72 verschillende effecten geschat kunnen worden, die elk al of niet een som zijn van een tweetal coëfficiënten van (3;2). In Tabel II (2^e kolom) is nu vermeld, welke coëfficiënten apart en welke tezamen geschat worden uit de waarnemingen die bij de proefopzet van Tabel I verricht kunnen worden.

Alvorens in te gaan op de eigenschappen van de proefopzet, zoals deze uit Tabel II volgen, willen wij eerst aangeven hoe de kleinste kwadraten schattingen $b_0, \ldots, b_{368}$ van de effecten berekend worden. Daartoe duiden wij de vector $(b_0, \ldots, b_{368})$ aan met b'. Voorts definiëren wij de 72 x 100 matrix X als volgt. De eerste kolom, die wij aan zullen geven met (0), bevat 100 dezelfde elementen, nl. +1. De tweede tot en met de negende kolom, die wij aangeven met resp. (1),...,(8) zijn de kolommen, genummerd 1,...,8 in Tabel I. Tenslotte definiëren wij de overige 63 kolommen, die wij aanduiden met achtereenvolgens (11),...,(88), (12),...,(78), (125),...,(368) als volgt: het element op de u-de rij van de kolom aangegeven door ab (resp. abc) is het product van de elementen die zich bevinden op de u-de rij en de kolommen aangegeven door (a) en (b) (resp. door (a), (b) en (c)). Zij y de vector der waarnemingsresultaten, dan is dus volgens (2.2;1):

$$b = (X'X)^{-1}X'y \qquad \text{- - - - - -} \qquad (3;3)$$

waarbij $(X'X)^{-1}$ wordt gevonden als:

$$(X'X)^{-1} = \begin{array}{c} \\ (0) \\ (1) \\ \vdots \\ (8) \\ (11) \\ \vdots \\ (88) \\ (12) \\ \vdots \\ (78) \\ (125) \\ \vdots \\ (368) \end{array} \begin{array}{c} \begin{array}{ccccccccccccccccccc} (0) & (1) & & (8) & (11) & & & (88) & (12) & & & (78) & (125) & & & (368) \end{array} \\ \left(\begin{array}{c|ccc|c|ccc|c|ccc|c|ccc}
1/20 & 0 & \cdots & 0 & -1/160 & \cdots & & -1/160 & 0 & \cdots & & 0 & 0 & \cdots & & 0 \\
0 & 1/80 & & & & & & & & & & & & & & \\
\vdots & & \ddots & & & & & & & & & & & & & \\
\vdots & & & 0 & & & & & & & & & & & & \\
0 & & & 1/80 & & & & & & & & & & & & \\
\hline
-1/160 & & & & 1/128 & & & & & & & & & & & \\
\vdots & & & & & \ddots & & & & & & & & & & \\
\vdots & & & & & & 0 & & & & & & & & & \\
-1/160 & & & & & & & 1/128 & & & & & & & & \\
\hline
0 & & & & & & & & 1/64 & & & & & & & \\
\vdots & & & & & & & & & \ddots & & & & & & \\
\vdots & & & & & & & & & & 0 & & & & & \\
0 & & & & & & & & & & & 1/64 & & & & \\
0 & & & & & & & & & & & & 1/64 & & & \\
\vdots & & & & & & & & & & & & & \ddots & & \\
\vdots & & & & & & & & & & & & & & 0 & \\
0 & & & & & & & & & & & & & & & 1/64
\end{array}\right) \end{array} \qquad (3;4)$$

Hierin is een blokmatrix, die door een nul wordt aangegeven, een submatrix, waarvan alle elementen $\equiv 0$ zijn.

De beschreven proefopzet heeft de volgende eigenschappen:

1) Het is mogelijk alle coëfficiënten van het tweedegraadsmodel (2.1;9) te schatten. Deze schattingen zullen alleen zuiver zijn, indien $\beta_{346}, \beta_{246}, \ldots, \beta_{137}$ alle $\equiv 0$ zijn (zie Tabel II). Dit is zeker het geval, indien het responsieoppervlak ter plaatse een tweedegraadsoppervlak is. De relatie (3.1) is zodanig gekozen, dat indien een β_{ij} tezamen wordt geschat met een $\beta_{i'j'k}$, in de meeste gevallen van deze $\beta_{i'j'k}$ op grond van chemische overwegingen mag worden aangenomen dat deze $\equiv 0$ zijn. Op deze wijze blijven schattingen van de β_{ij} zuiver, ook al is het responsieoppervlak geen tweede-, maar een derdegraadsoppervlak.
2) Indien het responsieoppervlak een oppervlak van een graad > 2 is, zal zich dit uiten in het optreden van systematisch van nul afwijkende b_{ijk}.
3) Uit (2.2;2) en (3.4) volgt dat elke schatting ongecorreleerd is met elke andere, met uitzondering van de b_0 en de b_{ii}. Deze zijn negatief gecorreleerd ($\rho = -0{,}32$), maar de b_{ii} zijn onderling wel ongecorreleerd. Deze ongecorreleerdheid wordt teweeggebracht door de afstand $\sqrt{8}$ van punten van de oktaëder tot het centrum en door het uitvoeren van 20 proeven in het centrum.
4) Indien de schattingen b berekend zouden zijn, kan men deze in (2.1;9) invullen en op grond van deze vergelijking voorspellingen doen ten aanzien van de responsie in een punt $x_1^*, \ldots, x_8^*$. Het is heel eenvoudig de variantie in de verwachte waarde van de responsie na te gaan tengevolge van het steekproefkarakter van de b's. Daar uit:

$$y = \underline{b}_0 + \underline{b}_1 x_1^* + \cdots + \underline{b}_8 x_8^* + \underline{b}_{11} x_1^{*2} + \cdots + \underline{b}_{88} x_8^{*2} + \underline{b}_{12} x_1^* x_2^* + \cdots + \underline{b}_{78} x_7^* x_8^* \qquad (3.5)$$

volgt:

$$var\{y\} = var\,\underline{b}_0 + \sum_i x_i^{*2}\, var\,\underline{b}_i + \sum_i x_i^{*4}\, var\,\underline{b}_{ii} + \sum_{i<j} x_i^{*2} x_j^{*2}\, var\,\underline{b}_{ij} + \sum_i x_i^{*2}\, cov\{\underline{b}_0, \underline{b}_{ii}\} + \text{andere covarianties,}$$

is volgens (2.2;2) en (3.4):

$$var\{y\} = \sigma^2\left\{\frac{1}{20} + \frac{1}{80}\sum_i x_i^{*2} + \frac{1}{128}\sum_i x_i^{*4} + \frac{1}{64}\sum_{i<j} x_i^{*2} x_j^{*2} - \frac{1}{160}\sum_i x_i^{*2}\right\}$$

ofwel

$$var\{y\} = \sigma^2\left\{\frac{1}{20} + \frac{1}{160}\sum_i x_i^{*2} + \frac{1}{128}\left(\sum_i x_i^{*2}\right)^2\right\}. \qquad (3.6)$$

De variantie van een voorspelde waarde in een punt is dus uitsluitend een functie van de afstand van het punt tot het centrum en niet afhankelijk van de ligging van het punt ten opzichte van de coördinaat-assen. Dit is de roteerbaarheids-eigenschap, waaraan de proefopzet zijn naam ontleent. Het belang van deze eigenschap is duidelijk: men weet a priori niet in welke richting het optimum zich zal bevinden, en wil dus in staat zijn in elke richting even nauwkeurig voorspellingen te doen.

5) De schattingen b zijn zeer eenvoudig te berekenen, dank zij de eenvoud van (3;4). Zo is bijv.

b_0 = gemiddelde van de waarnemingen van de 20 proeven, verricht in het centrum van de proefopzet.

$b_I = \frac{1}{80}$ [som van de waarnemingen bij proeven, waarbij de eerste factor op +1 niveau werd gehouden - som van de waarnemingen bij proeven, waarbij de eerste factor op - 1 niveau werd gehouden +
$\sqrt{8}$ (waarneming bij proefopzet met 1^e factor op $+\sqrt{8}$
- " " " " " " $-\sqrt{8}$)

$b_{II} = \frac{1}{640}$ som van de waarnemingen bij proeven van het $\frac{1}{4} \times 2^8$ - schema - $\frac{1}{160}$ som van overige waarnemingen +
$\frac{1}{16}$ (som van de waarnemingen bij de proeven met de eerste factor op $+\sqrt{8}$ of $-\sqrt{8}$).

Door de hierboven beschreven eigenschappen leent een roteer baar (orthogonaal) proefschema zich bij uitstek voor het opsporen van optimale condities.

4. Zuivering van schapenantidifterieserum

4.1. Eerste experiment

Het schapenantidifterieserum werd ten tijde van de experimenten nog niet op grote schaal geproduceerd. Ook waren slechts een gering aantal laboratoriumproeven verricht, zodat over het gehele zuiveringsproces zeer weinig bekend was. Zo was bijv. geen cijfermateriaal beschikbaar over de waarnemingsonnauwkeurigheden, die bij dit proces verwacht moesten worden.

Onder deze omstandigheden lag het voor de hand dat een oriënterend experiment van geringe omvang gewenst was, waarbij dit zo diende te worden ingericht, dat in een gunstig geval reeds hieruit enkele conclusies getrokken zouden kunnen worden ten aanzien van de ligging van het optimum, maar welk experiment in hoofdzaak gericht zou zijn op het verkrijgen van informatie, waarmee een groter en verdergaand experiment op een goede wijze kon worden opgezet.

Als eerste proefschema werd daarom door ons aanbevolen en dienovereenkomstig uitgevoerd een $\frac{1}{16}$ deel van een 2^8 factorieel schema in duplo, gekenmerkt door de relatie:

$$\text{eenheid} = 11 = \ldots = 88 = 1235 = 1246 = 1347 = 2348, \ldots \qquad (4.1;1)$$

Dit wil dus zeggen, dat 2 x 16 proeven uitgevoerd werden, waarbij de factoren op de (gestandaardiseerde) niveaux +1 en -1 gehouden werden, zoals in Tabel III is aangegeven.

Teneinde na te gaan welke mogelijkheden door deze proefopzet geboden worden om uit de hierbij te verrichten waarnemingen informatie te verkrijgen hoe een verbetering in de opbrengst resp. RF te bewerkstelligen zal zijn, beschouwen wij (2.1;9). Aangezien volgens (4.1;1) steeds geldt:

$$\left.\begin{array}{l} x_{1u}\, x_{2u}\, x_{3u}\, x_{5u} = +1 \\ x_{1u}\, x_{2u}\, x_{4u}\, x_{6u} = +1 \\ x_{1u}\, x_{3u}\, x_{4u}\, x_{7u} = +1 \\ x_{2u}\, x_{3u}\, x_{4u}\, x_{8u} = +1 \\ \text{en} \\ x_{1u}^2 = x_{2u}^2 = - - - = x_{8u}^2 = +1, \end{array}\right\} \qquad (4.1;2)$$

is het wat de proefopzet betreft, hetzelfde als (2.1;9) geschreven wordt als:

$$\begin{aligned} E\,\underline{y}_u = {} & (\beta_0 + \beta_{11} + \cdots + \beta_{88})\, x_{0u} + \beta_1\, x_{1u} + \cdots + \beta_8\, x_{8u} + \\ & + (\beta_{12} + \beta_{35} + \beta_{46} + \beta_{78})\, x_{1u}\, x_{2u} + (\beta_{13} + \beta_{25} + \beta_{47} + \beta_{68})\, x_{1u}\, x_{3u} + \\ & + (\beta_{14} + \beta_{26} + \beta_{37} + \beta_{58})\, x_{1u}\, x_{4u} + (\beta_{23} + \beta_{15} + \beta_{48} + \beta_{67})\, x_{2u}\, x_{3u} + \\ & + (\beta_{24} + \beta_{16} + \beta_{38} + \beta_{57})\, x_{2u}\, x_{4u} + (\beta_{34} + \beta_{17} + \beta_{28} + \beta_{56})\, x_{3u}\, x_{4u} + \\ & + (\beta_{45} + \beta_{18} + \beta_{27} + \beta_{36})\, x_{4u}\, x_{5u}. \qquad (4.1;3) \end{aligned}$$

Tabel III. Gestandaardiseerde waarde van de factoren, bij de proeven van het eerste experiment, en de bij de uitvoering verkregen waarnemingen

factoren									waarnemingen	
z_1	T_1	t_1	p	a	z_2	T_2	t_2	RIV	O	RF
+1	+1	+1	+1	+1	+1	+1	+1	8	23,6	1,36
								23	23,6	1,38
+1	+1	+1	-1	+1	-1	-1	-1	4	23,6	1,26
								25	23,6	1,24
+1	+1	-1	+1	-1	+1	-1	-1	7	60,8	1,50
								10	56,3	1,39
+1	+1	-1	-1	-1	-1	+1	+1	24	< 8,0	<1,18
								29	< 8,0	<0,92
+1	-1	+1	+1	-1	-1	+1	-1	11	12,0	1,06
								21	12,0	0,97
+1	-1	+1	-1	-1	+1	-1	+1	1	45,0	1,47
								26	40,0	1,21
+1	-1	-1	+1	+1	-1	-1	+1	15	26,6	1,49
								17	23,6	1,40
+1	-1	-1	-1	+1	+1	+1	-1	12	18,8	1,10
								13	18,8	1,12
-1	+1	+1	+1	-1	-1	-1	+1	20	31,5	1,52
								31	30,0	1,53
-1	+1	+1	-1	-1	+1	+1	-1	6	31,5	1,22
								19	35,5	1,29
-1	+1	-1	+1	+1	-1	+1	-1	9	< 5,5	< 0,81
								27	< 8,0	< 0,80
-1	+1	-1	-1	+1	+1	-1	+1	14	33,8	1,32
								16	30,0	1,26
-1	-1	+1	+1	+1	+1	-1	-1	28	35,5	1,35
								32	40,0	1,43
-1	-1	+1	-1	+1	-1	+1	+1	3	< 5,5	< 0,93
								18	< 8,0	<1,00
-1	-1	-1	+1	-1	+1	+1	+1	2	31,5	1,48
								22	28,0	1,30
-1	-1	-1	-1	-1	-1	-1	-1	5	45,0	1,39
								30	45,0	1,44

Hieruit volgt dat men bijv. β_0 en ieder der β_{ii} niet apart kan schatten, maar dat uitsluitend een schatting berekend kan worden van $\beta_0 + \sum_i \beta_{ii}$. Ook worden de β_{ij} in groepen van vier tezamen geschat, zoals dit is aangegeven in Tabel IV. In de eerste kolom staat hierbij de aanduiding voor de kleinste kwadratenschatting van de effecten in de tweede kolom.

Uit Tabel IV volgt nu dat indien het responsieoppervlak plaatselijk (redelijk) goed door een plat vlak te benaderen is, de coëfficiënten $\beta_0, \beta_1, ---, \beta_8$, die in de vergelijking van dit plat vlak voorkomen, zuiver te schatten zijn. In dit geval zal toepassing van de methode van de steilste weg naar factorencombinaties leiden, waarvan verwacht mag worden dat deze een hogere O, resp. RF, zullen geven. Wijkt het responsieoppervlak af van een plat vlak, dan zal dit tot uiting komen in het systematisch van nul verschillen van de b_{ij} . Past men in dit geval toch de methode van de steilste weg toe, dan zal de verwachte verbetering in het algemeen niet optreden.

De berekening van de kleinste kwadratenschattingen is zeer eenvoudig. Definiëren wij als X de matrix met als u -de rij ($u = 1,\ldots,16$) de elementen $X_{0u}, X_{1u}, ---, X_{8u}$, $X_{1u}X_{2u}, X_{1u}X_{3u}, ------, X_{7u}X_{8u}$, dan is:

$$(X'X) = 8\,I\,, \tag{4.1;4}$$

waarin I de 16 x 16 eenheidsmatrix is. De schattingen worden dan

$$\left.\begin{aligned} b_0 &= \frac{1}{8}\sum_{u=1}^{8} y_u \\ b_i &= \frac{1}{8}\sum_{u=1}^{8} x_{iu}\,y_u \\ b_{ij} &= \frac{1}{8}\sum_{u=1}^{8} x_{iu}\,x_{ju}\,y_u \end{aligned}\right\} \tag{4.1;5}$$

Uit (2.2;2) en (4.1;4) volgt dan ook direct, dat alle schattingen onderling ongecorreleerd zijn en dezelfde variantie bezitten.

In gezamenlijk overleg met het RIV werden de waarden +1 en -1 van de factoren gekozen als in Tabel V is aangegeven.

Aan het RIV was verzocht de proeven af te werken in de volgorde aangegeven door de nummering in de kolom: "no RIV" van Tabel III.

Tabel IV. Uit het eerste experiment berekende schattingen en de verwachtingen hiervan

berekende schatting	verwachting van de schatting	O	RF
b_0	$\beta_0 + \sum_i \beta_{11}$	26,1	1,17
b_1	β_1	- 0,2	0,01
b_2	β_2	- 0,3	- 0,03
b_3	β_3	- 0,4	0,03
b_4	β_4	1,3	0,09
b_5	β_5	- 5,6	- 0,06
b_6	β_6	8,4	0,15
b_7	β_7	-11,0	- 0,22
b_8	β_8	- 2,9	0,00
b_{12}	$\beta_{12} + \beta_{35} + \beta_{46} + \beta_{78}$	1,6	- 0,01
b_{13}	$\beta_{13} + \beta_{25} + \beta_{47} + \beta_{68}$	0,3	0,02
b_{14}	$\beta_{14} + \beta_{26} + \beta_{37} + \beta_{58}$	2,7	0,05
b_{23}	$\beta_{23} + \beta_{15} + \beta_{48} + \beta_{67}$	2,8	0,17
b_{24}	$\beta_{24} + \beta_{16} + \beta_{38} + \beta_{57}$	1,5	- 0,02
b_{34}	$\beta_{34} + \beta_{17} + \beta_{28} + \beta_{56}$	- 1,0	0,03
b_{45}	$\beta_{45} + \beta_{18} + \beta_{27} + \beta_{36}$	0,2	- 0,02
geschatte spreiding van ieder der b's		0,4	0,01

Tabel V. Niveaux der factoren, zoals deze in het eerste experiment betrokken werden

factor nummer	factor	+ 1	- 1
1	z_1	3,5	2,9
2	T_1	30°	20°
3	t_1	40'	20'
4	p	600 mg	200 mg
5	a	17%	13%
6	z_2	4,6	4,0
7	T_2	60°	50°
8	t_2	80'	40'

De verkregen waarnemingen zijn in Tabel III vermeld, de daaruit berekende schattingen in Tabel IV, tezamen met een schatting van de spreiding, berekend uit verschillen tussen duplo waarnemingen. Bij de berekening van de schattingen werden voor de proeven 24, 29, 3, 18, 9 en 27 de volgende waarden gebruikt:

proef no.	O	RF
24 en 29	1,5	0,50
3 en 18	0,5	0,50
9 en 27	1,5	0,50

Het is gebleken, dat de resultaten betrekkelijk ongevoelig zijn voor een andere keuze van deze getallen.

Ten aanzien van de resultaten (Tabel IV) merken wij de volgende punten op.

1) De factoren z_1, T_1 en t_1 vertonen geen invloed van betekenis. Er zijn hiervoor 3 mogelijke verklaringen:
   a) de factoren hebben inderdaad geen invloed,
   b) de factoren hebben wel invloed, maar de verschillen tussen de gekozen waarden ervan voor het +1 en -1 niveau zijn te gering om deze invloed duidelijk te laten blijken, of
   c) de factoren hebben wel invloed, maar de gekozen waarden voor het +1 en -1 niveau bevinden zich aan weerszijden van een voorwaardelijk maximum of minimum.

2) De factoren p, a, z_2, T_2 vertonen matige tot zeer grote lineaire effecten. Een bijzondere omstandigheid, die van veel belang is voor het opsporen van optimale condities is dat genoemde factoren een invloed in dezelfde richting op de O en de RF hebben. Verhoging van de hoeveelheid pepsine zal bijvoorbeeld volgens Tabel IV aanleiding geven tot een verhoging van de O en de RF.

3) De factor t_2 vertoont geen invloed op de RF, maar wel bij de opbrengst.

4) De interactieeffecten zijn zeker niet alle klein. Zo is b_{23} bij de RF groter dan het grootste lineaire effect.

De overweging dat slechts weinig proeven nodig zouden zijn om de steilste weg te onderzoeken leidde ertoe dat deze methode werd toegepast, ondanks het feit, dat het responsieoppervlak hoogstens bij grove benadering als een plat vlak kon worden voorgesteld.

De richtingscosinussen van de steilste weg in de 8 dimensionale ruimte waarin de niveaux van de factoren door +1 en -1 worden weergegeven zijn evenredig met $b_1, \ldots, b_8$. Door normering op 1 worden de richtingscosinussen zelf gevonden, zie Tabel VI.

Tabel VI. Richtingen van de steilste weg

factor	steilste weg voor opbrengst	steilste weg voor RF	weg gelegen in het midden tussen beide wegen	voorgestelde weg
1	- 0,014	0,045	0,016	0
2	- 0,020	- 0,101	- 0,062	0
3	- 0,026	0,118	0,047	0
4	0,083	0,306	0,198	0,200
5	- 0,366	- 0,222	- 0,2[illegible]	- 0,302
6	0,552	0,525	0,548	0,552
7	- 0,719	- 0,745	- 0,745	- 0,751
8	- 0,191	- 0,004	- 0,099	0

Getracht is in dit geval zowel een verbetering van de O als van de RF te verkrijgen door een weg te kiezen, die midden tussen de steilste wegen, berekend voor O en voor RF, in ligt.

De voorgestelde steilste weg van Tabel VI betekent dat proeven werden aanbevolen bij een aantal waarden van λ met de factorencombinatie :

nummer factor	factor	waarde van de factor
1	z_1	3,2
2	T_1	25°
3	t_1	30'
4	p	$(400 + 40\lambda)$ mg
5	a	$(15 - 0.6\lambda)\%$
6	z_2	$4.3 + 1/6\,\lambda$
7	T_2	$(55 - 3.8\lambda)^\circ$
8	t_2	60'

De resultaten van de proeven, die op grond van dit voorstel werden uitgevoerd zijn vermeld in Tabel VII.

Tabel VII. Waarnemingen van proeven verricht langs de voorgestelde steilste weg

Nummer proef	λ	Opbrengst	R.F.
1, 10	0	29,6, 29,6	1,42, 1,45
2, 11	1	47,0, 47,0	1,45, 1,53
3, 12	2	52,1, 52,1	1,31, 1,22
4, 13	3	54,7, 52,1	1,19, 1,17
5, 14	4	57,3, 54,7	1,01, 1,09
6, 15	5	57,3, 57,3	1,03, 0,94
7, 16	6	52,1, 52,1	0,85, 0,71
8, 17	7	52,1, 53,8	0,72, 0,66
9, 18	8	65,6, 65,6	0,95, 0,89

Uit Tabel VII blijkt dat de proeven langs de steilste weg een lagere RF te zien gaven dan die bij de eerste proefopzet (Tabel III). De gevolgtrekking, die gemaakt werd uit het feit, dat de verwachte toename in O en RF niet werd gerealiseerd, was dat de interacties tussen de factoren van een dergelijke orde van grootte waren, dat het responsieoppervlak niet door een plat vlak is te benaderen.

Een onderstelling, die in bovenstaande redenering stilzwijgend is gemaakt, is dat de kwaliteit van het uitgangsmateriaal in de ruim twee maanden, die tussen beide series proeven lagen, niet zou zijn veranderd. Het is achteraf gebleken (zie par. 5.6), dat deze onderstelling niet juist was. Inderdaad moet het als een fout in deze en de volgende proefopzet worden gezien, dat de genoemde onderstelling niet kon worden gecontroleerd.

Uit de verschillen tussen duplo waarnemingen die bij de beide groepen proeven verricht zijn, volgt als schatting van de variantie σ^2 van de waarnemingsonnauwkeurigheid.

voor de opbrengst: $s^2 = 3{,}087$
voor de RF: $s^2 = 0{,}0047$ (4.1;6)

ieder gebaseerd op 22 vrijheidsgraden. In verband met een later te bespreken kwestie wordt erop gewezen, dat in het hier besproken geval de duplowaarnemingen op éénzelfde dag verricht werden. Uit (4.1;6) volgt nu, dat indien twee waarnemingen van een factorencombinatie op één dag verricht worden, verwacht mag worden dat in circa 50% van de gevallen de verschillen tussen de waarnemingen van de opbrengst $\geq 1{,}7$ en tussen die van de RF $\geq 0{,}06$ zullen zijn.

De variantie (4.1;6) werd voldoende klein geacht om de in par. 2 beschreven experimenteertechniek verder te kunnen toepassen. Gezien het feit, dat het responsieoppervlak niet door een plat vlak te benaderen was, bleek het dus wenselijk een proefopzet van een dergelijke aard uit te voeren dat aan de hierbij te verkrijgen waarnemingen een tweedegraads oppervlak aangepast zou kunnen worden. Om deze redenen werd geadviseerd het roteerbare schema van paragraaf 3 uit te voeren.

4.2. Het tweede experiment

Het centrum, waar omheen dit experiment uitgevoerd zou worden, behoefde uiteraard niet samen te vallen met het punt met coördinaten 0,...,0. Men kan altijd trachten reeds verkregen informatie in rekening te brengen bij de keuze van het nieuwe centrum, om op deze wijze een gunstiger gebied van de factorenruimte in het onderzoek te betrekken. In Tabel VIII zijn de factorencombinaties opgesomd van die proeven, waarbij een RF $> 1{,}45$ was gevonden.

Tabel VIII. Gunstige factorencombinaties

RIV no's proef	factoren: z_1	T_1	t_1	p	a	z_2	T_2	t_2	O	R.F.
7 en 10 v/d 1e groep	+1	+1	-1	+1	-1	+1	-1	-1	58%	1,45
20 en 31 v/d 1e groep	-1	+1	+1	+1	-1	-1	-1	+1	30%	1,52
2 en 11 v/d 2e groep	0	0	0	$+\frac{1}{5}$	$-\frac{1}{3}$	$+\frac{1}{2}$	$-\frac{3}{4}$	0	47%	1,49

In overleg met het R.I.V. werd besloten het nieuwe centrum te kiezen bij 0, 0, 0, $+\frac{1}{2}$, $+\frac{1}{2}$, 0, $-\frac{1}{2}$ en 0. Hierbij werd op grond van chemische overwegingen (die buiten onze competentie vallen) het centrum van de 5-de factor (a) in tegenovergestelde richting verplaatst, dan die welke als meest gunstige uit de vorige proeven volgde (zie o.a. Tabel IV).

Om de notatie eenvoudig te houden is de oorsprong van het tot nu toe gebruikte coördinatenstelsel (waarbij de niveaux van de factoren van de proeven van de eerste proefopzet met +1 en -1 zijn aangegeven) verplaatst naar het nieuwe centrum, zodat in het nieuwe coördinatenstelsel de proefopzet weer om de oorsprong als centrum wordt uitgevoerd. Ook de eenheden op de assen zijn anders gekozen, en wel kleiner. Dit in verband met de overwegingen:

1) zelfs een kleinere eenheid geeft voor de relevante factoren nog wel duidelijk waarneembare effecten (zoals wij uit Tabel IV zien) en
2) hoe kleiner een eenheid, des te kleiner het door de proeven bestreken gebied en des te beter zal het responsieoppervlak benaderd kunnen worden door een tweedegraadsoppervlak.

In gezamenlijk overleg werden voor de eenheden de waarden gekozen als vermeld in Tabel IX.

Tabel IX. Overzicht van centrum, eenheid en niveaux van de proefopzet

Factor		centrum	eenheid	niveaux			
				$-\sqrt{8}$	- 1	+ 1	$+\sqrt{8}$
1	z_1	3,2	0,2	2,6	3,0	3,4	3,8
2	T_1	25°	$2,5^\circ$	18°	$22\frac{1}{2}^\circ$	$27\frac{1}{2}^\circ$	32°
3	t_1	30'	5'	16'	25'	35'	44'
4	p	500 mg	100 mg	217 mg	400 mg	600 mg	783 mg
5	a	16%	1%	13,17%	15%	17%	18,83%
6	z_2	4,3	0,2	3,7	4,1	4,5	4,9
7	T_2	$52\frac{1}{2}^\circ$	$2,5^\circ$	$45\frac{1}{2}^\circ$	50°	55°	$59\frac{1}{2}$%
8	t_2	60'	5'	46'	55'	65'	74'

De proeven 1-64 en 81 t/m 94 zijn in een volledig aselecte volgorde verricht, en wel zoals is aangegeven door de kolom "no RIV". In verband met organisatorische moeilijkheden konden de overige proeven slechts tezamen in één groep worden uitgevoerd, maar wel kon de volgorde binnen de groep aselect worden gekozen.

De waarneming verkregen bij het uitvoeren van deze proefopzet zijn vermeld in Tabel I, kolom D_1, terwijl de schattingen van de effecten, berekend als in 3 is aangegeven, vermeld zijn in Tabel II, kolom D_1.

4.3. De waarnemingen verricht in 0,...,0

In het centrum van de proefopzet zijn 20 proeven uitgevoerd, waarbij deze bovendien verdeeld waren over de drie dagen, die nodig waren voor de uitvoering van het experiment. Deze proeven zijn:

	MC no	RIV no
eerste dag	81 t/m 87	5, 9, 13, 16, 19, 22, 25
tweede dag	88 t/m 94	53,54, 60, 64, 65, 76, 77
derde dag	81 t/m 100	81,87, 89, 90, 95, 100 .

Uitgaande van de bij deze proeven verkregen waarnemingen is getoetst:

A) of een niveau-verschil optreedt tussen bepalingen op verschillende dagen verricht, dat niet door toevallige omstandigheden verklaarbaar is, en

B) of de variantie van de waarnemingsonnauwkeurigheden bij dit experiment, behoudens toevallige fluctuaties, hetzelfde is als die bij het eerste experiment (zie 4.1;6).

ad A Wij gaan uit van het model:

$$y_{ij} = \mu + \mu_i + \varepsilon_{ij} \qquad (4.3;1)$$

hetgeen wil zeggen dat de j -de waarneming op de i -de dag (i = 1,2,3; j = 1,...,7 voor i = 1,2; j = 1,...,6 voor i = 3) opgebouwd gedacht is uit een constant deel μ plus een dag-effect μ_i plus een stochastische responsiefluctuatie bij constante waarde van de factoren, ε_{ij} . Zonder verlies van algemeenheid geldt

$$\sum_i \mu_i = 0$$

en

$$E\, \varepsilon_{ij} = 0,$$

terwijl wij verder aannemen dat alle ε_{ij} onderling onafhankelijk en normaal verdeeld zijn met constante variantie σ^2. De nulhypothese

$$H_0 : \quad \mu_i = 0 \qquad \text{voor } i = 1,2,3$$

kan nu getoetst worden met behulp van de methode van de variantieanalyse. De betreffende schema's worden in dit geval:

voor de waarnemingen van de opbrengst: Tabel X
" " " " " RF : Tabel XI

Tabel X. Variantieanalyse met betrekking tot de waarnemingen van de O, verricht in O,...,O

bron van variantie	som van kwadraten	aantal vrijh. graden	gemiddelde som van kwadraten
tussen dagen	97,64	2	48,82
binnen dagen	100,73	17	5,935
totaal	198,73	19	

$F = \frac{48,82}{5,93} = 8,2$ $\quad \nu_1 = 2$, $\nu_2 = 17$ $\quad k = 0,005$

Tabel XI. Variantieanalyse met betrekking tot de waarnemingen van de RF, verricht in O,...,O

bron van variantie	som van kwadraten	aantal vrijh. graden	gemiddelde som van kwadraten
tussen dagen	0,0935	2	0,0467
binnen dagen	0,3462	17	0,0204
totaal	0,4397	19	

$F = \frac{0,0467}{0,0204} = 2,3$ $\quad \nu_1 = 2$, $\nu_2 = 17$ $\quad 0,15 > k > 0,10$

Er is dus geen reden om tot het bestaan van systematische niveauverschillen te besluiten wat de waarnemingen van de RF betreft, maar ten aanzien van de waarnemingen van de opbrengst kunnen wij de nulhypothese met grote betrouwbaarheid verwerpen en de conclusie trekken dat er tussen de verschillende dagen systematische verschillen in niveau optreden, die dus niet verklaarbaar zijn door toevallige fluctuaties tussen de waarnemingen op één dag verricht.

ad B

De twee schattingen van de variantie "binnen dagen", nl. uit het eerste experiment (4.1;6) en die uit de Tabellen X en XI kunnen vergeleken worden met behulp van een F -toets. Onder de nulhypothese, dat beide varianties gelijk zijn, heeft het quotient van beide schattingen een F -verdeling. Deze toets moet tweezijdig uitgevoerd worden. In dit geval vinden wij

1) voor O

$$F = \frac{5,935}{3,087} = 1,93, \text{ met resp. 17 en 22 vrijheidsgraden.}$$

De hierbij behorende overschrijdingskans is 0,17, zodat er geen aanwijzingen zijn, dat er een verschil is

2) voor de RF

$$F = \frac{0,0204}{0,0047} = 4,34, \text{ met resp. 17 en 22 vrijheidsgraden.}$$

De hierbij behorende overschrijdingskans is 0,002, zodat met zeer kleine onbetrouwbaarheid de nulhypothese verworpen mag worden ten gunste van de alternatieve hypothese dat de variantie "binnen dagen" bij het tweede experiment groter was dan bij het eerste. Achteraf is inderdaad gebleken, dat bij de experimentele procedure in het laboratorium enkele bronnen van onnauwkeurigheden, die normaliter wel optreden, bij het eerste experiment wel, en bij het tweede experiment niet, waren geëlimineerd.

Intussen moet de conclusie getrokken worden dat de variantie van de waarnemingsonnauwkeurigheden, die bij het zuiveringsprocédé optreden, groter is dan uit het oriënterende experiment is gebleken. Stellen wij $\mu_i = 0$ in (4.2;1) dan verkrijgen wij als schatting van σ^2 :

$$\left.\begin{array}{ll} \text{voor de opbrengst:} & s^2 = 10,44 \\ \text{voor de RF :} & s^2 = 0,0231 \end{array}\right\} \qquad (4.3;2)$$

ieder der getallen gebaseerd op 19 vrijheidsgraden. Dit wil zeggen dat in circa 50% van de gevallen tussen twee duplo-waarnemingen al of niet op dezelfde dag verricht, tussen de waarnemingen van de O verschillen $\geq$ 3,0 zullen optreden, en tussen die van de RF verschillen $\geq$ 0,14. Deze waarden steken wel zeer ongunstig af tegen die welke volgen uit (4.1;6). In par. (4.8) komen wij hierop terug.

4.4. De invloed van de factoren op de O en de RF

De lineaire effecten $b_1, \ldots, b_8$ in Tabel II komen, zoals te verwachten was, geheel overeen met die van Tabel IV, wat onderlinge verhouding en teken betreft. Een uitzondering hierop is de vijfde factor. Het grote lineaire effect van Tabel II in verhouding tot het effect b_5 in Tabel IV maakt het twijfelachtig of de chemische overwegingen, die ertoe leidden dat het centrum van de tweede proefopzet wat de vijfde factor betreft, in positieve richting verschoven werd, juist zijn geweest.

De grote lineaire effecten van de 5-de, 6-de en 7-de factor (a, z_2 en T_2) kan zeer duidelijk gedemonstreerd worden door vergelijking van de waarnemingen van die proeven, waarvoor $a=-1$, $z_2=+1$, $T_2=-1$, en die waarvoor $a=+1$, $z_2=-1$, $T_2=+1$ (Tabel XII). De invloed die uit deze waarnemingen duidelijk blijkt valt ook direct op bij het vergelijken van de waarnemingen verricht aan de proeven (MC no) 73 t/m 78.

Tabel XII. Lineaireeffecten van de 5-de, 6-de en 7-de factor

	opbrengst		RF	
	waarnemingen van proeven, waarbij:		waarnemingen van proeven, waarbij:	
	a = -1 z_2= +1 T_2= -1	a = +1 z_2= -1 T_2= +1	a = -1 z_2= +1 T_2= -1	a = +1 z_2= -1 T_2= +1
	46,9	10,0	1,65	0,98
	52,1	13,3	1,38	1,30
	57,3	13,3	1,39	0,94
	52,1	13,3	1,28	0,82
	46,9	18,8	1,30	1,06
	47,4	26,3	1,53	1,45
	37,4	22,1	1,03	1,08
	41,7	15,0	1,12	1,02
gemiddeld	47,7	16,5	1,33	1,08

Gezien het feit, dat de RF nog steeds lage - zeer lage waarden te zien gaf, hebben wij getracht uit de waarnemingen factorencombinaties af te leiden, waarvan verwacht mag worden, dat een hogere RF zal worden verkregen. Hierbij wordt evenwel gebruik gemaakt van het feit, dat de lineaire effecten van de O en de RF hetzelfde teken bezitten, zodat bij het maximaliseren van de RF verwacht mag worden, dat de O zeker niet zal dalen.

4.5. Het aangepaste oppervlak

Zoals uit Tabel II, kolom D_1, blijkt is het oppervlak, aangepast aan de waarnemingen van de RF:

$$y = 1{,}3330 - 0{,}0314\,x_1 + \cdots + 0{,}0088\,x_8 - 0{,}0149\,x_1^2 + \cdots$$
$$- 0{,}0030\,x_8^2 + 0{,}0386\,x_1 x_2 + \cdots - 0{,}0120\,x_7 x_8 \qquad (4.5;1)$$

Door de translatie:

$$\left.\begin{aligned}
x_1^* &= x_1 + 0{,}4995\\
x_2^* &= x_2 - 0{,}0089\\
x_3^* &= x_3 - 5{,}1651\\
x_4^* &= x_4 - 4{,}8562\\
x_5^* &= x_5 - 0{,}8972\\
x_6^* &= x_6 - 1{,}1694\\
x_7^* &= x_7 - 1{,}5227\\
x_8^* &= x_8 + 0{,}9132
\end{aligned}\right\} \qquad (4.5;2)$$

en de rotatie [1] (in matrixnotatie)

$$X = A\,x^* , \qquad (4.5;3)$$

waarin

$$A = \begin{pmatrix}
+0{,}16614 & -0{,}33052 & -0{,}31362 & +0{,}35784 & +0{,}47631 & -0{,}58901 & -0{,}03754 & +0{,}24805\\
-0{,}51249 & +0{,}44632 & -0{,}26209 & +0{,}09316 & -0{,}03763 & -0{,}33806 & +0{,}56127 & -0{,}17336\\
+0{,}07251 & +0{,}11740 & -0{,}53099 & +0{,}51014 & -0{,}46971 & +0{,}31339 & -0{,}10011 & +0{,}33152\\
+0{,}36233 & -0{,}00272 & +0{,}13048 & +0{,}08293 & -0{,}61491 & -0{,}56181 & -0{,}11420 & -0{,}37152\\
-0{,}10368 & +0{,}37753 & -0{,}12201 & +0{,}28236 & +0{,}32046 & +0{,}06841 & -0{,}60062 & -0{,}53292\\
-0{,}15561 & +0{,}49728 & +0{,}28136 & -0{,}13251 & -0{,}05870 & -0{,}31807 & -0{,}39061 & +0{,}61204\\
-0{,}69021 & -0{,}51679 & -0{,}14639 & -0{,}14736 & -0{,}25788 & -0{,}11752 & -0{,}36286 & -0{,}03759\\
+0{,}24871 & +0{,}14452 & -0{,}64692 & -0{,}69076 & +0{,}01824 & -0{,}07268 & -0{,}12456 & -0{,}02187
\end{pmatrix} \qquad (4.5;4)$$

gaat (4.5;1) over in:

$$\begin{aligned}
y - 1{,}3929 = &-0{,}05949\,X_1^2 - 0{,}05696\,X_2^2\\
&+0{,}03752\,X_3^2 - 0{,}02395\,X_4^2\\
&-0{,}01790\,X_5^2 + 0{,}01769\,X_6^2\\
&+0{,}00956\,X_7^2 - 0{,}00226\,X_8^2
\end{aligned} \qquad (4.5;5)$$

---

1) De berekening hiervan geschiedde op de Rekenafdeling van het Mathematisch Centrum, waarbij gebruik werd gemaakt van de electronische rekenmachine, de ARMAC.

A is uiteraard orthogonaal, zodat indien een vector X bekend is, de vector x^* gevonden wordt uit

$$x^* = A' X \qquad (4.5;6)$$

4.6. De mate van aanpassing

Of een tweedegraadsoppervlak over het geheel genomen als een redelijke benadering van het responsieoppervlak kan worden beschouwd, kan worden nagegaan door de nulhypothese te toetsen, dat alle effecten, waarvan de verwachtingen (Tabel II, 2-de kolom) alleen bestaan uit β's met drie indices, gelijk aan nul zijn. Onderstellen wij, dat de waarnemingsonnauwkeurigheden onafhankelijk en normaal verdeeld zijn met gelijke variantie σ^2, dan kan de nulhypothese getoetst worden door middel van de grootheid

$$F = \frac{64}{27} \frac{\sum b_{ijk}^2}{s^2} ,$$

waarin s^2 een schatting is van σ^2, gebaseerd op ν graden van vrijheid. Onder H_0 heeft F een F-verdeling met resp. 27 en ν graden van vrijheid. In het beschouwde geval is:

$$s^2 = 0{,}0231 \quad \text{met} \quad \nu = 19 \qquad \text{(volgens 4.3;2)}$$

dus

$$F = \frac{64}{27} \times \frac{0{,}00925981}{0{,}0231} = 0{,}95 ,$$

zodat, aangezien $F < 1$, er geen enkele reden is om aan te nemen, dat het responsieoppervlak voor de RF verschilt van een tweedegraadsoppervlak.

In vergelijking (4.5;5) wordt de RF "verklaard" door de variabelen $X_1, \ldots, X_8$. Men kan in principe van elke proef de x-coördinaten met behulp van (4.5;2) en (4.5;3) transformeren in de X-coördinaten en daarna met behulp van (4.5;5) bepalen welke waarde van de RF volgens het model bij deze proef verwacht zou kunnen worden. Uiteraard is dit vooral interessant voor die proeven, die een relatief hoge RF te zien gaven (Tabel XIII). Ook zijn er uiteraard proeven, waarbij een niet bijzonder hoge RF is waargenomen, maar bij welke volgens het model een RF $> 1{,}40$ verwacht had kunnen worden (Tabel XIV).

De afwijkingen tussen modelwaarde en waarneming moeten gezien worden als een afspiegeling van de waarnemingsonnauwkeurigheden. In Tabel XIII en XIV zijn deze verschillen inderdaad verklaarbaar, indien zij met (4.3;2) vergeleken worden (F = 1,0 met resp. 13 en 19 graden van vrijheid)

Uit Tabel XIII zien wij dat aan de factorencombinatie van proef 44 (RIV no 11) zeker aandacht geschonken dient te worden. De hoge waarde van de RF, waargenomen bij proef 74 (RIV no 84) wordt daarentegen niet geruggesteund door de modelwaarde.

Tabel XIII. Modelwaarden en waarnemingen van de RF van proeven, waarbij de waargenomen RF $\geq$ 1,45 was

proef no MC	proef no RIV	waargenomen waarde	model waarde	absolute waarde van het verschil
74	84	1,66	1,37	0,29
44	11	1,65	1,72	0,07
71	94	1,54	1,37	0,17
43	46	1,51	1,31	0,20
70	83	1,48	1,32	0,16
75	79	1,48	1,25	0,13
50	1	1,48	1,29	0,19
64	64	1,47	1,53	0,06
48	42	1,45	1,43	0,02
53	69	1,45	1,36	0,09

Tabel XIV. Modelwaarden en waarnemingen van de RF van proeven, waarbij de modelwaarde $\geq$ 1,45 is, terwijl de waargenomen RF < 1,45 was.

proef no MC	proef no RIV	waargenomen waarde	model waarde	absolute waarde van het verschil
27	68	1,39	1,50	0,11
4	35	1,39	1,48	0,09
60	59	1,31	1,45	0,14

4.7. Extrapolatie naar punten met hogere RF

Uit (4.5;5) blijkt dat verhoging van responsie te verkrijgen is door $X_1 = X_2 = X_4 = X_5 = X_8 = 0$ te houden, en X_3, X_6 en X_7 grote negatieve en positieve waarden te geven.

Teneinde na te gaan in welk deel van het X -coördinatenstelsel de proeven zijn uitgevoerd, zijn de coördinaten van proef 44 (R.I.V. no 11) omgerekend in de X -coördinaten. Dit gaf als resultaat:

$$\begin{aligned} X_1 &= +\ 0{,}0746 & X_5 &= -\ 0{,}1383 \\ X_2 &= -\ 0{,}1765 & X_6 &= +\ 1{,}5130 \\ X_3 &= +\ 3{,}1839 & X_7 &= +\ 1{,}9657 \\ X_4 &= -\ 0{,}1065 & X_8 &= +\ 7{,}1698\ , \end{aligned} \qquad (4.7;1)$$

waarbij volgens (4.5;5) een RF = 1,72 behoort.(zie Tabel XII). Inderdaad blijkt uit (4.7;1) dat X_1, X_2, X_4 en X_5 klein, en bijna = 0 zijn. X_8 heeft daarentegen een grote positieve waarde. Men kan nu consequent aan $X_8 = 0$ vasthouden bij het vaststellen van volgende proeven. Dit zou echter betekenen, dat men voor volgende proeven punten zou kiezen, die zo ver weg liggen van het door de proeven bestreken gebied, dat het model (4.5.5) te ver geëxtrapoleerd zou worden, en niet meer als zinvolle benadering van het responsieoppervlak in dit verre gebied kan gelden. Het is dus beter de eis: $X_8 = 0$ te laten vallen, indien men hierdoor bereiken kan dat men punten kan vinden met een gunstige (hogere) RF die dicht bij of in het door de proeven bestreken gebied liggen. Het verlies in RF ten gevolge van $X_8 \neq 0$ kan, vanwege de kleine coëfficiënt van X_8^2 in (4.5;5) zeer wel gecompenseerd worden door een kleine vergroting (in absolute zin) van X_3, X_6 of X_7.

Uit (4.5;5) volgt ook dat zowel positieve als negatieve waarden van X_3, X_6 en X_7 de responsie zouden verhogen. Wil de extrapolatie van het model (4.5;5) niet al te gewaagd zijn, dan dient men, zoals uit (4.7;1) volgt slechts positieve waarden van deze variabelen toe te laten.

Indien men als doel stelt een RF = 1,80 (of 2,00) te bereiken, dan heeft men nog een grote keuze van mogelijkheden voor de waarden, die X_3, X_6 en X_7 aan kunnen nemen. Zo kan o.a. van de combinatie:

$$X_1 = X_2 = X_4 = X_5 = 0$$
$$X_3 = +4{,}00$$
$$X_6 = +1{,}77$$
$$X_7 = +2{,}40 \text{ en}$$
$$X_8 = +7{,}00 ,$$

volgens (4.5;5) een RF = 2,00 verwacht worden. Uit (4.5;6) en (4.2;2) volgt dan, dat deze waarden van X overeenkomen met:

$$\begin{array}{lcl} x_1 = -0{,}40 & \text{ofwel} & z_1 = 3{,}12 \\ x_2 = +1{,}13 & & T_1 = 27{,}8^\circ \\ x_3 = -1{,}35 & & A_1 = 23' \\ x_4 = +1{,}48 & & p = 648 \text{ mg} \\ x_5 = -1{,}58 & & a = 14{,}4\,\% \\ x_6 = +1{,}07 & & z_2 = 4{,}51 \\ x_7 = -1{,}31 & & T_2 = 49{,}2^\circ \\ x_8 = +1{,}26 & & A_2 = 66' , \end{array} \qquad (4.7;2)$$

waarbij de overgang van de x_i naar de waarden van de factoren volgde uit Tabel IX.

Bij het opzetten en uitvoeren van dit type van controleproeven zijn twee punten van belang die in de volgende paragrafen besproken zullen worden.

4.8. Het aantal herhalingen van de controleproeven

In paragraaf 4.6 werd reeds aangegeven dat de factorencombinatie, die bij proef 44 (RIV no 11) is toegepast zeer zeker in aanmerking komt, nader bestudeerd te worden. Bovendien zullen behalve nieuwe waarnemingen aan deze combinatie van factoren ook proeven worden uitgevoerd van het type (4.7;2), en de uitkomsten hiervan zal men bij voorkeur vergelijken met die van proef 44 (RIV no 11) en de nog te verrichten herhalingen hiervan, om na te gaan of deze hoger zijn of niet. Op grond van de eis, dat indien er een niet zeer klein verschil is tussen de RF van beide factorencombinaties, men een redelijke kans wil hebben dit te ontdekken, kan onder gebruikmaking van reeds bekende gegevens het aantal benodigde herhalingen worden vastgesteld.

Hiertoe onderstellen wij dat de waarnemingen normaal verdeeld zijn; voorts dat de schattingen van σ^2 in (4.3;2) gebaseerd zijn op ∞ vrijheidsgraden, zodat de σ^2 exact bekend is. Wij nemen vervolgens aan, dat de te volgen procedure zal zijn het verschil van gemiddelden van de waarnemingen van de RF van proef 44 en die van (4.7;2) te vergelijken met de σ^2 (eenzijdige toets met onbetrouwbaarheidsdrempel 0,10). Uit deze gegevens is het aantal benodigde herhalingen af te leiden (zie Tabel XV).

Tabel XV. Aantal benodigde herhalingen als functie van het onderscheidingsvermogen van de toets

verschil in RF \ kans om het verschil te constateren	70%	90%
0,10	14	31
0,20	4	8

De getallen in Tabel XV geven dus aan:

1) dat proeven te verrichten om de toeneming van de RF te constateren zeker in viervoud dienen te geschieden,
2) dat een toename van 0,10 in de RF in de regel niet te constateren zal zijn en derhalve
3) dat, zoal een toename van de RF met redelijke zekerheid geconstateerd kan worden, het met de grote waarde van de variantie van de waarnemingsonnauwkeurigheid (4.3;2) onmogelijk zal zijn met een redelijk aantal experimenten een optimum anders dan zéér grof te localiseren.

4.9. Gewenste nauwkeurigheid, waarmee de factoren dienen te worden aangehouden

Een belangrijk punt is in hoeverre kan men de waarden genoemd in (4.7;2) afronden zonder de betekenis van deze proef te doorkruisen. Dit kan bijvoorbeeld nagegaan worden door er vanuit te gaan dat een proef zal worden uitgevoerd bij

$$\left.\begin{array}{ll} x_1 = 3{,}2 & a = 15{,}4\ \% \\ T_1 = 28\tfrac{1}{2}^\circ & x_2 = 4{,}4 \\ t_1 = 27' & T_2 = 49^\circ \\ p = 690\ mg & t_2 = 67', \end{array}\right\} \qquad (4.9;1)$$

en uit (4.5;1) te bepalen welke invloed de verwachte RF ondergaat indien een van deze factoren anders wordt gekozen, terwijl de overige constant gehouden worden. Daar de beschouwde veranderingen in elk van de factoren slechts zeer klein zijn, kan men ook (4.5;1) naar x_i partieel differentiëren en in de verkregen vergelijking (4.9;1) invullen. Vermenigvuldiging van deze waarde met de (in -coördinaten) voorgestelde verandering van de waarde van de factor geeft de betreffende te verwachten verandering in de responsie: Tabel XVI.

Tabel XVI. Invloed van kleine wijzigingen van waarden van de factoren

factor	verandering van de waarde van de factor	bijbehorende verandering in de waarde van de responsie	geadviseerde nauwkeurigheid
z_1	0,1	0,00	1 cijfer v/d 1e decimaal
T_1	1°	0,02	$\frac{1}{2}^o$
t_1	2'	0,04	$\frac{1}{2}$'
p	40 mg	0,05	10 mg
a	0,4%	0,06	0,05%
z_2	0,1	0,03	5 cijfers v/d 2e decimaal
T_2	1°	0,02	$\frac{1}{2}^o$
t_2	2'	0,05	$\frac{1}{4}$'

4.10. De controleproeven

Een aantal controleproeven, opgesteld op eenzelfde basis als (4.7;2) zijn uitgevoerd. De resultaten van deze proeven kwamen evenwel in genen dele overeen met de verwachtingen. Zo werd in plaats van een verwachte RF=1,80 bij een aantal proeven een gemiddelde RF = 1,33 (nr 109) en bij andere proeven in plaats van een RF = 1,90 een RF = 1,21 gevonden (nr 110). Deze discrepanties zijn zo groot, dat zij zich niet laten verklaren uit afwijkingen van het model. Kort nadat deze proeven waren uitgevoerd, werd gevonden dat er een groot verloop in het uitgangsmateriaal was opgetreden, zodat de controleproeven in wezen niet vergelijkbaar waren met de uit het tweede experiment opgestelde model.

4.11. De ammoniumsulfaatconcentratie bij de tweede uitzouting

Juist voordat de controleproeven zouden worden uitgevoerd, werd door het R.I.V. gevonden dat een zeer belangrijke factor niet in het onderzoek betrokken was geworden. Dit was nl. de concentratie van de ammoniumsulfaat bij de tweede uitzouting. Een aantal proeven werden uitgevoerd om een indruk te krijgen van de invloed van deze nieuwe factor. Daar deze niet op advies van het Mathematisch Centrum verricht zijn, zullen deze hier verder niet beschreven worden.

Bij deze proeven werd een combinatie van factoren gevonden, waarbij het gezuiverde product aan de practijkeisen qua zuiverheid en opbrengst voldeed. In dit stadium werd daarom het onderzoek afgebroken.

4.12. Conclusies

De tot op heden verrichte experimenten moeten gezien worden als de eerste fase van een onderzoek naar de optimale condities voor de zuivering van schapenantidifterieserum. Het is juist door deze experimenten duidelijk geworden dat een volledig onderzoek dieper dient te gaan. Daarbij zal vooral speciale aandacht besteed moeten worden aan:

1) het verkleinen van de variantie van de waarnemingsonnauwkeurigheid door een stapsgewijs onderzoek van de experimentele procedure, en
2) het nagaan, welke factoren van invloed zijn op het proces.

Zou het niet mogelijk blijken de variantie van de waarnemingsonnauwkeurigheden, zoals opgegeven in (4.3;2), aanzienlijk te verkleinen, dan zal het, zoals reeds in paragraaf 4.8 is besproken wel niet mogelijk zijn met een niet al te groot aantal experimenten een optimum anders dan zeer grof te localiseren. Wij vestigen er de nadruk op dat dit het geval zal zijn, onafhankelijk van het feit of men een meer verfijnde experimenteertechniek toepast, dan wel de tot op heden gebruikelijke "one factor at a time"-methode.

5. Zuivering van paardenantidifterieserum

5.1. Inleiding

Het onderzoek, waarbij de waarnemingen zijn verricht, had als doel na te gaan of de factorencombinatie, waarbij het zuiveringsproces van paardenantidifterieserum geschiedt, optimaal is en zo niet, welke factorencombinatie een hogere RF geeft bij eenzelfde opbrengst.

In tegenstelling tot het schapenantidifterieserum was het paardenantidifterieserum dus reeds in productie. De betreffende chemici meenden wel, dat indien het proces niet optimaal zou blijken te zijn, de optimale condities slechts weinig zouden verschillen van de waarden van de factoren, die bij de productie worden gebruikt.

Dit wil dus ook zeggen, dat er weinig hoop is reeds iets te kunnen bereiken door het uitvoeren van een klein experiment, waarbij uit de te verkrijgen waarnemingen de coëfficiënten van een lineair model te schatten zouden zijn. Immers in de omgeving van een optimum zal men zeker rekening moeten houden met de kromming van het responsieoppervlak. Wil het mogelijk zijn aan de waarnemingen een tweedegraadsvergelijking aan te passen, dan moet het uit te voeren experiment uiteraard van grotere omvang zijn. Door ons werd dus geadviseerd de proefopzet uit te voeren, die reeds in hoofdstuk 2 beschreven is.

Als centrum voor deze proefopzet werd de combinatie van factoren gekozen, zoals deze toegepast worden in het productieproces. De niveaux van de factoren zijn in Tabel XVII vermeld.

Tabel XVII. Niveaux van de factoren

		Centrum	eenheid	$-\sqrt{8}$	-1	$+1$	$+\sqrt{8}$
p_H	z_1	3,2	0,2	2,6	3,0	3,4	3,8
temp.	T_1	25°	2,5°	18°	22½°	27½°	32°
tijd	t_1	30'	5'	16'	25'	35'	44'
pepsine	p	400 mg	100 mg	117 mg	300 mg	500 mg	683 mg
am.sulf.	a	18%	1%	15,17%	17%	19%	20,83%
p_H	z_2	4,3	0,2	3,7	4,1	4,5	4,9
temp.	T_2	55°	2,5°	48°	52½°	57½°	62°
tijd	t_2	60'	5'	46'	55'	65'	74'

De verkregen waarnemingen zijn vermeld in Tabel I, kolom D_2.

5.2. De proeven verricht in O,...,O

Zoals uit Tabel I blijkt zijn er twintig proeven verricht in O,...,O, zodat uit de waarnemingen hierbij verricht een waarde van de O en RF te verkrijgen zijn, die een indruk geven van de zuivering en het rendement van het productieproces. Als gemiddelde van de twintig waarnemingen vinden wij:

$$\begin{aligned} O &: 51{,}81\% \\ RF &: 1{,}961 \end{aligned} \qquad (5.2;1)$$

Evenals bij schapenantidifterieserum kan ook nu worden nagegaan of er systematische verschillen optreden tussen de waarnemingen, die gedaan zijn op ieder van de drie dagen, welke nodig waren voor de uitvoering van het experiment. De wijze waarop dit kan worden onderzocht is vermeld in 4.3. De variantie-analyseschema's worden:

voor de O : Tabel XVIII

voor de RF: Tabel XIX

Tabel XVIII. Variantieanalyse met betrekking tot waarnemingen van de O, verricht in O,...,O

bron van variantie	som van kwadraten	aantal vrijheids-graden	gemiddelde som van kwadraten
tussen dagen	28,86	2	14,43
binnen dagen	77,00	17	4,53
totaal	105,86	19	

$$\left.\begin{aligned} F &= \frac{14{,}43}{4{,}53} = 3{,}2 \\ \nu_1 &= 2,\ \nu_2 = 17 \end{aligned}\right\} \rightarrow k = 0{,}07.$$

In geen van beide gevallen kan de nulhypothese verworpen worden: er is dus geen reden tot het bestaan van systematische niveauverschillen tussen de verschillende dagen te besluiten.

De schatting van de variantie van de waarnemingsonnauwkeurigheid (indien in (4.3;1) $\mu_i = 0$ gesteld wordt) is:

$$\left.\begin{aligned} O: &\quad s^2 = 5{,}572 \quad \nu = 19 \\ RF: &\quad s^2 = 0{,}0152 \quad \nu = 19 \end{aligned}\right\} \qquad (5.2;2)$$

hetgeen een lagere getalswaarde is dan de overeenkomstige schattingen (4.3;2) voor schapenantidifterieserum. Evenwel kan niet bewezen worden dat de schattingen (5.2;2) een kleinere variantie vertegenwoordigen dan (4.3;2) ($F = 1{,}9$ en $1{,}5$, bijbehorende overschrijdingskansen $k = 0{,}10$, resp. $0{,}25$).

Tabel XIX. Variantieanalyse met betrekking tot de waarnemingen van de RF, verricht in O,...,O

bron van variantie	som van kwadraten	aantal vrij-heidsgraden	gemiddelde som van kwadraten
tussen dagen	0,0471	2	0,0235
binnen dagen	0,2415	17	0,0142
totaal	0,2886	19	

$$\left.\begin{array}{l} F = \dfrac{0{,}0235}{0{,}0142} = 1{,}65 \\ \nu_1 = 2, \quad \nu_2 = 17 \end{array}\right\} \longrightarrow \quad k = 0{,}20 \text{ (ongeveer)}$$

5.3. De invloed van de factoren

De schattingen van de effecten van de factoren zijn vermeld in Tabel II, kolom D_2.

Een zeer typisch verschil met de schattingen van de effecten van kolom D_1 (schapenantidifterieserum) is dat de lineaire effecten, voor zover zij van enige betekenis zijn, bij de O en de RF tegengestelde tekens bezitten. Bovendien zijn voor de belangrijkste factoren nl. 5, 6 en 7 de verhoudingen van de schatting van O tot die van de RF van dezelfde orde van grootte, nl. 23, 14 en 13. Dit wijst er dus zeker op, dat indien wij het maximum van de RF willen trachten te vinden bij constante O, dit maximum zeker niet veel hoger zal liggen dan (5.2;1). Hierbij hebben wij even afgezien van de invloed van de interacties, die evenwel toch een kleinere invloed hebben dan de lineaire effecten van factoren 5, 6 en 7.

5.4. Het aangepaste oppervlak

Zoals uit Tabel II, kolom D_2 blijkt, is de vergelijking van het aangepaste oppervlak voor de waarnemingen van de opbrengst, uitgedrukt in de coördinaten, waarin de factorencombinaties in Tabel I ook zijn aangegeven:

$$\begin{aligned} O - 51{,}81 = & -0{,}6037\,x_1 + \cdots - 0{,}0963\,x_8 - 0{,}7702\,x_1^2 + \cdots - 0{,}0375\,x_8^2 \\ & - 0{,}1578\,x_1 x_2 + \cdots + 0{,}3109\,x_7 x_8 . \end{aligned} \qquad (5.4;1)$$

Aan de RF zal een lineaire vergelijking aangepast worden. Hoewel de mate van kromming van het responsieoppervlak, zoals deze in de coëfficiënten b_{ij} tot uitdrukking komt, voor O en RF weinig verschillen, is toch de lineaire benadering gebruikt en wel om de volgende redenen.

1) Wij willen de RF omhoog trachten te brengen. Het blijkt nu dat in het voor ons interessante gebied van de factorenruimte de RF volgens het tweedegraadsmodel altijd hoger is dan de verwachte waarde, die uit het lineaire deel van dit tweedegraadsmodel zou volgen. Heeft een punt dus een gunstige (hoge) RF volgens het lineaire model, dan zal dit zeker ook zo zijn volgens het tweedegraadsoppervlak, dat uit de waarnemingen als beste aanpassing kan worden berekend.

2) Indien ook aan de RF een tweedegraadsvergelijking wordt aangepast, wordt de interpretatie hiervan, tezamen met(5.4;1), zeer moeilijk. Er bestaat nl. in het algemeen geen orthogonale transformatie zodanig dat beide tweedegraadsvergelijkingen reduceren tot een som van χ_i^2 .

De vergelijking van het aangepaste oppervlak is nu voor de RF:

$$RF - 1{,}961 = +0{,}0034\,x_1 + \cdots + 0{,}0119\,x_8 \,. \qquad (5.4;2)$$

Door middel van de translatie

$$\left.\begin{aligned} x_1^* &= x_1 - 2{,}1728 \\ x_2^* &= x_2 + 2{,}3544 \\ x_3^* &= x_3 - 10{,}1464 \\ x_4^* &= x_4 - 3{,}1512 \\ x_5^* &= x_5 - 7{,}4899 \\ x_6^* &= x_6 - 10{,}4302 \\ x_7^* &= x_7 + 8{,}0306 \\ x_8^* &= x_8 - 2{,}0021 \end{aligned}\right\} \qquad (5.4;3)$$

en de rotatie

$$X = B\,x^* \,, \qquad (5.4;4)$$

waarin X' de vector$(X_1, \cdots, X_8)$en $x^{*\prime}$ de vector$(x_1^*, \cdots, x_8^*)$ is en B wordt gegeven door:

$$B = \begin{pmatrix} +0{,}74704 & -0{,}14664 & +0{,}07388 & -0{,}52220 & -0{,}26802 & +0{,}15058 & +0{,}12258 & +0{,}18097 \\ +0{,}54287 & +0{,}29205 & -0{,}18698 & +0{,}20289 & +0{,}39123 & -0{,}53086 & -0{,}31104 & -0{,}11070 \\ -0{,}18055 & +0{,}26672 & +0{,}29182 & -0{,}35640 & +0{,}52832 & -0{,}16871 & +0{,}28683 & +0{,}54242 \\ +0{,}22424 & +0{,}00709 & -0{,}44711 & +0{,}35012 & +0{,}36367 & +0{,}47956 & +0{,}51036 & +0{,}06881 \\ +0{,}19599 & +0{,}22598 & +0{,}59358 & +0{,}48199 & -0{,}28092 & -0{,}16399 & +0{,}46368 & -0{,}07110 \\ -0{,}03390 & -0{,}08363 & +0{,}11994 & -0{,}38109 & +0{,}30772 & -0{,}07522 & +0{,}33693 & -0{,}78637 \\ +0{,}00859 & +0{,}79701 & +0{,}07986 & -0{,}14233 & -0{,}06680 & +0{,}49359 & -0{,}23986 & -0{,}18010 \\ +0{,}15721 & -0{,}35878 & +0{,}54901 & +0{,}19680 & +0{,}43088 & +0{,}40110 & -0{,}39929 & -0{,}02108 \end{pmatrix}$$

gaat (5.4;1) over in:

$$O = 62{,}0516 - 0{,}97305\,X_1^2 - 0{,}86539\,X_2^2 + 0{,}65977\,X_3^2 + 0{,}41349\,X_4^2$$
$$- 0{,}35585\,X_5^2 - 0{,}32829\,X_6^2 - 0{,}19675\,X_7^2 - 0{,}01533\,X_8^2 . \qquad (5.4;5)$$

(5.4;2) wordt na de translatie (5.4;3) en de rotatie (5.4;4)

$$RF = -0{,}274 - 0{,}02489\,X_1 + 0{,}08592\,X_2 + 0{,}12494\,X_3 + 0{,}04016\,X_4 +$$
$$+ 0{,}03159\,X_5 + 0{,}08822\,X_6 - 0{,}14063\,X_7 - 0{,}08325\,X_8 . \qquad (5.4;6)$$

B is uiteraard orthogonaal. Indien waarden voor $X_1, \ldots, X_8$ gegeven zijn, vindt men dus $x_1^*, \ldots, x_8^*$ uit:

$$x^* = B' X , \qquad (5.4;7)$$

en met behulp van (5.4;3) vindt men nu de bijbehorende x_i's .

5.5. Extrapolatie naar punten met hogere RF

Uit de vergelijkingen (5.4;5) en (5.4;6) blijkt dat een absoluut optimum niet bestaat in de onmiddellijke omgeving van het door de experimenten bestreken gebied. Men kan nl. onder de voorwaarde dat de opbrengst = 51,81, elke waarde van de RF verkrijgen door bijv. X_8 groter negatief en X_3 groter positief te maken. Het is dus zinvol [illegible] onderzoeken van welke waarde van x_i verwacht mag worden, da[illegible] de hiermee overeenkomstige factorencombinatie een O = 51,8 en een RF > 1,96 zal geven.

De coördinaten van het centrum van de proefopzet zijn:

$X_1^0 = -0{,}014$	$X_5^0 = +0{,}245$
$X_2^0 = +1{,}096$	$X_6^0 = +2{,}621$
$X_3^0 = -1{,}797$	$X_7^0 = -4{,}718$
$X_4^0 = -0{,}798$	$X_8^0 = -17{,}952$

Onder de waarnemingen in Tabel I bevinden zich enkele, waarbij de opbrengst niet veel afwijkt van 52%, terwijl de RF hoger dan 1,96 is. Voor deze proeven zijn in Tabel XX opgenomen: de waargenomen waarden en de waarden, die men volgens het model 5.4;1 en 5.4;2 (ofwel 5.4;6 en 5.4;7) zou kunnen verwachten. In de laatste kolom is bovendien vermeld de verwachte waarde van de RF, indien deze wordt berekend uit de aangepaste tweedegraads vergelijking.

Tabel XX. Enkele waargenomen waarden en hun verwachtingen volgens het model

Proef nr.		waargenomen waarden		waarden volgens model (5.4;6) en (5.4;7)		verwachte waarde van RF volgens het tweedegraadsmodel
MC	RIV	O	RF	O	RF	RF
15	3	51,7	2,21	50,6	1,86	1,95
56	52	58,1	2,13	60,7	1,87	1,95
20	66	55,6	2,14	52,3	1,87	1,95
59	34	50,5	2,36	50,4	2,17	2,24

Alleen proef 34 heeft dus volgens het model een opbrengst, die weinig afwijkt van 51,8, maar die een RF bezit, die duidelijk hoger is dan 1,96. De X -coördinaten van deze proef zijn:

$$\begin{aligned} X_1 &= -1,15 & X_5 &= +0,69 \\ X_2 &= +0,59 & X_6 &= +3,13 \\ X_3 &= -3,15 & X_7 &= -6,23 \\ X_4 &= -0,63 & X_8 &= -19,31 . \end{aligned} \qquad (5.5;1)$$

Teneinde punten met nog hogere RF te vinden, zal men enkele van de X_i moeten variëren. Hiervoor komen in aanmerking:

1) X_8, omdat een negatieve waarde van X_8 de RF vergroot, maar door zéér kleine coëfficiënt van X_8^2 in (5) vermindert de opbrengst slechts weinig.

2) X_3 en X_4 : door de positieve coëfficiënt van X_3^2 en X_4^2 in (5.4;6) kan een kleine wijziging van X_3 en X_4 gemakkelijk de vermindering in opbrengst ten gevolge van het variëren van X_8 te niet doen. Bovendien kunnen X_3 en X_4 er ook voor zorgen, dat de RF groter wordt.

Getracht is tezamen met X_1^o, X_2^o, X_5^o, X_6^o en X_7^o een combinatie van X_3, X_4 en X_8 te vinden, die:

1) een RF = 2,30 geeft volgens (5.4;7)
2) een O = 51,81 " " (5.4;6)
3) waarvoor $(X_3 - X_3^o)^2 + (X_4 - X_4^o)^2 + (X_5 - X_5^o)^2$ minimaal is, hetgeen wil zeggen, dat wij de factoren zo weinig mogelijk willen veranderen.

De oplossing hiervoor is

$$X_3 = +\ 1{,}10$$
$$X_4 = -\ 2{,}28$$
$$X_8 = -17{,}76,$$

hetgeen kan worden weergegeven in de x -coördinaten als

$x_1 = -\ 0{,}72$	dus	$x_1 = 3{,}2 - 0{,}72 \times 0{,}2 = 3{,}06$		
$x_2 = +\ 0{,}61$		$T_1 = 25 + 0{,}61 \times 2{,}5 = 26\frac{1}{2}^{\circ}$		
$x_3 = +\ 2{,}06$		$t_1 = 30' + 2{,}06 \times 5' = 40'$		
$x_4 = -\ 1{,}13$		$p = 400 - 1{,}13 \times 100 = 287$ mg		(5.5;3)
$x_5 = +\ 0{,}83$		$a = 18 + 0{,}83 \times 1 = 18{,}83\%$		
$x_6 = -\ 1{,}33$		$x_2 = 4{,}3 - 1{,}33 \times 0{,}2 = 4{,}03$		
$x_7 = +\ 0{,}44$		$T_2 = 55 + 0{,}44 \times 2{,}5 = 56^{\circ}$		
$x_8 = +\ 1{,}49$		$t_2 = 60' + 1{,}49 \times 5 = 67\frac{1}{2}'$ (67)		

5.6. De proeven, verricht ter bevestiging van de conclusies van 5.5

Een klein experiment is uitgevoerd, teneinde te trachten de conclusies beschreven in 5.5 te bevestigen. Dit experiment omvatte:

1) een proef in viervoud bij 0,...,0
2) een proef in viervoud bij de factorencombinatie 59 (RIV no 34)
3) een proef in viervoud bij de factorencombinatie (5.5;3).

Tengevolge van een misverstand werd daarbij in (5.5;3) $T_2 = 53\frac{1}{2}$ in plaats van 56° genomen.

De resultaten zijn vermeld in Tabel XXI.

Tabel XXI: waarnemingen verricht aan de bevestigingsproeven

factoren-combinatie	0,...,0		MC 59 (RIV 34)		(5.5;3) ($T_2 = 53\frac{1}{2}$)	
	Opbrengst %	RF	Opbrengst %	RF	Opbrengst %	RF
	54,6	1,84	48,9	1,94	48,9	2,11
	54,6	1,84	51,1	1,85	47,6	2,04
	51,1	1,73	51,1	1,85	47,6	2,15
	53,6	1,63	52,7	1,76	46,3	2,15
gemiddeld	53,5	1,76	51,0	1,85	47,6	2,11

De volgende conclusies kunnen uit de waarnemingen worden getrokken.

1) Van de waarnemingen in O,...,O zal uiteraard verwacht worden dat behalve toevallige afwijkingen deze waarnemingen overeenkomen met die van proef nr 81-100 van Tabel I. Deze overeenstemming kan worden getoetst met behulp van de toets van Student (De waarnemingen worden hierbij normaal verdeeld ondersteld). Dit geeft voor de opbrengst (met behulp van (5.2;2)):

$$t = \frac{53{,}5 - 51{,}8}{s\sqrt{\frac{1}{20} + \frac{1}{4}}} = 1{,}3 \qquad \nu = 19 \qquad \longrightarrow \qquad k = 0{,}20$$

en voor de RF:

$$t = \frac{1{,}96 - 1{,}76}{s\sqrt{\frac{1}{20} + \frac{1}{4}}} = 9 \qquad \nu = 19 \qquad \longrightarrow \qquad k \ll 10^{-4}.$$

Ten aanzien van beide groepen van waarnemingen van de opbrengst kan niet worden besloten tot het bestaan van andere dan zuiver toevallige afwijkingen. Ten aanzien van de RF kan daarentegen met aan zekerheid grenzende waarschijnlijkheid gezegd worden dat er een duidelijk verloop is opgetreden in het uitgangsmateriaal gedurende de maanden, die lagen tussen de uitvoering van het grote experiment en van de zojuist beschreven proeven. Het effect van het verloop kon in het laboratorium onderzocht worden en bleek van de orde van grootte van 5% (in twee maanden).

2) Corrigeren wij de gemiddelde waarde, gevonden bij de herhalingen van proef 59 (RIV 34), voor het onder 1 beschreven verloop door bij de RF 5% op te tellen, dan vinden wij

$$\left.\begin{array}{l} O = 51{,}0 \\ RF = \ \ 1{,}94 \end{array}\right\} \qquad (5.6;1)$$

waaruit de conclusie getrokken moet worden, dat de waarde (5.6;1) niet te rijmen is (wat RF betreft) met de modelwaarde, opgegeven in Tabel XX. Vergelijken wij de RF van (5.6;1) met die van proef 59 (RIV 34) dan vinden wij

$$t = \frac{2{,}36 - 1{,}94}{s\sqrt{\frac{1}{4} + 1}} = 3{,}04 \qquad \nu = 19 \qquad \longrightarrow \qquad k = 0{,}008,$$

zodat er een aanwijzing is dat ook hier - na correctie - , een systematische afwijking optreedt. Wij komen hierop nog terug.

3) Ten aanzien van de gecorrigeerde waarden van RF en O van proef 101 (O: 47,6, RF: 2,31) merken wij slechts op dat de RF wel de gewenste gunstige waarde heeft, maar dat dit ten koste is gegaan van een ontoelaatbaar grote verlaging van de O.

In verband met het verloop in het uitgangsmateriaal zijn dezelfde proeven nogmaals uitgevoerd met een verse charge serum. Bovendien zijn dezelfde proeven ook nog eens uitgevoerd met de oude charge van het serum, maar met de correcte waarde van T_2 in (5.5;3). De waarnemingen zijn vermeld in Tabel XXII.

Tabel XXII. Waarnemingen verricht aan proeven betreffende vers en oud plasma

factoren-combinatie	O,...,O		MC 59 (RIV 34)		(5.5;3)(T_2=53½)	
	Opbrengst %	R.F.	Opbrengst %	R.F.	Opbrengst %	R.F.
	41,8	1,80	45,4	1,94	40,5	2,08
oud	51,6	1,90	43,4	2,07	43,4	2,13
plasma	54,2	1,91	47,5	2,00	39,5	2,00
	50,0	1,87	46,8	1,98	41,8	2,14
gemiddeld	49,4	1,87	45,8	2,00	41,3	2,09
	56,2	2,11	54,0	2,36	46,6	2,51
vers	54,0	2,13	54,0	2,25	48,8	2,38
plasma	48,8	2,18	51,4	2,16	46,4	2,53
	56,2	2,18	51,8	2,23	50,1	2,48
gemiddeld	53,8	2,15	52,8	2,25	47,9	2,43

Uit Tabel XXII (en XXI) kunnen de volgende conclusies worden getrokken:

1) De toename van de opbrengst, resp. RF gevonden bij de herhalingen van proef 59 (RIV 34) ten opzichte van de proeven in O,...,O is

	O	RF
in Tabel XXI:	- 3,6	+ 0,13
" " XXII, oud plasma:	- 1,0	+ 0,10
" " XXII, vers plasma:	- 2,5	+ 0,09

De overeenstemming tussen deze drie paren getallen is zeer goed te noemen. De conclusie is dan ook dat de factorencombinatie waarbij proef 59 (RIV no 34) is uitgevoerd niet beschouwd kan worden als een combinatie, die gunstiger is dan de huidige condities. Dit geldt in nog sterkere mate voor de factorencombinatie (5.5;3), waarbij de toename gevonden ten opzichte van de huidige condities is:

	O	RF
oud plasma	- 8,1	+ 0,21
vers plasma	- 5,9	+ 0,28

2) De waarnemingen in Tabel XXII zijn gevonden met behulp van een voor verloop gecorrigeerde titer. Het is daarom zinvol de hypothese te toetsen dat de waarden van de RF van de proeven verricht in O,...,O in Tabel XXII niet verschillen van die in Tabel I. De toetsingsgrootheid is voor deze toets

$$A = \frac{1,96 - 1,87}{s\sqrt{\frac{1}{4} + \frac{1}{20}}} = 1,34,$$

met 19 vrijheidsgraden. De hierbij behorende waarde van de overschrijdingskans is 0,20, zodat er geen enkele reden is, aan te nemen, dat de correctie voor het verloop niet goed is geweest.

5.7. Conclusies

Hoewel volgens de vergelijking van het aangepaste oppervlak bij gelijkblijvende O een hogere RF dan bij de huidige productie wordt verkregen mogelijk zou moeten zijn is het niet gelukt een dergelijke factorencombinatie ook werkelijk te vinden.

Evenals bij het schapenantidifterieserum speelt de zeer grote variantie van de waarnemingsonnauwkeurigheid bij deze discrepantie tussen theoretische modelwaarden en de practijkwaarnemingen waarschijnlijk een zeer grote rol. Een voortzetting van dit onderzoek zou slechts dan zinvol zijn indien deze variantie aanzienlijk zou kunnen worden verlaagd. Dat dit onderzoek toch tot zover tot uitvoering is gekomen, vindt zijn oorzaak in het feit, dat het experiment, waarbij de honderd waarnemingen werden verricht uitgevoerd is tijdens de statistische analyse van het schapenantidifterieserum-experiment, zodat toen nog niets bekend was over het feit, dat de verwachte variantie der waarnemingsonnauwkeurigheid een onderschatting was van de werkelijkheid.

6. Paardenantitetanusserum

6.1. Inleiding

Het onderzoek, waarbij de waarnemingen zijn verricht, had als doel na te gaan of de factorencombinatie, waaronder het zuiveringsproces van paardenantitetanusserum verloopt, optimaal is en zo niet, welke factorencombinaties een hogere opbrengst geven bij eenzelfde reinigingsfactor (RF).

Evenals dit het geval was bij het paardenantidifterieserum, meende men dat indien de zojuist genoemde factorencombinatie niet optimaal zou blijken te zijn, de optimale combinatie slechts weinig hiervan zou verschillen. Op grond hiervan is, zoals in 5.1 is uiteengezet, het juister te beginnen met een experiment dat omvangrijker is dan in 4.1 voor schapendifterieserum is beschreven. Ons advies was dan ook de proefopzet uit te voeren, zoals deze beschreven is in hoofdstuk 3.

Als centrum van deze proefopzet werd weer de factorencombinatie gekozen, die toegepast wordt in het productieproces. De niveaux zijn in Tabel XXIII vermeld.

Tabel XXIII. Niveaux van de factoren

		Centrum	eenheid	$-\sqrt{8}$	- 1	+ 1	$+\sqrt{8}$
p_H	z_1	3,2	0,2	2,6	3,0	3,4	3,8
temp.	T_1	25°	2,5°	18°	22½°	27½°	32°
tijd	t_1	30'	5'	16'	25'	35'	44'
pepsine	p	400 mg	100 mg	117 mg	300 mg	500 mg	683 mg
am.sulf.	a	16%	1%	13,17%	15%	17%	18,83%
p_H	z_2	4,3	0,2	3,7	4,1	4,5	4,9
temp.	T_2	55	2,5	48	52½	57½	62
tijd	t_2	60'	5'	46'	55'	65'	74'

De waarnemingen zijn vermeld in kolom D_3 van Tabel I.

6.2. Waarnemingen verricht in 0,...,0

Evenals bij de andere sera, zijn ook hier 20 proeven verricht in het centrum der proefopzet, en wel zijn van deze 20 proeven een aantal uitgevoerd op elk der drie dagen, die ook hier nodig waren voor de uitvoering van het experiment. Aan de hand van de waarnemingen, die bij deze proeven zijn verricht, kan als in 4.3 en 5.2 worden nagegaan of er systematische niveauverschillen optreden tussen de bepalingen verricht op ieder der drie dagen.

De wijze waarop dit kan worden onderzocht is beschreven in 4.3. De variantieanalyse-schema's zijn voor de waarnemingen

van de opbrengst: Tabel XXIV,
van de RF : Tabel XXV.

Tabel XXIV. Variantieanalyse met betrekking tot waarnemingen van O verricht in O,...,O

bron van variantie	som van kwadraten	aantal vrij-heidsgraden	gemiddelde som van kwadraten
tussen dagen	20,62	2	10,31
binnen dagen	33,82	17	1,99
totaal	54,44	19	

$$\left.\begin{array}{l} F = 5{,}2 \\ \nu_1 = 2, \quad \nu_2 = 17 \end{array}\right\} \longrightarrow \quad k = 0{,}02$$

Tabel XXV. Variantieanalyse met betrekking tot waarnemingen van de RF verricht in O,...,O

bron van variantie	som van kwadraten	aantal vrij-heidsgraden	gemiddelde som van kwadraten
tussen dagen	0,0413	2	0,0206
binnen dagen	0,0976	17	0,00574
totaal	0,1389	19	

$$\left.\begin{array}{l} F = 3{,}6 \\ \nu_1 = 2, \quad \nu_2 = 17 \end{array}\right\} \longrightarrow \quad k = 0{,}04$$

Uit Tabel XXIV en XXV blijkt dus dat er een zeer duidelijke aanwijzing bestaat dat er niveauverschillen opgetreden zijn. De gemiddelde waarnemingen op ieder der dagen zijn weergegeven in Tabel XXVI.

Tabel XXVI. Gemiddelde van de waarnemingen verricht in O,...,O

	O	RF
eerste dag	51,7	1,78
tweede dag	53,6	1,87
derde dag	51,3	1,88

Voor de opbrengst is de tweede dag gemiddeld een hoge waarde verkregen, voor de reinigingsfactor de eerste dag een lage waarde, die in dit geval de oorzaak zijn van de kleine overschrijdingskansen in Tabel XXIV en XXV.

Bij de aanpassing van de tweedegraadsvergelijking, die in 6.3 ter sprake zal komen, is geen rekening gehouden met het optreden van de dagverschillen. Op de consequenties hiervan komen wij in 6.5 terug.

Ziet men af van de dagverschillen, dan volgt uit de waarnemingen verricht in centrum als schatting van de O en RF die verkregen worden bij het productieproces:

$$\left.\begin{aligned} O &= 52{,}22 \\ RF &= 1{,}845 \end{aligned}\right\} \qquad (6.2;1)$$

Als schatting van de variantie der waarnemingsonnauwkeurigheden vinden wij, eveneens afgezien van de dagverschillen:

$$\left.\begin{aligned} &\text{voor de O:} \quad s^2 = 2{,}865 \\ &\text{voor de RF:} \quad s^2 = 0{,}00731 \end{aligned}\right\} \qquad (6.2;2)$$

Bij toetsing wordt gevonden dat er een duidelijke aanwijzing is dat deze varianties kleiner zijn dan bij schapenantidifterieserum

$$\begin{aligned} F &= 3{,}64 \text{ resp. } 3{,}16 \\ k &= 0{,}01 \text{ resp. } 0{,}02 \end{aligned} \qquad (5.2;2)$$

maar er is geen reden aan te nemen, dat (6.2;2) verschilt van de waarden van (5.2;2) voor paardenantidifterieserum (F = 1,94 resp. 2,07, k = 0,10 resp. 0,08).

6.3. Het aangepaste oppervlak

De schattingen van de effecten van de factoren zijn vermeld in Tabel II, kolom D_3.

Uitgedrukt in het coördinatensysteem, waarop de factorencombinaties van de proeven ook reeds betrokken waren, kunnen de opbrengst en de RF dus voorgesteld worden door de volgende tweedegraadsvergelijkingen:

$$\begin{aligned} O - 52{,}22 = &- 0{,}8187\,x_1 + \cdots + 0{,}0800\,x_8 - 0{,}4605\,x_1^2 + \cdots + \\ &+ 0{,}3207\,x_8^2 - 0{,}6221\,x_1 x_2 + \cdots - 0{,}8909\,x_7 x_8 \end{aligned} \qquad (6.3;1)$$

en

$$\begin{aligned} RF - 1{,}845 = &- 0{,}0058\,x_1 + \cdots + 0{,}0101\,x_8 + 0{,}0121\,x_1^2 + \cdots + \\ &+ 0{,}0064\,x_8^2 + 0{,}0106\,x_1 x_2 + \cdots + 0{,}0025\,x_7 x_8 . \end{aligned} \qquad (6.3;2)$$

Bij de analyse van de overeenkomstige hiervoor beschreven experimenten kon steeds nog de onderstelling gemaakt worden, dat of de opbrengst of de reinigingsfactor voor te stellen was door een <u>lineaire</u> vergelijking in de $x_1, \cdots, x_8$. In dit geval blijkt het bij de berekeningen van essentieel belang te zijn beide responsies door een <u>tweedegraads</u> model voor te stellen. Hierdoor wordt de interpretatie van (6.3;1) en (6.3;2) veel moeilijker; er bestaat, zoals reeds gezegd is, geen orthogonale transformatie, zodanig dat (6.3;1) en (6.3;2) beide kunnen worden geschreven als een som van termen in de X_i^2 (i = 1,...,8).

Om deze reden is alleen vergelijking (6.3;1) getransformeerd op de hoofdassen. Dit is ook al geschied, omdat wij bij een constante RF de O willen maximaliseren. Het blijkt dan nl. iets gunstiger te zijn (6.3;1) te transformeren. Daarbij doet zich de moeilijkheid voor dat de matrix van (6.3;1) singulier blijkt te zijn. Om deze reden is de transformatie op de hoofdassen van (6.3;1) slechts gedeeltelijk uitgevoerd, nl. alleen de rotatie. Door middel van de rotatie:

$$X = C x \qquad (6.3;3)$$

waarin $X' = (X_1, \cdots, X_8)$, $x' = (x_1, \cdots, x_8)$ en

$$C = \begin{pmatrix} -.5069 & -.3625 & +.1654 & +.2836 & +.0431 & -.3130 & +.5265 & +.3561 \\ -.0280 & -.0023 & -.2921 & +.2251 & +.3153 & -.0931 & +.3871 & -.7780 \\ -.0551 & +.3167 & -.3793 & +.2284 & -.8097 & -.0461 & +.2068 & -.0102 \\ -.2093 & +.7785 & +.0244 & -.2280 & +.3230 & +.1363 & +.3546 & +.2212 \\ -.6990 & +.0020 & +.0233 & +.2280 & -.0192 & +.5482 & -.3593 & -.1698 \\ +.3192 & -.0367 & -.4423 & +.5741 & +.3104 & +.3279 & +.0044 & +.4095 \\ -.0635 & +.3864 & +.1724 & +.4858 & +.1291 & -.5954 & -.4573 & -.0270 \\ -.3176 & -.1078 & -.7190 & -.3837 & +.1594 & -.3320 & -.2572 & +.1470 \end{pmatrix} \qquad (6.3;4)$$

gaat (6.3;1) over in:

$$\begin{aligned} O - 52,22 = & -1,3085\, X_1^2 - 0,6199\, X_1 \\ & +0,7253\, X_2^2 - 1,1578\, X_2 \\ & -0,5226\, X_3^2 + 0,4044\, X_3 \\ & +0,4990\, X_4^2 - 0,7819\, X_4 \\ & -0,3371\, X_5^2 + 2,4374\, X_5 \\ & +0,1971\, X_6^2 + 0,3001\, X_6 \\ & -0,1352\, X_7^2 - 0,1342\, X_7 \\ & - 0,9219\, X_8 . \end{aligned} \qquad (6.3;5)$$

Toepassing van de rotatie (6.3;3) op vergelijking (6.3;2) geeft evenwel opnieuw een niet homogene tweedegraadsvorm in $X_1, \ldots, X_8$:

$$RF - 1{,}845 = c_1 X_1 + \cdots + c_8 X_8 + c_{11} X_1^2 + \cdots + c_{88} X_8^2 + \\ + c_{12} X_1 X_2 + \cdots + c_{78} X_7 X_8 . \qquad (6.3;6)$$

Het zal in verband met de aard van het onderzoek niet nodig blijken te zijn alle coëfficiënten c_{ij} $(i \neq j)$ te berekenen. De andere coëfficiënten zijn:

$c_1 = +0{,}1414$	$c_{11} = +0{,}0351$
$c_2 = +0{,}0950$	$c_{22} = +0{,}0140$
$c_3 = -0{,}0136$	$c_{33} = +0{,}0203$
$c_4 = +0{,}0649$	$c_{44} = -0{,}0026$
$c_5 = -0{,}1429$	$c_{55} = +0{,}0033$
$c_6 = -0{,}0229$	$c_{66} = +0{,}0071$
$c_7 = +0{,}0175$	$c_{77} = +0{,}0152$
$c_8 = +0{,}0319$	$c_{88} = -0{,}0091$.

6.4. Extrapolatie naar punten met hogere O

Allereerst zullen wij onderzoeken welke gevolgen het heeft, indien een bepaalde $X \neq o$ gekozen wordt, terwijl alle andere = 0 gehouden worden.

a) Uit (6.3;5) blijkt dat elke waarde van $X_1 \neq o$ een verlaging geeft van de opbrengst beneden de waarde 52,22. Indien men een hoge opbrengst wil krijgen, zal X_1 bij voorkeur = 0 gekozen worden.

b) Voor $X_2 >$ 0 zal de RF groter worden dan 1,845. De opbrengst zal eveneens groter worden dan 52,22 als $X_2 > 1{,}6$ of als $X_2 < 0$.

c) $X_3 \neq$ 0 zal altijd een verlaging van de opbrengst betekenen. Men zal bij voorkeur X_3 = 0 kiezen.

d) De opbrengst wordt groter dan 52,22 als $X_4 < 0$ of als $X_4 > 1{,}5$, de RF neemt toe als $X_4 > 0$.

e) Indien de waarde van X_5 niet al te groot gekozen wordt, dan draagt $X_5 >$ 0 sterk bij tot verhoging van de opbrengst, maar de RF wordt tegelijkertijd lager. Bij $X_5 < 0$ is dit precies omgekeerd.

f) De variabelen X_6 en X_7 hebben in de vergelijkingen (6.3;5) en (6.3;6) kleine coëfficiënten. Zo zij al een zodanige waarde aan kunnen nemen, dat èn O èn RF zullen stijgen, zal deze stijging gering zijn.

g) Negatieve waarden voor X_8 dragen bij tot verhoging van de opbrengst, terwijl zij tevens de RF verlagen. Waarden $>$ O leiden tot een toename van O en daling van de RF.

Uit deze beschouwingen blijkt dus dat slechts de variabele X_2 en X_4 mee kunnen helpen èn de opbrengst èn de RF-factor te verhogen. Deze verhoging van de RF kan dan worden gecompenseerd door een vermindering ten gevolge van een variabele, die door het feit, dat deze variabele $\neq$ O gekozen kan worden, de opbrengst aanzienlijk zal verhogen. Voor dit laatste komt vooral X_5 in aanmerking. Wij stellen voor de andere variabelen = O te kiezen.

Na berekening van de coëfficiënten c_{25}, c_{24} en c_{45} in (6.3;6) wordt (6.3;6) voor $X_1 = X_3 = X_6 = X_7 = X_8 = 0$:

$$\begin{aligned} RF - 1{,}845 = {} & 0{,}0950\, X_2 + 0{,}0140\, X_2^2 \\ & + 0{,}0649\, X_4 - 0{,}0026\, X_4^2 \\ & - 0{,}1429\, X_5 + 0{,}0033\, X_5^2 \\ & - 0{,}0301\, X_2 X_4 - 0{,}0105\, X_2 X_5 \\ & + 0{,}0144\, X_5 X_4 , \end{aligned} \qquad (6.4;1)$$

terwijl (6.3;5) wordt:

$$\begin{aligned} O - 52{,}22 = {} & + 0{,}7253\, X_2^2 - 1{,}5178\, X_2 \\ & + 0{,}4990\, X_4^2 - 0{,}7819\, X_4 \\ & - 0{,}3371\, X_5^2 + 2{,}4374\, X_5 . \end{aligned} \qquad (6.4;2)$$

Men kan zich nu afvragen, welke waarden voor X_2, X_4 en X_5 gekozen dienen te worden, zodanig dat

$$RF = 1{,}845$$
$$O = 52{,}22 + 5{,}00$$

en waarbij

$$X_2^2 + X_4^2 + X_5^2 = \text{minimaal} \qquad (6.4;3)$$

Dit is dus het punt met gelijke RF en een opbrengst 5,00 hoger dan in het centrum, welk punt tevens zo dicht mogelijk bij het centrum ligt.

Het minimum van (6.4;3) is een zeer vlak minimum, wij vinden als benaderde oplossing:

$$X_2 = +1{,}80 \qquad X_4 = -1{,}70 \qquad X_5 = +1{,}00 \;, \qquad (6.4;4)$$

waarbij dus tevens $X_1 = X_3 = X_6 = X_7 = X_8 = 0$. De waarde van het minimum (6.4;4) is hier dus 7,1.

Uit (6.3;3) volgt dan, dat van het punt gekenmerkt door de coördinaten:

x_1	= - 0,39	ofwel	z_1 = 3,2 - 0,39 x 0,2	= 3,12
x_2	= - 1,31		T_1 = 25 - 1,31 x 2,5	= 21,7°
x_3	= - 0,54		t_1 = 30 - 0,54 x 5	= 27,3'
x_4	= + 1,02		a = 400 + 1,02 x 100	= 502 mg (6.4;5)
x_5	= 0,00		b = 16 - 0,00 x 1%	= 16%
x_6	= + 0,15		z_2 = 4,3 + 0,15 x 0,2	= 4,33
x_7	= - 0,27		T_2 = 55 - 0,27 x 2,5	= - 0,7
x_8	= + 1,95		t_2 = 60' + 1,95 x 5'	= 70'

men mag verwachten een gelijke RF en een 5,00 hogere opbrengst te zullen waarnemen als in het centrum van de oorspronkelijk uitgevoerde proefopzet.

Men kan eenvoudig nagaan bij welke van de verrichte proeven een RF is waargenomen die ongeveer 1,84 was en een opbrengst, hoger dan 52%. De resultaten van deze proeven worden in Tabel XXVII vergeleken met de modelwaarden, die men verkrijgt door in (6.3;1) en (6.3;2) de factorencombinaties van de betreffende proeven in te vullen. Hierbij blijkt dat het model zeer redelijk de gevonden waarden verklaart. De toename in opbrengst ten opzichte van de huidige condities is evenwel klein.

De proeven 33, 38 en 46 (RIV no's: 2, 6 en 43) (onderste regels, Tabel XXVII) zijn proeven waarbij de RF aanzienlijk hoger is dan die van de huidige condities, terwijl de opbrengst ongeveer gelijk was.

Vergelijken wij de overeenkomstige modelwaarden, dan blijkt dat een vrij aanzienlijke verbetering in de RF mogelijk is ten koste van een geringe verlaging van de opbrengst. De vraag rijst dan of het ook mogelijk is, een verbetering in de RF te verkrijgen, zonder een vermindering van opbrengst. Dit blijkt inderdaad (volgens het model) mogelijk te zijn. Vult men in (6.3;5) en (6.3;6) in: $X_1 = 0$, $X_2 = +2{,}00$, $X_3 = \cdots = X_8 = 0$, dan bepalen deze waarden van X de factorencombinatie:

$$
\begin{array}{lll}
x_1 = -0{,}05 & \text{ofwel} & z_1 = 3{,}2 - 0{,}05 \times 0{,}2 = 3{,}19 \\
x_2 = +0{,}00 & & T_1 = 25 + 0{,}00 \times 2{,}5 = 25{,}0^\circ \\
x_3 = -0{,}58 & & t_1 = 30 - 0{,}58 \times 5' = 27{,}1' \\
x_4 = +0{,}45 & & p = 400 + 0{,}45 \times 100 = 445 \text{ mg} \\
x_5 = +0{,}63 & & a = 16 + 0{,}63 \times 1\,\% = 16{,}63\,\% \qquad (6.4;6) \\
x_6 = -0{,}18 & & z_2 = 4{,}3 - 0{,}18 \times 0{,}2 = 4{,}26 \\
x_7 = +0{,}77 & & T_2 = 55^\circ + 0{,}77 \times 2{,}5 = 57^\circ \\
x_8 = -1{,}55 & & t_2 = 60 - 1{,}55 \times 5' = 52'
\end{array}
$$

waarbij men een gelijke opbrengst en een verhoging in RF van 0,25 ten opzichte van de huidige condities kan verwachten. Reeds bij de berekeningen, die leidden tot het opstellen van (6.4;4) werd het duidelijk, dat het eenvoudiger zal zijn de RF te verhogen bij gelijkblijvende O, dan omgekeerd.

Tabel XXVII. Enkele waarnemingen en de betreffende modelwaarden

Proef no.		waargenomen waarden		model-waarden	
MC	RIV	RF	O	RF	O
50	1	1,86	58,5	1,82	55,1
43	46	1,86	55,8	1,81	53,4
56	52	1,86	54,3	1,71	55,5
20	66	1,85	57,1	1,82	56,4
33	2	2,38	52,5	2,25	50,0
38	6	2,16	56,0	1,92	56,2
46	43	2,16	52,3	2,26	49,6

6.5. Invloed van de dagverschillen

Teneinde na te gaan welke invloed de dagverschillen, die in 6.2 besproken zijn, hebben, is bij alle waarnemingen van de RF die de eerste dag verricht zijn (t.w RIV no's 1-40) 0,10 opgeteld, en van alle waarnemingen van de opbrengst, die de tweede dag zijn verricht (RIV no's 40-79) 2,1 afgetrokken. Met deze gecorrigeerde waarden van de waarnemingen zijn nogmaals de effecten van Tabel II berekend. Zodoende kan men een vergelijking van de tweede graad aanpassen aan de waarnemingen van de O en de RF (analoog aan (6.3;1) en (6.3;2)). In deze vergelijkingen is daarna de combinatie (6.4;6) ingevuld, hetgeen als verwachte waarde van de toename van de opbrengst ten opzichte van de opbrengst in het centrum van de proefopzet, gaf: + 0,6. De analoge waarde van de RF is + 0,24. Op grond hiervan mogen wij zeggen, dat het in vrijwel aselecte volgorde uitvoeren van de proeven de invloed van de dagverschillen zo goed als geëlimineerd heeft.

6.6. Proeven uitgevoerd ter bevestiging van de conclusies van 6.4 en 6.5

Enkele proeven werden uitgevoerd teneinde na te gaan in hoeverre de voorspelde toenames in RF en O zich ook inderdaad zouden voordoen. Deze proeven omvatten:

1) een proef in viervoud bij de factorencombinatie van het huidige productieproces,
2) een proef in viervoud bij (6.4;5); proef 101,
3) " " " " bij (6.4;6): " 102,

De waarnemingen zijn vermeld in Tabel XXVIII.

Tabel XXVIII. Waarnemingen verricht bij de proeven, uitgevoerd ter bevestiging van de conclusies van 6.4 en 6.5

factoren-combinatie	Opbrengst %	R.F.	Opbrengst %	R.F.	Opbrengst %	R.F.
oud plasma mengsel	53,6	1,94	55,6	2,04	51,6	2,30
	55,6	1,96	55,6	1,87	51,6	2,25
	53,0	1,93	51,6	1,91	51,6	2,30
	52,2	1,96	53,6	1,92	50,6	2,14
gemiddeld	53,6	1,95	54,1	1,93	51,4	2,25
nieuw plasma	53,6	1,90	53,6	1,92	54,0	2,39
	55,8	2,01	55,8	1,86	52,6	2,36
	55,8	2,00	55,8	1,96	51,5	2,40
	53,6	1,90	55,8	1,86	53,6	2,35
gemiddeld	54,7	1,95	55,3	1,90	52,9	2,37

Ten aanzien van Tabel XXVIII merken wij op:

1) Een verhoging van de opbrengst is bij de waarnemingen van proef 101 ten opzichte van de proeven verricht in O,...,O niet opgetreden. Zoals reeds eerder is opgemerkt, is het vinden van een combinatie van factoren, waarvan men een hogere opbrengst mag verwachten, zeer moeilijk; het is dan ook niet verrassend, dat deze verhoging zich niet realiseerde. Dit staat uiteraard ook in verband met de onzekerheid in de modelvergelijkingen ten gevolge van de grote variantie van de waarnemingsonnauwkeurigheden.

2) Een zeer duidelijke verbetering in RF is verkregen bij proef 102, evenwel ten koste van een geringe verlaging van de opbrengst. Het product van de waarde van RF en O is bij proef 102: 129, 124, 124 en 126, terwijl dit bij de proef verricht in O,...,O is: 101, 112, 116, 102. In deze situatie moeten wij reeds met de toets van Wilcoxon de hypothese, dat er geen verschil zou zijn, verwerpen (k = 0,03).

Een schatting van de verbetering van de RF, behorende bij de condities (6.4;6) is: + 0,36 (voorspelde waarde: + 0,25), een schatting van de vermindering aan O, is daarbij + 2,0 (voorspelde waarde - 0,0).

3) Alvorens een beslissing te nemen ten aanzien van het toepassen van (6.4;6) bij het productieproces adviseren wij nog een klein experiment uit te voeren, waarbij de O en de RF behorende bij de condities van proef 33 (RIV no 2) vergeleken kunnen worden met die, behorende bij (6.4;6). Zoals uit Tabel XXVII blijkt, behoort bij proef 33 (RIV no 2) een grote verbetering in RF (waargenomen waarde + 0,43, modelwaarde + 0,30) en een geringe vermindering van de opbrengst (waargenomen 0,0, modelwaarde - 2,0). Het kan dus van belang zijn, de beide factorencombinaties te vergelijken ten aanzien van de daarbij te verkrijgen O en RF.

4) Wij zouden tevens willen voorstellen bij het experiment, waartoe wij onder 3) adviseren nog enkele factorencombinaties te onderzoeken. De condities (6.4;6) zijn te verkrijgen door in (6.4;1) en (6.4;2) $X_1 = X_3 = \cdots = X_8 = 0$ en X_2 = + 2,00 te stellen. Het zou zeker zeer nuttig zijn, na te gaan of bij hogere waarde van X_2, de volgens (6.4;1) en (6.4;2) optredende verhoging van de opbrengst en van de RF zich inderdaad ook realiseren.

6.7. Conclusies

Bij het zuiveringsproces voor paardenantitetanusserum is een factorencombinatie gevonden die bij een geringe verlaging van de opbrengst een in verhouding grote toename van de reinigingsfactor bewerkstelligt.

Een advies is gegeven om de mogelijkheid van een nog verder gaande verbetering van de RF te onderzoeken. Hierbij behoeft tevens geen verlies van opbrengst verwacht te worden.

Ook in dit geval is het niet zinvol bij de huidige experimentele procedure het onderzoek voort te zetten. De reden hiervoor - t.w. de grote variantie van de waarnemingsonnauwkeurigheden - is reeds op vele plaatsen ter sprake gekomen. Wij verwijzen hiervoor bijvoorbeeld naar 4.12 en 5.7.

Literatuur

[1] BLOEMENA, A.R., en Ph. VAN ELTEREN: Proefopzet voor het opsporen van optimale condities bij serumzuivering (eerste fase). Rapport S 198 van de Statistische Afdeling van het Mathematisch Centrum, Mei 1956.

[2] BLOEMENA, A.R., en Ph. VAN ELTEREN: Verwerking en Interpretatie van de waarnemingsuitkomsten, verkregen bij het uitvoeren van de proefopzet beschreven in S 198, Rapport S 198A van de Statistische Afdeling van het Mathematisch Centrum, Mei 1956.

[3] BLOEMENA, A.R., Samenvatting van de resultaten van de statistische analyse van de waarnemingen betreffende de zuivering van paardendifterieserum. Rapport S 1956-62(1) van de Statistische Afdeling van het Mathematisch Centrum, December 1956.

[4] BLOEMENA, A.R., Samenvatting van de resultaten van de statistische analyse van de waarnemingen betreffende de zuivering van paardentetanusserum. Rapport S 1957-604(1) van de Statistische Afdeling van het Mathematisch Centrum, Maart 1957.

[5] BLOEMENA, A.R., Experimentele bepaling van optimale condities. Rapport S 204 (OV 6) van de Statistische Afdeling van het Mathematisch Centrum.

[6] BOX, G.E.P., en K.B. WILSON: On the experimental attainment of optimum conditions. J.R.S.S., 1951, B 13, 1 - 45.

[7] BOX, G.E.P.: Multifactor designs of first order, Biometrika, 1952, 39, 49.

[8] BOX, G.E.P., L.R. CONNOR, W.R. COUSINS, O.L. DAVIES, F.R. HINSWORTH, G.P. SILLITO: Design and analysis of experiments, Oliver and Boyd, Edinburough, 1953.

[9] BOX, G.E.P., Bepalen van optimale condities door experimenten op bedrijfsschaal, Lezing gehouden voor de Conference on Statistical Quality Control, Parijs, Juli 1955.

[10] BOX, G.E.P., The exploration and exploitation of response surfaces, Biometrics, 1954, 10, 16 - 61.

[11] BOX, G.E.P., en P.V. YOULE: The exploration and exploitation of response surfaces: An example of the link between the fitted surface and the basic mechanism of the system, Biometrics, 1955, no 3, 11, 287 - 323.

[12] BOX, G.E.P. and J.S. HUNTER, Multi-factor experimental designs, Technical Report No 6, Mimeo Series no 92, Institute of Statistics, University of North-Carolina.

[13] DAVIES, O.L., en W.A. HAY: The construction and use of fractional factorial designs in industrial research Biometrics, 1950, 6, 233 - 249.

[14] READ, D.R.: The design of chemical experiments, Biometrics, 1954, 10, 1 - 16.

[15] VAN IJZEREN, J.:De theoretische zijde van de methode der kleinste kwadraten, Statistica, 1954, 8.1, 21 - 47.