Project 505.0420
Ontwikkeling methoden van onderzoek voor het aantonen en bepalen van mycotoxinen en marine toxinen
Projectleider: ir L.G.M.Th. Tuinstra

Rapport 91.02 Januari 1991

Model voor het berekenen van ijklijnen

J.M.P. van Trijp en A.H. Roos

Medewerkers: dr W.G. de Ruig en ir A.A.M. Jansen
(Groep Landbouwwiskunde, DLO)

Goedgekeurd door: ir L.G.M.Th. Tuinstra

Rijks-Kwaliteitsinstituut voor land- en tuinbouwprodukten (RIKILT)
Bornsesteeg 45, 6708 PD Wageningen
Postbus 230, 6700 AE Wageningen
Telefoon 08370-75400
Telex 75180 RIKIL
Telefax 08370-17717

VERZENDLIJST

INTERN:
directeur
sectorhoofden
afdelingshoofden onderzoekafdelingen
projectleider (ir L.G.M.Th. Tuinstra)
programmabeheer en informatieverzorging (2x)
circulatie
afdeling Organische Contaminanten (5x)
bibliotheek

EXTERN:
Dienst Landbouwkundig Onderzoek
Directie Wetenschap en Technologie
Directie Voedings- en Kwaliteitsaangelegenheden
ir A.A.M. Jansen, Groep Landbouwwiskunde, DLO
Produktschap voor Veevoeder
Comité van Graanhandelaren

ABSTRACT

Model voor het berekenen van ijklijnen

Regression model for computing of a calibration series

Report 91.02   January 1991

J.M.P. van Trijp and A.H. Roos

State Institute for Quality Control of Agricultural Products
P.O. Box 230, Wageningen, the Netherlands

1 table, 1 figure, 1 annexe, 1 reference

Linear regression computing of a calibration series by means of an equal weight least squares method is not allowed, when a linearity plot model is being used to identify calibration point outliers.
It is shown that a weighted linear regression forced through the origin, may be a proper approach for computing of a calibration series.
If this weighted model applies, the required computation is reduced to means and standard errors of ratios.

Keywords: linear regression, linearity, calibration, ratios

INHOUD Blz

SAMENVATTING

Bij de bepaling van het aflatoxine B_1 gehalte werd met behulp van een standaardreeks een ijklijn geconstrueerd volgens de kleinste kwadraten methode met toepassing van een wegingsfactor g=1. Een lineariteitsplot wordt gebruikt om te selecteren welke ijkpunten van de standaardreeks worden opgenomen in de ijklijn. Deze twee procedures gaan uit van een verschillend model voor de meetfouten.
Het toepassen van de lineariteitsplot voor elk ijkpunt met een wegingsfactor is gelijkwaardig aan het uitvoeren van een gewogen kleinste kwadratenmethode met een wegingsfactor $g=1/x^2$.
De kleinste kwadratenmethode is getoetst met twaalf aflatoxine standaardreeksen en gaf bij toepassing van lineaire regressie met een wegingsfactor (g=1) en met een wegingsfactor ($g=1/x^2$), een vergelijkbare betrouwbaarheid voor de geschatte richtingscoëfficiënt (b) van de ijklijn. Een grotere betrouwbaarheid van de richtingscoëfficiënt wordt verkregen door de ijklijn door de oorsprong te berekenen.

1 INLEIDING

Bij de kwantificering van het aflatoxine B_1 gehalte wordt een vijfpunts ijklijn toegepast. De berekening van deze ijklijn gebeurt door middel van lineaire regressie volgens de kleinste kwadraten methode.
Elk ijkpunt dat in de ijklijn opgenomen wordt, is eerst getest op voldoende lineariteit. Dit gebeurt met een lineariteitsplot. Hierbij wordt nagegaan of de respons/massa ratio van het ijkpunt, niet meer dan $\pm$ 10% afwijkt van de gemiddelde respons/massa ratio van de vijf ijkpunten. Van deze ijkpunten mag maximaal een ijkpunt vervallen.
Bij het berekenen van monsters, hetzij via de ijklijn, hetzij via afzonderlijke standaarden, kwamen niet te verklaren verschillen in berekend gehalte aan het licht.
Naar aanleiding van de verschillen is nagegaan in hoeverre de kriteria voor het toepassen van de ijklijn aangepast moesten worden. Omdat dezelfde problematiek ook binnen andere afdelingen zal spelen is e.e.a. vastgelegd in een rapport.

2 REKENMODELLEN

De volgende punten zijn nader uitgewerkt.
- Vaststellen selectiekriterium voor het opnemen van een ijkpunt in de ijklijn.
- Vaststellen welke wiskundige vergelijking optimaal de ijklijn beschrijft.
- Opstellen protocol voor de berekening van de ijklijn.

2.1 Lineaire regressie (wegingsfactor g=1)

2.1.1 Lineariteitsplot

Een lineariteitsplot is een grafische constructie van de ijklijn zodanig dat snel op linear gedrag getoetst kan worden (zie fig. 1).
De lineariteitsplot wordt als volgt geconstrueerd.

x-as : massa component geïnjecteerd (pg)

y-as : relatieve verhouding tussen de respons/massa ($= \%\ y_i/x_i$)

De gemiddelde respons van de ijkpunten wordt op 100% gesteld. De afzonderlijke ijkpunten worden hieraan getoetst.

% y_i/x_i voor elke i berekent men als volgt:

$$\left[\frac{y_i/x_i}{\left[\frac{\sum_{i=1}^{N} (y_i/x_i)}{N} \right]} \right] * 100\% = \%\ y_i/x_i$$

y_i = respons ijkpunt (i = 1 .... N)
x_i = massa ijkpunt (pg) (i = 1 .... N)
N = totaal aantal ijkpunten

Bij een perfecte lineariteit zal elk ijkpunt % y_i/x_i = 100% geven. In de praktijk zijn er meestal afwijkingen. Het is in de loop der tijd een goede aanname gebleken om bij aflatoxine B_1 analyses een tolerantie van ± 10% toe te staan. Daarom wordt een verband als lineair geaccepteerd als de afwijking niet groter is dan 100 ± 10%.
Een ijkpunt dat buiten deze band valt, wordt niet meegenomen in de berekening van de vergelijking van de ijklijn. Valt er meer dan één ijkpunt buiten de band (d.w.z. de ijklijn zou geconstrueerd worden uit minder dan vier punten), dan moet een nieuwe ijkreeks geanalyseerd worden. In tabel 1 is aan de hand van een reële ijklijn een rekenvoorbeeld uitgewerkt:

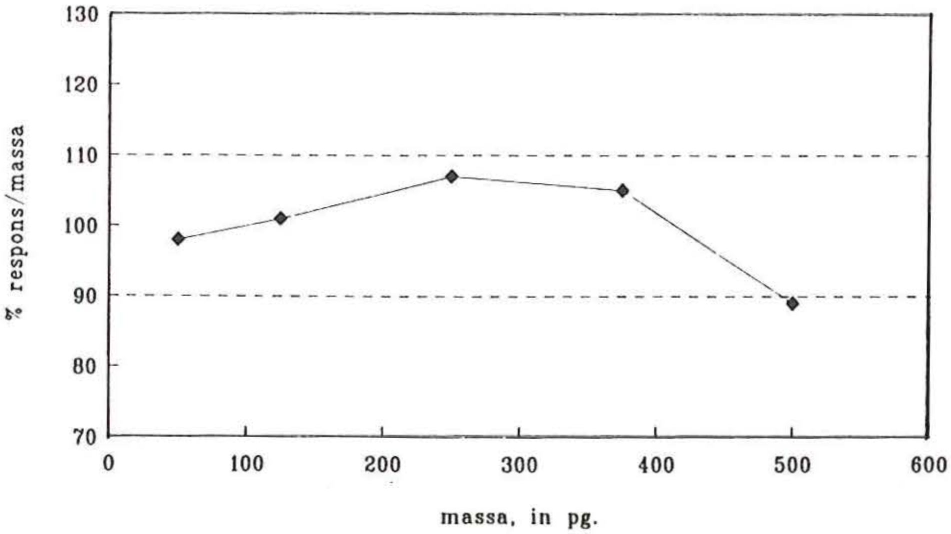
Een ijkpunt bestaat uit de variabelen x en y:

x = massa (pg)
y = respons (hoogte- of oppervlak eenheden)

Tabel 1 Berekening lineariteitsplot

	x (pg)	y (mm)	y/x	% y/x (= $\frac{y/x}{\text{gemiddelde } y/x}$ x 100%)
1	50	22	0,440	98
2	125	57	0,456	101
3	250	120	0,480	107
4	375	177	0,472	105
5	500	200	0,400	89
Gemiddeld			0,4496	100

Grafisch uitgezet levert de lineariteitsplot figuur 1 op.



Figuur 1 Lineariteitsplot

Uit de lineariteitsplot is snel op te maken dat het ijkpunt van 500 pg buiten de band valt, (% y/x = 89). Dit punt wordt dus verworpen en niet meer toegepast voor het berekenen van de ijklijn. De overige punten voldoen wel, zodat uit deze punten de vergelijking van de ijklijn berekend kan worden.

2.1.2 Berekenen van de ijklijn

De data uit 2.1.1 worden gebruikt voor het berekenen van de ijklijn. De ijklijn wordt voorgesteld door y = b x + a
waarin

y = respons	a = afsnede y-as
x = massa	b = richtingscoëfficiënt van de ijklijn

Standaard wordt de ijklijn berekend volgens de methode der kleinste kwadraten.

$$b = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N} (x_i - \bar{x})^2} = \frac{\sum_{i=1}^{N} x_i y_i - (\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i)/N}{\sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)^2/N}$$

en

$$a = \bar{y} - b\bar{x} = (\sum_{i=1}^{N} y_i)/N - (b \sum_{i=1}^{N} x_i)/N$$

Na invoeren van de punten 1 t/m 4 resulteert dit in de vergelijking

$$y = 0,48\ x - 2,00 \qquad (1)$$

Vergelijking (1) wordt gebruikt voor het berekenen van onbekende monsters met bekende respons. Bij de methode der kleinste kwadraten zoals hier toegepast, wordt aan elke respons y een evengroot gewicht toegekend (g=1). Dan hebben punten met een massa, die veel van het gemiddelde afwijkt een relatief grote invloed op de richtingcoëfficiënt (= helling); in het geval van regressie door de oorsprong (a=0) krijgen de punten met een grote massa relatief veel gewicht.

De lineariteitsplot gaat uit van een lijn door de oorsprong (a=0) en kent aan de respons y echter een wegingsfactor

$g = \frac{1}{x^2}$ toe.

Met andere woorden, de invloed van ijkpunten met grote massa is dus omgekeerd evenredig met het kwadraat van de massa. Dit betekent dat ijklijnberekening volgens lineaire regressie met a=0 en met een wegingsfactor g=1 onverenigbaar is met gebruik van de lineariteitsplot.

De lineaire regressie, die de vergelijking van de ijklijn berekent, moet daarom vervangen worden door gewogen lineaire regressie door de oorsprong als het tenminste redelijk is te veronderstellen, dat de standaardafwijking van de meetfouten in y evenredig toeneemt met x.

2.2 Gewogen lineaire regressie (wegingsfactor $g=1/x^2$)

2.2.1 Enkele modelberekeningen

In bijlage A zijn twaalf ijklijnen van de aflatoxine B_1 bepaling uitgerekend. Bij elke ijklijn is van tevoren elk ijkpunt getoetst door middel van de lineariteitsplot, punten die niet aan het lineariteitskriterium voldoen zijn geschrapt.

Met het statistische pakket SPSS (1) zijn per ijklijn vier regressieberekeningen uitgevoerd.

lineaire regressie g = 1

1. y = bx + a
2. y = bx

gewogen lineaire regressie, $g = \frac{1}{x^2}$

3. y = bx + a
4. y = bx

Het programma heeft per ijklijn en per case, de vergelijking van de lijn en een "standard error of estimate"(S_b) bepaald. Bij vergelijking tussen de gevonden standaardfouten valt op, dat S_b kleiner is indien de grafiek (vergelijking van de lijn) door de oorsprong gaat (a=0). Voorts is de betrouwbaarheid van b bij beide keuze van de gewichten vergelijkbaar.

Een grotere betrouwbaarheid van de richtingscoëfficiënt wordt verkregen door de grafiek door de oorsprong te berekenen.
Als restrictie moet aangetekend worden dat dit slechts geldt voor het geteste lineaire gebied; m.a.w. monsters met een respons buiten de ijklijn moeten verdund worden totdat de respons binnen de ijklijn valt.

2.2.2 Berekenen van de ijklijn

In de eerstegraads vergelijking

$$y = bx \qquad (g = \frac{1}{x^2})$$

wordt de richtingscoëfficiënt b op eenvoudige wijze als volgt berekend:

$$b = \frac{\sum_{i=1}^{N} y_i/x_i}{N}$$

Bereken van elk ijkpunt de richtingscoëfficiënt b, sommeer deze en deel de verkregen waarde door het aantal gebruikte ijkpunten N. In het geval van tabel 1/figuur 1, geldt N = 4.

Aangezien b nu bekend is, kan men uit een meetwaarde y van een onbekend monster de bijbehorende waarde van x berekenen:

$$x = \frac{y}{b} ,$$

x is de massa van de meetcomponent in het onbekende monster.
De berekening voor de ijklijn uit tabel 1/figuur 1, verloopt identiek met de berekening in tabel 1, tot en met de kolom y/x; de richtingscoëfficiënt b is dus 0,4496.

De vergelijking van de lijn is nu:

$$y = 0{,}4496 \ x \quad \text{of} \quad x = \frac{y}{0{,}4496} \qquad (2)$$

3 CONCLUSIE/AANBEVELING

Bij de bepaling van het aflatoxine B_1 gehalte werd tot nu toe met behulp van een standaardreeks een ijklijn geconstrueerd volgens de standaardprocedure voor lineaire regressie. Een lineariteitsplot wordt gebruikt om vooraf te selecteren welke ijkpunten van de standaardreeks worden opgenomen in de ijklijn. Deze twee procedures gaan uit van een verschillend model voor de waarnemingen.
Het toepassen van de lineariteitsplot veronderstelt, dat de ijklijn door de oorsprong gaat en dat elk ijkpunt een gewicht $g=1/x^2$ heeft. Op een dergelijke wijze getoetste ijkpunten kunnen derhalve, door middel van gewogen lineaire regressie $g=1/x^2$ gebruikt worden voor het schatten van een eerste graads vergelijking zonder constante, in de vorm $y = bx$.

Het protocol voor de ijklijnberekening luidt als volgt:

- stel een lineariteitsplot op
- toets de lineariteit (90-110%)
- bereken de richtingscoëfficiënt b uit $y = bx$ volgens

$$b = \frac{\sum_{i=1}^{N} y_i/x_i}{N}$$

- bereken in onbekende monsters x volgens

$$x = \frac{y}{b}$$

Het wordt aangeraden deze methode als standaardrekenwijze binnen het RIKILT toe te passen, indien voor het berekenen van gehalten met behulp van een ijklijn een lineariteitsplot als redelijk geacht selectiecriterium wordt gehanteerd.

LITERATUUR

1 Anonymus, SPSS/PC+ V.2.0 Base Manual for the IBM PC/XT/AT and PS/2 SPSS Inc, 444 N Michican Ave., Chigago Il. 60611

BIJLAGE A IJklijn berekeningen volgens lineaire regressie (g=1) resp. gewogen lineaire regressie ($g=1/x^2$)

ijklijn	y (mm)	x (mg/kg)	g ($1/x^2$)	lineaire regressie (g=1) y = bx + a	y = bx	gewogen lineaire regressie ($g=1/x^2$) y = bx + a	y = bx
1	119	0,005	40000	y = 22060x + 17	y = 23160x	y = 23110x + 5	y = 23596x
1	240	0,010	10000				
1	365	0,015	4444	S_b= 1470	S_b= 566	S_b= 1660	S_b= 462
1	445	0,020	2500				
2	48	0,002	250000	y = 22297x + 12	y = 23105x	y = 23520x + 2	y = 23850x
2	125	0,005	40000				
2	240	0,010	10000	S_b= 840	S_b= 463	S_b= 1129	S_b= 444
2	360	0,015	4444				
2	445	0,020	2500				
3	44	0,002	250000	y = 21550x + 3	y = 21735x	y = 21811x + 1	y = 21983x
3	115	0,005	40000				
3	215	0,010	10000	S_b= 240	S_b= 125	S_b= 298	S_b= 267
3	325	0,015	4444				
3	435	0,020	2500				
4	58	0,002	250000	y = 28999x - 2	y = 28834x	y = 28649x + 1	y = 28767x
4	144	0,005	40000				
4	285	0,010	10000	S_b= 421	S_b= 200	S_b= 467	S_b= 158
4	425	0,015	4444				
4	584	0,020	2500				
5	44	0,002	250000	y = 22949x - 3	y = 22692x	y = 22763x - 2	y = 22367x
5	109	0,005	40000				
5	230	0,010	10000	S_b= 339	S_b= 172	S_b= 375	S_b= 281
5	340	0,015	4444				
6	91	0,005	40000	y = 19120x - 1	y = 19053x	y = 19559x - 6	y = 18950x
6	196	0,010	10000				
6	285	0,015	4444	S_b= 444	S_b= 149	S_b= 545	S_b= 287
6	380	0,020	2500				

BIJLAGE A (vervolg) IJklijn berekeningen volgens lineaire regressie (g=1) resp. gewogen lineaire regressie ($g=1/x^2$)

ijklijn	y (mm)	x (mg/kg)	g ($1/x^2$)	lineaire regressie (g=1) y = bx + a	y = bx	gewogen lineaire regressie ($g=1/x^2$) y = bx + a	y = bx
7	49	0,010	10000	y = 4979x + 11	y = 5135x	y = 5331x - 4	y = 5201x
7	134	0,025	1600				
7	265	0,050	400	S_b= 405	S_b= 193	S_b= 460	S_b= 158
7	424	0,075	178				
7	479	0,100	100				
8	61	0,010	10000	y = 5065x + 15	y = 5278x	y = 5213x + 10	y = 5566x
8	146	0,025	1600				
8	267	0,050	400	S_b= 108	S_b= 83	S_b= 143	S_b= 175
8	405	0,075	178				
8	515	0,100	100				
9	44	0,002	250000	y = 21948x + 11	y = 22716x	y = 23563x - 2	y = 23167x
9	122	0,005	40000				
9	235	0,010	10000	S_b= 1477	S_b= 722	S_b= 1776	S_b= 632
9	368	0,015	4444				
9	428	0,020	2500				
10	36	0,010	10000	y = 3315x + 9	y = 3442x	y = 3497x + 2	y = 3553x
10	93	0,025	1600				
10	169	0,050	400	S_b= 310	S_b= 148	S_b= 330	S_b= 119
10	290	0,075	178				
10	320	0,100	100				
11	31	0,002	250000	y = 17184x - 2	y = 16983x	y = 17414x - 4	y = 16625x
11	85	0,005	40000				
11	170	0,010	10000	S_b= 133	S_b= 92	S_b= 210	S_b= 375
11	255	0,015	4444				
12	49	0,002	250000	y = 24064x - 1	y = 23963x	y = 23717x + 2	y = 24003x
12	121	0,005	40000				
12	234	0,010	10000	S_b= 362	S_b= 170	S_b= 414	S_b= 203
12	355	0,015	4444				
12	485	0,020	2500				