

RAPPORT OVER EEN STATISTISCH ONDERZOEK NAAR HET VERBAND TUSSEN HET VOLUME VAN BROOD EN ENIGE PHYSISCHE EN CHEMISCHE EIGENSCHAPPEN VAN HET DAARVAN GEBRUIKTE MEEL? verricht door het Mathematisch Centrum te Amsterdam in opdracht van de Stoommeelfabriek "Holland" te Amsterdam.

1. Doel van het onderzoek.

Het onderzoek had ten doel, na te gaan, of het volume van brood (de z.g. "bakaard") indirect kan worden bepaald uit een aantal fysische en/of chemische eigenschappen van het meel, zoals die op het laboratorium van de Stoommeelfabriek "Holland" te Amsterdam met behulp van de apparatuur van Brabender en quantitatief-chemische analyse gemeten worden.

2. Algemeen overzicht.

Het onderzoek werd verricht door den heer H. de Jager, cand.math., onder leiding van Prof. Dr. D. van Dantzig.

Nadat de heer de Jager de in het laboratorium der Stoommeelfabriek "Holland" toegepaste meetmethoden bestudeerd en het beschikbare statistische materiaal systematisch samengevat had, werden eerst enige correlatiecoëfficiënten berekend. Enkele daarvan bleken vrij hoog te zijn, waardoor de mogelijkheid geopend werd, dat voortzetting van het onderzoek tot een positief resultaat zou leiden. De aanwezigheid van een zelfs vrij hoge correlatiecoëfficiënt tussen twee grootheden garandeert echter nog geenszins het bestaan van een bruikbaar verband tussen deze grootheden, in dier voege, dat meting van één van beide tot een betrouwbare prognose voor de andere zou leiden. Daarom werden de verkregen meetresultaten in een aantal puntdiagrammen tegen de te bepalen grootheid, de broodonttrek, uitgezet. Daarbij bleken de verschillende puntgroepen telkens in twee vrij duidelijk gescheiden delen uiteen te vallen, die bij groepen monsters van sterk verschillende samenstelling bleken te behoren. Binnen elk dezer groepen bleek in de meeste gevallen van een correlatie geen sprake te zijn, voor zover althans uit het vrij beperkte materiaal is af te lezen (een restrictie, die ook verder steeds moet worden in acht genomen). Dit betekent dus, dat de gevonden correlatiecoëfficiënten geheel of bijna geheel op rekening te stellen zijn van de verschillen in samenstelling. Anders gezegd: de apparatuur van Brabender kan alleen het verschil constateren tussen (b.v.) Amerikaans en inlands meel, maar daarvoor is zij niet nodig, daar de samenstelling vooraf al bekend is. Een fijnere analyse van verschillende meelsoorten van een zelfde type vermag deze apparatuur (voor zover uit het beschikbare materiaal is af te lezen) echter niet te geven. Beziat men b.v. het diagram van de rekenergie na 135 minuten, dan blijkt deze bij alle monsters van samenstelling I boven 65 cm^2 en bij alle monsters van samenstelling II daarbeneden te liggen. Echter komt een rekenergie tussen 90 en 95 cm^2 zowel bij de beste Amerikaanse soorten ($0 = 123,7$) als bij de slechtste ($0 = 105,6$) inlandse soorten voor.

Toch schijnt in dit diagram in de puntgroep die bij de inlandse soorten behoort van links naar rechts enige stijging aanwezig te zijn, hetgeen op een positieve correlatie wijst. Teneinde na te gaan, in hoeverre deze inderdaad aanwezig was, werd de reproduceerbaarheid van de afzonderlijke metingen onderzocht. Daartoe werden enige keren volkomen gelijke monsters aan de metingen onderworpen. Daarbij bleek, dat deze reproduceerbaarheid zeer gering is, d.w.z. dat volkomen gelijke monsters nog vrij aanzienlijk van elkaar verschillende meetresultaten kunnen geven. Dit geldt zowel voor de broodonttrek, waarbij een spreiding van 1 cm. werd gevonden als bij de verschillende fysisch-chemische grootheden. B.v. bedroeg deze spreiding bij de rekenergie 6 cm^2 . Deze grote variabiliteit bij gelijke monsters zag ieder-

re poging tot vaststelling ener correlatie-coëfficiënt volkomen illusoir maken, zodat daarvan werd afgezien.

Analoge beschouwingen gelden voor de andere fysische en chemische grootheden. B.v. liggen bij de Amerikaanse soorten waarden van het glutengehalte tussen 34% en 36% over bijna het gehele gebied der broodsoorten, n.l. van $O = 108,7$ tot $O = 123,7$ verspreid.

Ter nadere toelichting worden onder de punten 3 - 6 nog enige details over het onderzoek vermeld, waarna in punt 7 een samenvatting en slotconclusie gegeven wordt.

3. Omvang.

Het ter beschikking staande materiaal omvatte alle proevenseries van monsters uit het jaar 1946 (t/m 3 Jan. '47) waarvoor zowel volume en/of omtrek van het brood als ook één of meer der bedoelde grootheden gemeten was. Dit materiaal bestond uit 85 series, die reeds dadelijk in de volgende groepen waren te onderscheiden:

- A. 55 reeksen van algemene melanges, waarvan al dan niet enige procenten gries en/of patentmeel was getrokken.
- B. 24 reeksen van patentmeel.
- C. 5 reeksen van monsters van andere meelfabrieken.
- D. 1 reeks van een mengsel van A en B bloem (n.l. van melange 57) in onbekende verhouding.

Van de groepen C en D werd verder geen gebruik gemaakt, daar hun omvang te klein was; toevoeging van groep C bij A (hetgeen nog het meest voor de hand ligt) had misschien een onnauwkeurigheid in groep A kunnen veroorzaken. Gezien de geheel verschillende eigenschappen van de groepen A en B - hetgeen bij onderzoek direct opvalt - werden deze groepen bij alle onderzoeken afzonderlijk behandeld.

De volgende grootheden werden in het onderzoek betrokken:

- 1. Volume c.q. omtrek (in de zin van Jörgensen gemeten) van het brood.
- 2. Oppervlakte (rekenergie), hoogte (rekweerstand) en lengte (rekbaarheid) van het extensogram, met de extensograaf van Brabender, op vier verschillende tijdstippen bepaald.
- 3. Ontwikkeling van CO_2 gas in de fermentograaf, in het 1ste, 2de, 3de en 4de uur.
- 4. Gehalte aan as, gluten, maltose en $KBrO_3$ c.q. $(NH_4)_2 SO_4$ in het meel.
- 5. Gehalte aan NaCl-oplossing in deeg van bepaalde consistentie, ontwikkeltijd daartos en valeur.

Tenslotte werd de samenstelling der melanges (Amerikaanse of inlandse tarwe, rogge, soja) in aanmerking genomen.

4. Nauwkeurigheid der waarnemingen.

Alvorens een verband te onderzoeken tussen twee of meer grootheden, is het noodzakelijk, een beeld te verkrijgen van de nauwkeurigheid, die aan ieder gegeven moet worden toegekend. Indien b.v. zou blijken, dat de spreiding in verschillende metingen voor een bepaalde grootheid van eenzelfde monster (of gelijk te achten monsters) gelijk aan of zelfs groter dan deze spreiding voor verschillende monsters was, dan zou dit gegeven te onnauwkeurig bepaald zijn om tot enig verband ervan met andere grootheden te kunnen besluiten.

In een tweetal voor dat doel genomen proeven werden voor de brood-omtrek en de gegevens van het extensogram deze spreidingen onderzocht (22 en 24 Jan. '47). Een beschrijving van omstandigheden en waarnemingen is in de dossiers van het laboratorium der Stoommeelfabriek "Holland" te vinden. De gegevens zijn te gering in aantal, om de spreiding nauwkeurig te bepalen, maar voldoende om de grootte-orde aan te geven, zoals in Tabel 1 is geschied.

Kolom 1 vermeldt de gegevens: $\emptyset$ = omtrek, Re = rekenergie (oppervlakte van het extensogram), Rw = rekweerstand (hoogte van het extensogram) en Rb = rekbaarheid (lengte van het extensogram). Kolom 3 geeft een ruwe waarde voor de spreiding, gemeten in de in kolom 2 aangegeven eenheden. Bij de spreiding in de omtrek is te bedenken, dat deze (evenals de andere gegevens) betrekking heeft op broden van gelijk behandelde gelijk te achten monsters, niet op de verschillende broden, die als gebruikelijk van eenzelfde monster worden gebakken - welke laatste spreiding slechts 0,4 cm bedraagt. De gegevens van de Re , Rw en Rb hebben betrekking op het derde extensogram (na 135 minuten). Bij de andere extensogrammen zijn deze spreidingen - evenals hun verhoudingen tot de overeenkomstige gemiddelden - ongeveer half zo groot.

TABEL 1.

grootheid	eenheid	spreiding van gelijke monsters	spr. van al- le monsters	spr. van de ind. monsters
$\emptyset$	cm	1,0	5,2	2,2
Re	cm ²	6	28	8
Rw	cm	0,4	2,1	0,6
Rb	min.	1,3	3,8	1,7

Ter vergelijking zijn in kolom 4 de overeenkomstige spreidingen bij alle monsters van groep A aangegeven. Duidelijk blijken de waarden van kolom 4 groter te zijn dan die van kolom 3; de spreiding voor alle monsters kan slechts voor een gedeelte uit de onnauwkeurigheid in de bepaling worden verklaard en is voor het grootste deel aan essentiële verschillen in de samenstelling der verschillende monsters toe te schrijven. De betekenis van de 5e kolom zal later blijken.

5. Gang van het onderzoek.

Bij enige proevenreeksen was het volume bepaald; bij de meeste echter was de omtrek gemeten. Teneinde alle gegevens vergelijkbaar te maken, werden bij die proeven, waar een rechtstreekse omtrekbeoordeling ontbrak, deze volgens een door den heer H.G. Bolk aangegeven betrekking uit het volume afgeleid. Daarbij werd om praktische redenen aangenomen, dat dit verband éénduidig functioneel is, d.w.z. dat uit het volume de omtrek ondubbelzinnig en nauwkeurig valt af te leiden.

De gebruikelijke methode om het verband op te sporen tussen enige grootheden is nu het bepalen van de correlatiecoëfficiënten.

Bij de gegevens van groep A werden verscheidene correlatiecoëfficiënten bepaald. Wezenlijk van na verschillend bleken de correlatiecoëfficiënten te zijn, welke betrekking hebben op de grootheden Re en Rw van het extensogram. (De indices bij deze symbolen in het volgende geven het tijdstip van meting aan; zo betekent Re_{180} de energie na $4 \times 45 = 180$ minuten). De correlatiecoëfficiënt $_{180}$ van de omtrek met Re_{135} bleek ruim 0,7, die met Rw_{135} ongeveer 0,5 te zijn. Teneinde de afhankelijkheid van de verschillende grootheden nauwkeuriger na te gaan bleek het gewenst, de bijbehorende puntdiagrammen te tekenen. Zulke diagrammen werden gemaakt van de omtrek met de volgende grootheden: Rekenergie Re , rekweerstand Rw en rekbaarheid Rb , alle drie na 1,2,3 en 4 maal 45 minuten bepaald. Bovendien de verschillen $Re_{180} - Re_{90}$, $Rw_{180} - Rw_{90}$, $Rb_{180} - Rb_{90}$.

Deze grootheden werden gekozen als maten voor het tempo, waarmede genoemde eigenschappen van het deeg zich ontwikkelen. (Een dergelijke keuze is natuurlijk altijd enigszins willekeurig; we overtuigden er ons van, dat een andere keuze der tijdstippen of het vervangen van het verschil door het quotient geen betere resultaten zou hebben opgeleverd). Voorts de ontwikkeling F_1 t/m F_4 van CO_2 -gas in de vier uren, alsmede $F_4 - F_2$ en $F_1 + F_2 + F_3 + F_4$; vervolgens het gehalte aan as, gluten en maltose in meel, dat van H_2O met $NaCl$ in deeg, benevens de waarde. Bij de gegevens omtrent het gehalte aan harde Amerikaanse tarwe, rogge, soja en dat aan $KBrO_3$ en bij de ontwikkeltijd werden voor een ruwe verdeling in een paar ondergroepen de omtrekken vergeleken, daar deze grootheden slechts enkele waarden bleken aan te nemen.

Enige van de diagrammen zijn aan dit rapport toegevoegd; van de zojuist genoemde groeperingen geven de tabellen 2 en 3 een voorbeeld:

TABEL 2.

samenstelling	gemiddelde omtrek (in cm.)	spreiding van de omtrek (in cm.)
I	117,9	3,4
II a	111	2
II b	110,1	2,2

TABEL 3.

omtrek (in cm.)	105-108	108-111	111-114	114-117	117-120	120-123
aantal	-	1	2	13	7	7
waarnemingen	-	3	1	1	-	-
bij samenst.	3	9	5	1	-	-
Totaal	3	13	8	15	7	7

123-126	totaal
2	32
-	5
-	18
2	55

Hierin is de samenstelling onderscheiden naar de percentages aan harde Amerikaanse tarwe, zachte Amerikaanse- of inlandse tarwe en rogge.

I bevat 85-100% harde Amerikaanse tarwe, geen rogge (melanges 29 t/m 62)
 IIa " 35-52% " " " " (melanges 63 t/m 67)
 II b " 20-50% " " " en 10-30% " (melanges 68 t/m 83)

Uit deze tabellen blijkt, dat de samenstelling in belangrijke mate de omtrek beïnvloedt; de onderscheiding in IIa en IIb blijkt daarvoor en voor het volgende echter niet van belang. Derhalve zullen deze twee samenstellingen onder de aanduiding II worden verenigd. (De kleinste omtrek van samenst. I, melange 47 blijkt nu zo sterk van de andere af te wijken, dat aan een waarnemingsfout kan worden gedacht.)

Beschouwen we thans de diagrammen. Op het eerste gezicht blijken

Re en lw in verband met de onttrek te staan. Een nadere beschouwing leert echter het volgende: De monsters van samenstelling I gedroegen zich veelal geheel anders dan die van samenstelling II. Wat de Re betreft, deze zijn bij I steeds boven 65 cm², bij II steeds 135 eronder. Evenzo vertoont het glutengehalte een scheiding bij ongeveer 32%. Bij andere gegevens weer bestaat geen volledige scheiding of is de samenstelling niet uit de diagrammen af te lezen. Soms onderscheiden I en II zich door de grootte van de spreiding. Dat niet één van beide ondergroepen in elk opzicht heterogener van samenstelling is blijkt daaruit, dat de spreiding in de F₄ het grootst is bij I, de spreiding in het glutengehalte daarentegen bij II.

We zien dus, dat verscheidene gegevens in staat zijn te bepalen, of een monster tot samenstelling I dan wel II behoort, en dat op deze wijze indirect enig verband met de onttrek bestaat.

In de diagrammen zijn nu de punten die tot I en II behoren, met een x resp. een o aangegeven. Dit verschaft inzicht in het verband in elk der beide ondergroepen afzonderlijk. In elk diagram blijkt echter steeds, dat hoge waarden van een bepaalde grootheid zowel met grote als met kleine broodontrekken binnen deze groep gepaard kunnen gaan en dat hetzelfde voor de lage waarden geldt. Beschouwen we als voorbeeld het diagram van Re₁₃₅ nader, waar nog onder alle gegevens het sterkste verband met de broodontrek schijnt te bestaan, nader. De scheiding tussen de monsters x in het rechter-boven gedeelte en die der monsters o in het linker-onder quadrant is duidelijk. Binnen elk dezer ondergroepen is het echter niet mogelijk, uit het gegeven der Re₁₃₅ (d.w.z. uit de hoogte der punten) een conclusie omtrent de onttrek (d.w.z. de andere afmeting) af te leiden. Reeds blijkt uit Tabel 1, kolom 5, dat de spreidingen van alle monsters van samenstelling II- en tot deze behoorden de monsters, waarvan de spreiding in kolom 3 van deze tabel vermeld was- voor het grootste deel aan deze "toevallige" component te wijten zijn. Een nauwkeuriger bepaling dan de reeds aangeduide ruwe verdeling in twee groepen lijkt door middel van de Brabender-apparatuur niet mogelijk. En waar de samenstelling der monsters reeds van te voren bekend was en niet eerst experimenteel behoeft te worden bepaald, komen we tot het resultaat, dat de bepaling van de in de aanvang bedoelde grootheden, voor zover uit de ter beschikking staande gegevens kan worden afgeleid, geen waarde heeft voor de bepaling van de broodontrek.

De genoemde verdeling in twee groepen kan de gevonden correlatiecoëfficiënten volledig verklaren. Valt n.l. het materiaal uiteen in twee gelijke groepen met betrekking tot een veranderlijke (b.v. de onttrek) en neemt de andere veranderlijke (Re enz.) voor elke monster van eenzelfde groep dezelfde waarde aan, die echter voor beide groepen verschilt, dan wordt een correlatiecoëfficiënt van $\frac{1}{2} \sqrt{3} = 0,866$ gevonden. Het is evenwel duidelijk, dat deze tweede veranderlijke slechts tussen twee gebieden van de eerste veranderlijke kan onderscheiden, doch daarbinnen geen enkele uitspraak kan doen.

6. Opmerkingen.

In verband met deze negatieve conclusies zijn twee vragen nog van belang:

1e. Zijn de aantallen waarnemingen voor elke onderzochte grootheid voldoende om zeker te zijn van het negatieve resultaat of zal uitbreiding van deze aantallen wellicht nog nuttig kunnen zijn voor het gestelde doel?

2e. Kunnen andere dan de tot nog toe onderzochte grootheden of combinaties van de onderzochte grootheden nog positieve resultaten opleveren?

Wat de eerste vraag betreft: hoewel het materiaal van 32 resp. 23 melanges uiteraard gering is, is het wel voldoende om te constateren, dat de correlatie-coëfficiënten slechts weinig van nul zullen verschillen.

Voor enkele grootheden is het aantal waarnemingen bij ondergroep II echter voldoende. Voor de waarde zullen nadere bepalingen vermoedelijk wel niets opleveren, gezien de afwezigheid van correlatie bij de monsters van samenstelling I en de aanwijzing daartoe bij ondergroep II; dit vermoeden is echter geenszins een zekerheid. Slechts voor het asgehalte zijn er enige aanwijzingen te vinden, dat dit bij de monsters van samenstelling II met de omtrek zou kunnen samenhangen. Welke die samenhang is, daarover is uit het geringe aantal waarnemingen (9) natuurlijk geen conclusie te trekken. Daar het asgehalte vrij eenvoudig te bepalen is, menen wij tot voortzetting dezer bepalingen te moeten adviseren.

Voor de onderzoekingen in groep B geldt in grove trekken hetzelfde als bij groep A. Het resultaat is ook thans, dat uit de beschikbare gegevens niet tot enig nut van de metingen voor de bepaling van de omtrek kan worden geconcludeerd.

Met betrekking tot de tweede vraag kregen wij bij oppervlakkige beschouwing van de gegevens omtrent vochtgehalte, oventemperatuur, rijstijd en zeefproef de indruk, dat ook deze voor het gestelde doel van zeer geringe waarde zijn. Met betrekking tot de mogelijkheid, dat sommige combinaties der gemeten grootheden een belangrijk juistere prognose zouden mogelijk maken dan deze grootheden afzonderlijk, menen wij te moeten opmerken, dat het verkrijgen van uitsluitend daaromtrent een aanzienlijk omvangrijker statistisch onderzoek zou vereisen, terwijl het ons hoogst twijfelachtig voorkomt, of dit tot een positief resultaat zou leiden. Wij menen derhalve te moeten adviseren, zulk een uitgebreider onderzoek vooralsnog niet te entameren, tenzij het mocht gelukken, de meetnauwkeurigheid, d.i. de reproduceerbaarheid der afzonderlijke meetresultaten aanzienlijk te verhogen.

Tenslotte merken wij op, dat dit gehele onderzoek betrekking heeft op omtrekmeting. In hoeverre de uitgevoerde proefnemingen van belang zijn voor andere doeleinden (bijv. voor het bepalen van toe te voegen hoeveelheden $KBrO_3$) en dus toch zullen behoeven te worden verricht, valt buiten het terrein van dit onderzoek.

7 Samenvatting en conclusies.

Het ter beschikking gestelde statistische materiaal geeft geen aanleiding tot de onderstelling, dat op de bepaling van rekenergie, rekweerstand, rekbaarheid, CO_2 -ontwikkeling, waarde, as-, gluten- en maltosegehalte een zelfs maar matig betrouwbare prognose omtrent de broodomtrek kan worden gebaseerd, al schijnt deze mogelijkheid ten aanzien van het asgehalte bij inlandse meelsoorten niet geheel uitgesloten. De in enkele gevallen gevonden vrij hoge correlatie-coëfficiënten zijn geheel aan de verschillende samenstellingen der melanges toe te schrijven, d.w.z. voor zover uit het beschikbare materiaal kan worden afgelezen, is de Brabender-apparatuur niet in staat, een fijnere analyse van meelsoorten te constateren dan met de indeling in de samenstellingen I en II overeenkomt.

Tot verdere toepassing van de bepaling der genoemde fysische en chemische grootheden kan alleen geadviseerd worden, indien het gelukt, de meetmethoden zodanig te verbeteren, dat de variabiliteit bij herhaalde metingen van een grootheid aan eenzelfde monster aanzienlijk verminderd wordt. Ook voortzetting van het statistisch onderzoek door bepaling van multiple correlaties e.d. zou eerst daarna zinvol zijn, al zou vergelijking van het beschikbare vrij beperkte materiaal met het materiaal van andere laboratoria ook thans reeds zin hebben, daar de mogelijkheid niet uitgesloten moet worden geacht, dat sommige variaties in de apparatuur door vergrote meetnauwkeurigheid tot een geheel ander beeld zouden leiden.

Amsterdam, Mei 1947.









