

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
AFDELING MATHEMATISCHE STATISTIEK

EEN NIEUWE SIMULATIEMETHODE
GEILLUSTREERD AAN DE HAND VAN EEN WALSPROBLEEM

door R. D. Boschloo

Mathematisch Centrum, Amsterdam

rapport S 407 (SB 4)



Maart 1969

The Mathematical Centre at Amsterdam, founded the 11th of February, 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Hoofdstuk 1.

Een nieuwe simulatie methode geïllustreerd aan de hand van een walsprobleem.

§1. Inleiding

Wanneer men processen bestudeert waarin stochastische variabelen een rol spelen dan zal men eerst trachten problemen die voorkomen in het wiskundig model analytisch aan te pakken. Mocht evenwel het model voor een analytische aanpak te gecompliceerd zijn dan resten er, indien men de pogingen niet wil opgeven, twee manieren om alsnog tot een oplossing te geraken.

De ene manier bestaat uit het vereenvoudigen van het wiskundig model. Aangezien bij het kiezen van het oorspronkelijk model reeds vaak vereenvoudigingen hebben plaatsgevonden (b.v. door het opzettelijk verwaarlozen van details of door het maken van veronderstellingen over verdelingen enz.) is er aan de oplossingsmethode geen fundamenteel nieuw element toegevoegd; alleen zullen het model en de werkelijkheid minder overeenkomst vertonen.

De andere manier waarop het onderzoek toch voortgang kan hebben is het toepassen van simulatietechnieken, dwz. dat men het proces numeriek gaat nabootsen. De einduitkomsten van het nagebootste proces zullen - wanneer het model juist is gekozen - dezelfde kansverdeling bezitten als de uitkomsten die men in werkelijkheid zal vinden.

Ook bij het toepassen van simulatietechnieken boet het oorspronkelijk model aan nauwkeurigheid in evenals bij het vereenvoudigen van het model. Dit is gemakkelijk in te zien als men een willekeurige stochastisch variabele uit het oorspronkelijke model beschouwt en nagaat hoe deze variabele zich manifesteert in de simulatie:

Zoals gebruikelijk worden tijdens de simulatie met behulp van de inverse verdelingsfunctie trekkingen uit de gewenste verdeling gedaan. Het zal echter duidelijk zijn dat sommige eigenschappen van de oorspronkelijke verdeling van de variabele - die exact zijn vastgelegd door de parameters van deze verdeling - zelfs met een grote verzameling trekkingen niet meer precies is weer te vinden. Om de gedachte te bepalen beschouwe men b.v. de verwachting μ van een variabele: Deze verwachting is uit de verzameling trekkingen niet meer exact te bepalen:

Met moet volstaan met schattingen van de verwachting.

Hieruit zou men kunnen concluderen dat de rechtvaardiging van het gebruik van simulatietechnieken alleen te baseren is op het feit dat een analytische oplossingsmethode niet of althans moeilijk te realiseren is.

Men kan echter de mogelijkheden van simulatie ook op een meer positieve manier beschouwen. Ook wat betreft de exactheid. Daartoe is het nodig zich niet te beperken tot het oorspronkelijke wiskundige model want daarin zal - zoals we zagen - simulatie het wat betreft de exactheid altijd afleggen tegen een analytische aanpak. Naar zal blijken laat simulatie echter meer vrijheid bij de keuze van het model, zodat men in staat is een model te kiezen dat beter met de werkelijkheid overeenkomt.

Ten overvloede zij opgemerkt dat wel is waar een exacte analytische aanpak de onderzoeker in staat stelt om binnen het wiskundig model exacte uitspraken te doen, maar dat deze exactheid beperkt blijft tot het model daar de onzekerheid over juistheid van het model met betrekking tot de werkelijkheid blijft bestaan, zelfs - zoals bekend - na statistische toetsingen.

De genoemde betere keuze van het model die bij simulatie soms mogelijk is wordt verkregen door een grotere overeenkomst tussen de verdeling van de stochastische variabelen in het model en de corresponderende verdelingen in de werkelijkheid.

Daartoe zal eerst nagegaan worden hoe men bij een analytische aanpak tot de keuze van de verdelingen in het model komt.

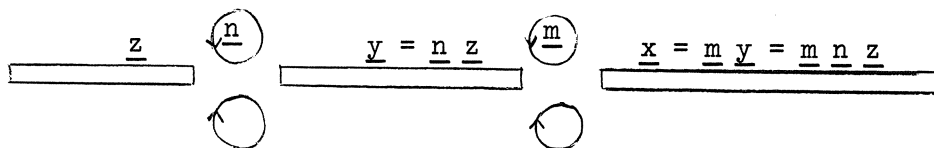
De keuzemogelijkheid blijft daar beperkt tot de handvol standaard verdelingen.

De voorkeur voor een bepaalde verdeling kan gebaseerd zijn op theoretische overwegingen. Dit is een solide basis.

Een andere solide basis is een keuze op grond van een statistisch onderzoek waarbij echter zij opgemerkt dat uit het niet kunnen verwerpen van een hypothese over de soort verdeling niet betekent dat de hypothese juist is.

Een zeer twijfelachtige basis wordt gevormd door "practische overwegingen met het oog op de berekeningen" of door het zonder argumentatie b.v. normaal onderstellen van een verdeling.

De voor een analytische aanpak noodzakelijke keuze van een verdelingssoort zal dus vaak de werkelijkheid geweld aandoen. Dit is vrijwel zeker het geval als de keuze nog verder beperkt is doordat slechts één of enkele verdelingen een analytische aanpak mogelijk maken. Soms is het zelfs helemaal niet mogelijk de verdelingen in het model zo te kiezen dat men alleen met standaard verdelingen te maken heeft. In een walsprobleem dat aanleiding was tot deze scriptie kwam (zeer vereenvoudigd) het volgende proces voor



Ijzerlengte z ondergaat bij de eerste wals een verlenging van n en bij de tweede nog eens een verlenging m . Met het oog op bepaalde knipstrategieën voor de lengte x (zie later) zou het probleem zeer gecompliceerd worden als x niet verdeeld is volgens een van de standaardverdelingen. Verder is het niet onlogisch te veronderstellen dat m en n dezelfde soort verdeling hebben. De beperking die men zich wat betreft de verdeling van x oplegt en de omstandigheid dat het nauwkeurig meten van de lengte z en y door temperatuur en vorm van het metaal worden bemoeilijkt maakten dat uitgaande van de verdelingen voor x en m en n gezocht werd naar de bijpassende verdeling voor y en vervolgens voor z .

Wanneer men mogelijke afhankelijkheden tussen de variabelen verwaarloost, zijn lognormale verdelingen een oplossing daar het product van twee onderling onafhankelijke lognormale verdelingen weer lognormaal is. Dit lijkt het enige type standaard verdeling waarmee het bovenstaande model is op te bouwen. Hoogstwaarschijnlijk zal bij het werken met standaard verdelingen, deze beperkte keuze de werkelijkheid geweld aandoen.

§ 2. De gebruikelijke simulatie methoden.

Veronderstel dat men trekkingen wil doen uit de verdeling van $\underline{x}$. Deze verdeling is onbekend maar men heeft de beschikking over een steekproef van n waarnemingen $x_i (i=1, \dots, n)$ uit $\underline{x}$.

Wanneer men deze steekproef niet gebruikt om na te gaan welke van de standaardverdelingen het best bij deze steekproef aansluit (zoals in het voorgaande gesteld is zal dit in het algemeen een vereenvoudiging van het model inhouden) dan lijkt de weg van de analytische aanpak afgesloten evenals de weg van - de zeer gebruikelijke - simulatie met behulp van standaardverdelingen. Een andere en rekenkundig eenvoudige methode is de simulatie met behulp van de n waarnemingen in de steekproef, waarbij men dus als gesimuleerde trekkingen uit $\underline{x}$ de opvolgende waarden van x_i kiest. Dit is vooral een goede methode als er afhankelijkheid tussen de opvolgende waarnemingen is wat bijvoorbeeld het geval is als x_i het aantal zonuren op de i -de dag voorstelt. Wanneer er tussen opvolgende trekkingen x_i geen afhankelijkheid bestaat kan men willekeurig uit de n waarnemingen trekkingen doen.

Een nadeel van deze methode is het feit dat de informatie over de verdeling van $\underline{x}$ in zo'n groot aantal getallen (de n waarnemingen in de steekproef) is vastgelegd.

Wanneer men in de werkelijkheid met een continue verdeling te maken heeft dan is een opvallend verschil tussen de uitkomsten in de werkelijkheid en die in de simulatie dat er in de laatste slechts een beperkt aantal verschillende uitkomstwaarden voorkomt.

De hier te poneren methode heeft deze laatstgenoemde nadelen niet daar er met behulp van $x_1, \dots, x_n$ een continue verdelingsfunctie wordt geconstrueerd die door een beperkt aantal parameters is vast te leggen.

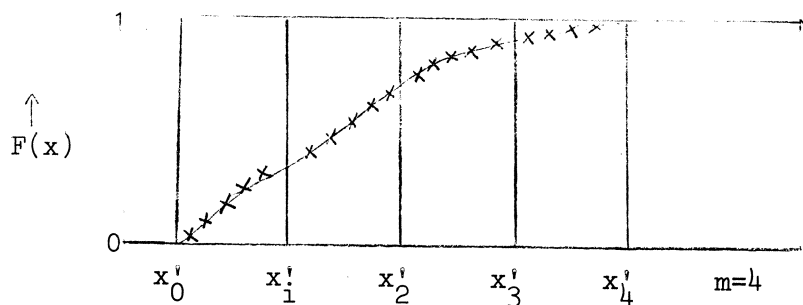
Deze verdelingsfunctie kan met willekeurige nauwkeurigheid aan een gegeven aantal punten worden aangepast. Zelfs eventuele meertoppigheid van de verdelingsdichtheid vormt geen bezwaar, terwijl de inverse verdelingsfunctie eenvoudig uit de verdelingsfunctie is af te leiden.

§ 3. De nieuwe simulatie methode.

Onderstel dat men een verzameling van n waarden van x heeft waarvoor de waarde van de verdelingsfunctie $y_i = F(x_i)$ $i=1, \dots, n$ bekend is. De functie $F(x)$ zelf is onbekend.

De x -as wordt zodanig in m intervallen verdeeld dat ieder interval hetzelfde aantal waarden x_i bevat. Stel x_j^* $j=0, 1, \dots, m$ vormen de grenzen van de m intervallen.

Zie de figuur:



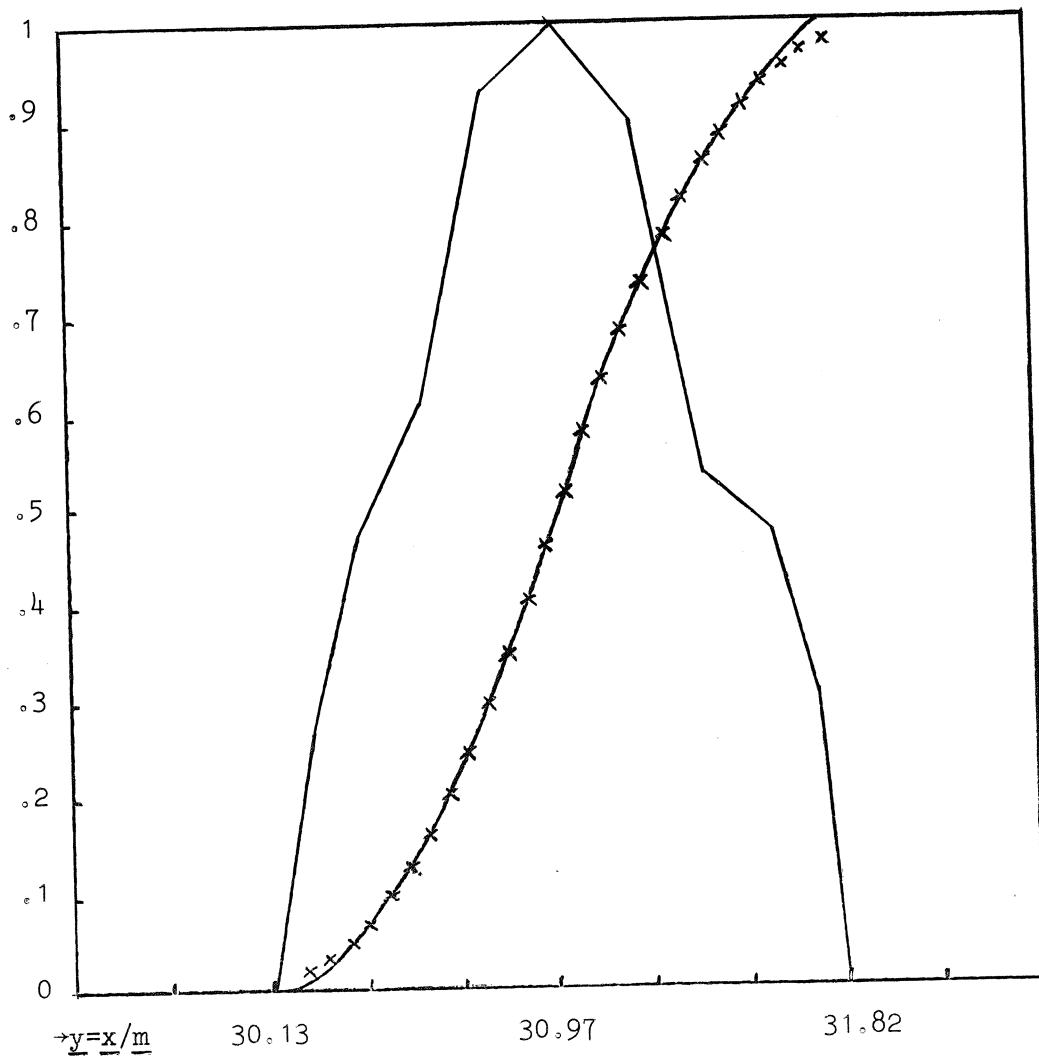
De grondgedachte is nu om beginnend in het midden voor ieder interval met behulp van de kleinste kwadraten methode een kwadratische functie aan te passen langs de punten (x_i, y_i) waarbij geëist wordt dat in de grenspunten x_j^* waar twee opvolgende functies in elkaar overgaan de afgeleiden van beide functies aan elkaar gelijk zijn.

Het ALGOL-60 programma waarmee dit gebeurt is vrij lang en gecompliceerd en zal hier niet besproken worden.

De resultaten van twee aanpassingen zijn te zien in de figuren 1 en 2. In deze figuren zijn respectievelijk de verdelingen van de ijzerlengten $\underline{z}$ en $\underline{y}$ afgebeeld, die zijn bepaald uit de verdelingen van $\underline{x}$, $\underline{m}$ en $\underline{n}$ op de volgende manier:

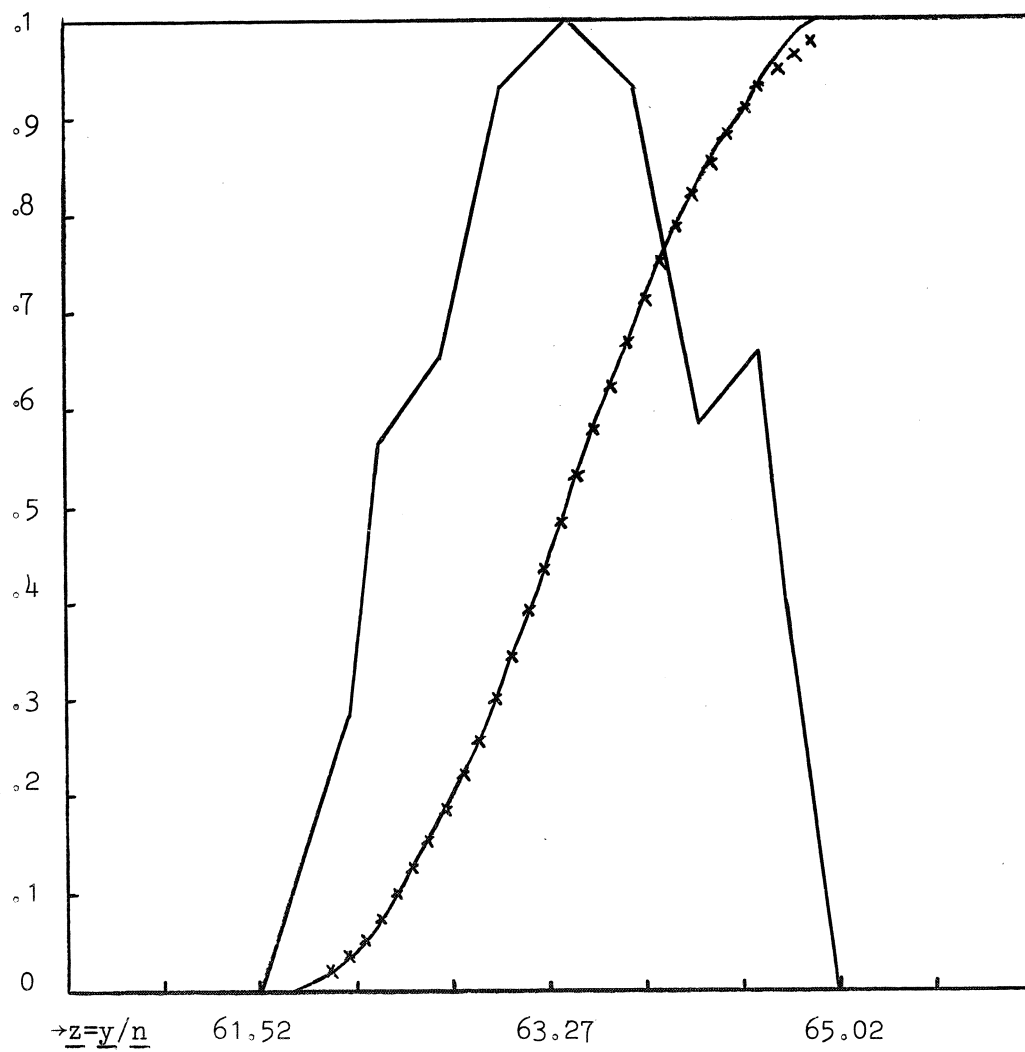
De verdelingsfunctie van b.v. $\underline{y}$ is aangepast aan een aantal punten die zijn berekend door middel van de verdelingen van de eindlengte $\underline{x}$ en de verlengingsfactor $\underline{m}$. Zoals bekend $\underline{y} = \underline{x} / \underline{m}$.

De verdeling van x is normaal ondersteld en voor de verdelingsfunctie van $\underline{m}$ is een functie gekozen die uit 6 aaneensluitende kwadratische functies bestaat die differentieerbaar zijn in de overgangspunten.



Figuur 1.

De aanpassing van de verdelingsfunctie van $\underline{y}$ aan de berekende punten.
 Op 10 intervallen zijn aansluitende kwadratische functies aangepast.
 Ook is de verdelingsdichtheid afgebeeld in de figuur.



Figuur 2.

De aanpassing van de verdelingsfunctie van z aan de berekende punten.
 Op 10 intervallen zijn aansluitende kwadratische functies aangepast.
 Ook is de verdelingsdichtheid afgebeeld in de figuur.

Stel $\underline{x}$ heeft de verdelingsdichtheid $f(x)$ en
 $\underline{m}$ heeft de verdelingsdichtheid $g(m)$.

Veronderstel verder dat $\underline{x}$ en $\underline{m}$ onderling onafhankelijk zijn (deze onderstelling is waarschijnlijk niet geheel in overeenstemming met de werkelijkheid) dan is het mogelijk de verdelingsfunctie $H(y)$ van $\underline{y}$ op te schrijven:

$$\begin{aligned} H(y) = P(\underline{y} \leq y) &= \int \int_{\underline{x}/\underline{m} \leq y} g(m) f(x) \, dm \, dx = \\ &= \int_{-\infty}^{\infty} g(m) \left\{ \int_{-\infty}^{ym} f(x) \, dx \right\} dm = \\ &= \int_{-\infty}^{\infty} g(m) F(ym) \, dm. \end{aligned}$$

Voor de dichtheid geldt dan

$$h(y) = H'(y) = \int_{-\infty}^{\infty} g(m) f(ym) m \, dm.$$

Daar g en f bekend zijn, is $h(y)$ met behulp van een numerieke integratie methode voor iedere waarde van y te berekenen.

Hetzelfde geldt voor $H(y)$.

Men is dus in staat op numerieke wijze voor een aantal argumenten y_i de bijbehorende functie waarden $H(y_i)$ te berekenen.

Door de verzameling punten $(y_i, H(y_i))$ is vervolgens op de in het voorgaande uiteengezette methode een aantal aansluitende kwadratische functies aangepast (zie figuur 1).

Voor de verdeling van $\underline{z}$ is analoog te werk gegaan, echter met dit verschil dat de convolutie integraal hier genomen is over twee per interval kwadratische functies.

Op de bovenstaande manier is een oplossing gevonden voor de verdelingen van het walsprobleem.

Door het aantal intervallen groot te kiezen verkrijgt men een willekeurige kleine afwijking van de kwadratische kromme tot de berekende punten.

Opmerking 1. De figuren laten zien dat in de staarten de aanpassing het minst goed is.

Dit komt doordat aan de kwadratische functies op de buitenste intervallen de eis werd gesteld dat de afgeleide (Dus de dichtheidsfunctie) nul is in de uiterste grenspunten.

Deze eis is betrekkelijk willekeurig.

Opmerking 2. Wanneer de coëfficiënten van een kwadratische functie bekend zijn dan is het vinden van de inverse verdelingsfunctie - die nodig is voor het doen van trekkingen in de simulatie - vrij eenvoudig.

Voor de eenduidigheid van de inverse functie is monotonie vereist. In het ALGOL-programma is daarom een beveiliging opgenomen die ertegenwaakt dat er op een interval een functie wordt gekozen die niet sterk monotoon stijgend is op dat interval.

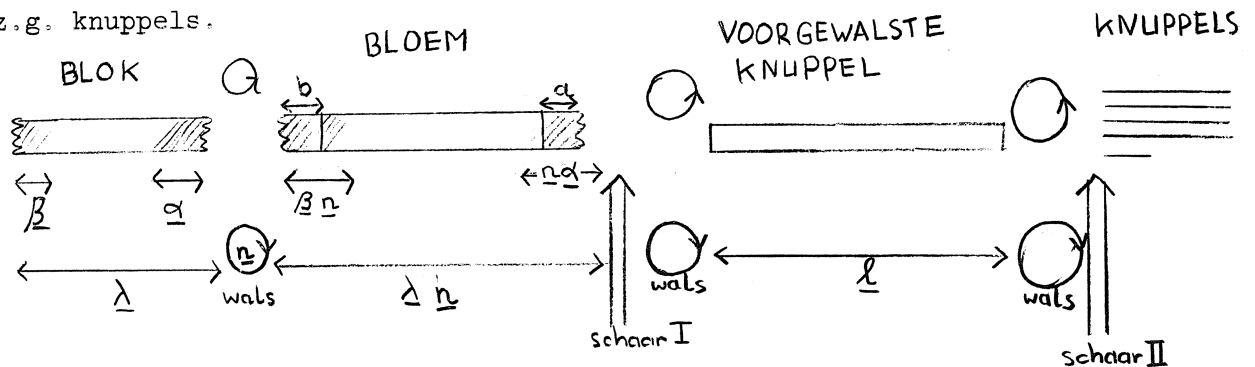
Hoofdstuk II De simulatie van het walsproces.

§1. Procesbeschrijving.

Bij de Hoogovens komt het onderstaande walsproces voor waarbij de probleemstelling is een tweetal ijzerscharen zodanig af te stellen dat de som van het verlies aan afkeurijzer en het verlies t.g.v. extra bewerkingen minimaal is.

Schaar I dient om het z.g. dubbelijzer (DY) dat zich aan de uiteinden van het metaal bevindt grotendeels weg te knippen. Teveel wegnippen betekend echter verlies aan goedijzer (GY).

Schaar II knipt na de laatste wals het metaal in handelbare stukken, de z.g. knuppels.



Voor de eerste walsing zijn α en β de lengtes DY.

Schaar I knipt de vaste lengtes a en b van de bloem af. a slaat op het voorste stuk.

Vóór de laatste wals wordt de lengte λ gemeten en aan de hand van deze λ wordt, volgens de strategie $s(\lambda)$, schaar II afgesteld op het knippen van knuppels met lengte $s = s(\lambda)$.

Daar geldt $11,5 < s < 12$ meter is niet iedere lengte restloos te delen (b.v. lengtes 120 en 126,5 meter kunnen niet restloos in knuppels van een toegelaten lengte gedeeld worden).

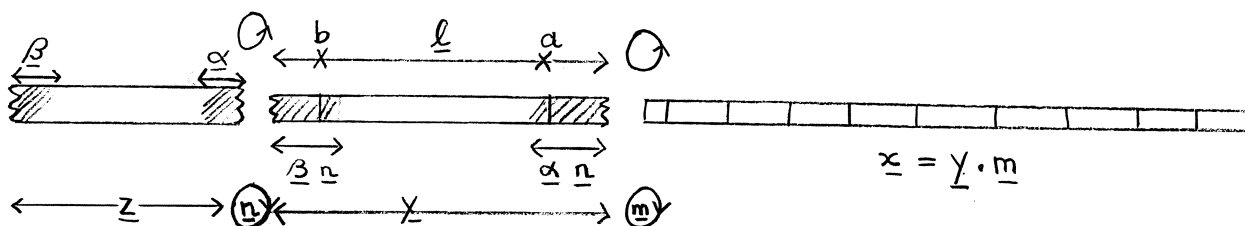
De verlengingsfactoren van de verschillende walsen en de ijzerlengten zijn zeer waarschijnlijk allen stochasten die onderlinge afhankelijkheden kunnen vertonen.

Aangezien de voor deze scriptie uitgevoerde simulatie slechts ten doel heeft een demonstratie te geven van de geconstrueerde simulatie gereedschappen is in het simulatiemodel een wals weggelaten.

Er blijven daardoor 2 walsingen over en deze maken voldoende duidelijk dat het toevoegen van de derde walsing aan de simulatie, geen problemen zal geven.

Het weglaten van nog een walsing is overwogen. Daar het oorspronkelijke uitgangsmodel daardoor echter niet zo eenvoudig te herkennen was is deze vereenvoudiging van het model niet aangebracht.

Het model ziet er nu als volgt uit



De uiteindelijk afgeleverde knuppels hebben niet allen de lengte s . De restknuppel en de beginknuppel (waar soms DY afgebrand moet worden) zullen - al naar gelang de knuppellengte k - tot een van de volgende categorieën behoren:

Categorie	lengte GY in meters	kosten per knuppel	afbrandkosten per ton DY	verlies in op- brengst per ton GY
DY-vrije knuppel	$10 < k < 12$	(opbrengst maximaal)		
Afbrand knuppel	$10 < k < 12$	C_2		
Proef 13 knuppel	$8 < k < 10$	C_{33}	C_{31}	C_{32}
Afkeur knuppel	$k < 8$			C_1

Het verschil tussen de eerste en de tweede categorie is dat de eerste categorie geen DY bevat en de tweede wel.

Het verlies aan GY in de afgeknipte stukken a en b is ook C_1 per ton.

De totaal opbrengst is maximaal als de kosten en het verlies aan GY minimaal zijn.

Gevraagd wordt de optimale beslissingen a en b en de strategie $s(1)$ bepalen.

Een oplossing zonder een bijgevoegd gekwantificeerd model dat de mogelijkheid biedt om de gevolgen van een eventuele andere beslissing te voorspellen heeft echter voor de leiding van de walserij weinig waarde.

§2 Het kiezen van een geschikt model (overzicht)

In het oplossingsschema kan men 3 fasen onderkennen:

- I Het bepalen van de vereiste verdeling in het bijzonder die van y en z . (met behulp van 2 ALGOL-programma's)
- II Het bepalen van de knipstrategie $s(1)$ (2 ALGOL-programma's)
- III Met behulp van het in I en II bepaalde, de simulatie uitvoeren voor verschillende waarden van a en b (1 ALGOL-programma).

Ad I. In hoofdstuk 1 bleek dat het gebruik van standaardverdelingen grote problemen geeft.

Als alternatief werd toen een nieuwe methode geïntroduceerd die is toegepast voor het vinden van de verdelingen voor y en z als de verdelingen van x , m en n gegeven zijn.

In hoofdstuk I is ook uiteengezet waarom het voordelen heeft het model van achter naar voren op te bouwen dus beginnend bij x en m via de dan te vinden y en n naar z .

Ten aanzien van m en n kan men iedere verdeling kiezen die men wenst. Om dit te illustreren is in het hier gesimuleerde model een per interval kwadratische verdelingsfunctie gekozen die ook continu en differentieerbaar is in de interval overgangen.

Aan het bestaan van mogelijke afhankelijkheden tussen de optredende variabelen is in deze scriptie weinig aandacht besteed. Dit mede doordat cijfermateriaal hierover grotendeels ontbreekt.

Daar er van achteren naar voren gewerkt wordt is het mogelijk een verdeling voor x te kiezen.

Wanneer x een normale verdeling heeft wordt de uitwerking van het model sterk vereenvoudigd zoals in §3 zal blijken.

Ad II. Mede door het normaal onderstellen van de verdeling van x is het mogelijk een knipstrategie te berekenen indien men de aanname maakt dat er geen DY aanwezig is. (Dit gebeurt in §4)

Dit is een zeer onrealistische aanname.

Het is echter eenvoudig aan te tonen dat eventueel DY aan het begin van het te walsen metaal geen invloed op de knipstrategie heeft. (Dit gebeurt in §5). Ten overvloede is dit door een simulatie gecontroleerd en geconstateerd. Voor het DY aan het eind wordt een correctie op de strategie aangebracht die door simulatie is bepaald (Dit wordt beschreven in §6)

Ad III. Met behulp van de gevonden strategie $s(1)$ en de gevonden verdelingen is het uitvoeren van de simulatie nu mogelijk geworden.

Voor verschillende waarden van de kaplengtes a en b wordt het verlies aan GY + de kosten door extra bewerkingen gesimuleerd en afgedrukt.

Tevens worden de percentages proef 13 knuppels, afkeurknuppels en afbrandknuppels onder de begin en eindknuppels berekend en afgedrukt (Deze resultaten worden vermeld in §7).

Het verlies in de beginknuppel als functie van a wordt dus ook door simulatie gevonden, hetgeen strikt genomen niet nodig is. Dit verlies had ook analytisch bepaald kunnen worden zoals aan het slot van de scriptie in §8 uiteengezet is.

§3 De normale verdeling voor x ;

μ_i is de verwachte totale lengte (de som van de lengtes van de knuppels en de restknuppel) bij de i -de realisatie van het proces.

Deze μ_i wordt bepaald aan de hand van de tijdens het proces gemeten lengte l_i (zie §1 van hoofdstuk II)

Het verband tussen μ_i en l_i is deterministisch; wanneer men een lengte l gemeten heeft dan ligt daardoor de waarde van de verwachte totale lengte $\mu = f(l)$ vast.

Hierbij dient echter het volgende te worden opgemerkt:

In de praktijk vertoont de functie $f(l)$ van tijd tot tijd nog wel eens verschillen (b.v. door slijtage aan de walsen), waardoor het de voorkeur verdient het personeel de vrije hand te laten bij het voorspellen van μ . Dit blijkt n.l. betere resultaten te geven dan het uit tabellen laten aflezen van functiewaarden met l als argument. In deze scriptie zal dan ook de af te leiden knipstrategie als functie van μ worden geschreven en niet als functie van l .

Na de laatste walsing kan men drie verdelingen onderscheiden die alle drie in verband staan met de totale lengte.

Allereerst is er de verdeling $F(x)$ van de totale lengte $\underline{x}$.

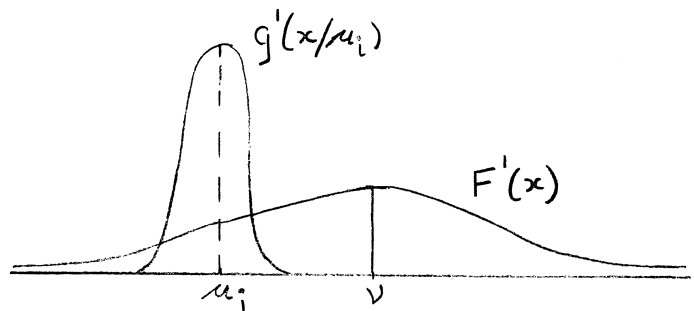
Verder de verdeling $H(\mu)$ van de verwachting $\underline{\mu}$, die evenals $\underline{l}$ stochastisch is. En tenslotte de voorwaardelijke verdeling $G(x|\mu)$ van $\underline{x}$ als $\underline{\mu} = \mu$; wanneer men een uitkomst μ verwacht zal de in werkelijkheid gerealiseerde $\underline{x}$ verdeeld zijn volgens $G(x|\mu)$.

Bewezen kan worden dat als $\underline{\mu}$ uit $\mathcal{N}(\nu, \tau^2)$

en $\underline{x}|\mu_i$ uit $\mathcal{N}(\mu_i, \sigma^2)$ getrokken

zijn, dat dan "x voorwaardelijk" een $\mathcal{N}(\nu, \sigma^2 + \tau^2)$

verdeling heeft.



Het normaal onderstellen van de verdelingen is een vereenvoudiging van het model die het berekenen van de knipstrategie als functie van de verwachte lengte μ sterk vergemakkelijkt (zie de volgende paragraaf waar van het normaal verdeeld zijn van $x|\mu_i$ gebruik wordt gemaakt).

§4. De optimale strategie berekening bij afwezigheid van DY.

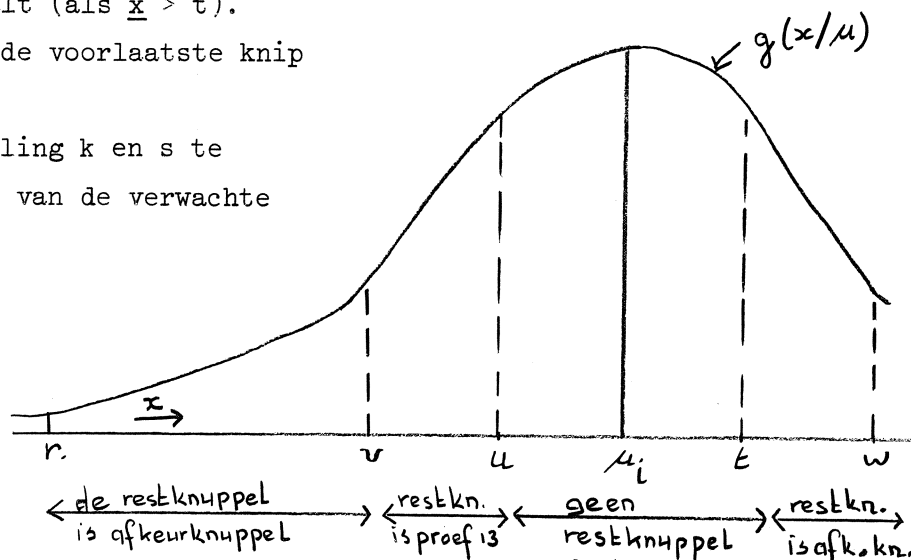
Stel k is het aantal knuppels van lengte s dat men tracht te verkrijgen.

$t = s$. k is het punt waar de laatste knip valt (als $\underline{x} > t$).

Bij r valt dan de voorlaatste knip

dus $t - r = s$.

Het is de bedoeling k en s te kiezen op grond van de verwachte uitkomst μ_i .



In de figuur is r de voorlaatste knip. $r = t - s = \frac{k-1}{k} t$.

$v = r + 8$ en $w = t + 8$. Valt de totale lengte $\underline{x}$ tussen r en v of tussen t en w dan is de restknuppel een afkeurknuppel wat een verlies betekend van C_1 per ton GY wat overeenkomt met c_1 per meter GY.

$u = r + 10$. Wanneer de totale lengte x tussen v en u terechtkomt veroorzaakt de restknuppel, die een proef 13 knuppel is, een verlies van C_{32} per ton GY.

Het beschouwde interval (r, w) is zonodig uit te breiden.

Zoals reeds is opgemerkt moet t gekozen worden op grond van de verwachte totale lengte μ_i . De waarden v , u en w zijn bepaald uitgaande van t en zijn dus functies van t : b.v. $v(t) = r + 8 = t \times (k - 1)/k + 8$, enz.

Bij gegeven μ_i en gegeven $g(x|\mu)$ is het te verwachten verlies een functie van t :

$$V(t|\mu) = c_1 \int_{r(t)}^{v(t)} (x - r(t)) \cdot g(x|\mu) dx + c_{32} \int_{v(t)}^{u(t)} (x - r(t)) \cdot g(x|\mu) dx + \\ + c_1 \int_t^{w(t)} (x - t) \cdot g(x|\mu) dx.$$

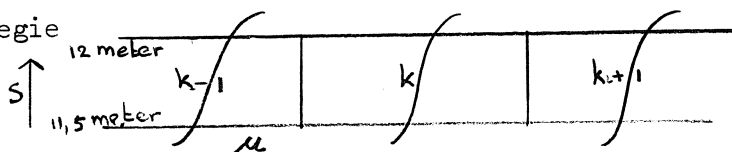
De vergelijking

$\frac{\partial}{\partial t} \mathcal{V}(t|\mu) = 0$ levert de optimale $t^* = ks^*$ waarvoor $\mathcal{V}(t^*|\mu)$ minimaal is.

De berekening van t^* kan men voor verschillende waarden van μ herhalen.

Vervolgens kan men de strategie functie $t^* = \mathcal{T}(\mu)$ of $s^* = t^*/k$ in de vorm van een polynoom construeren. Voor iedere waarde van k krijgt men een polynoom. De totale strategie

krijgt de vorm :



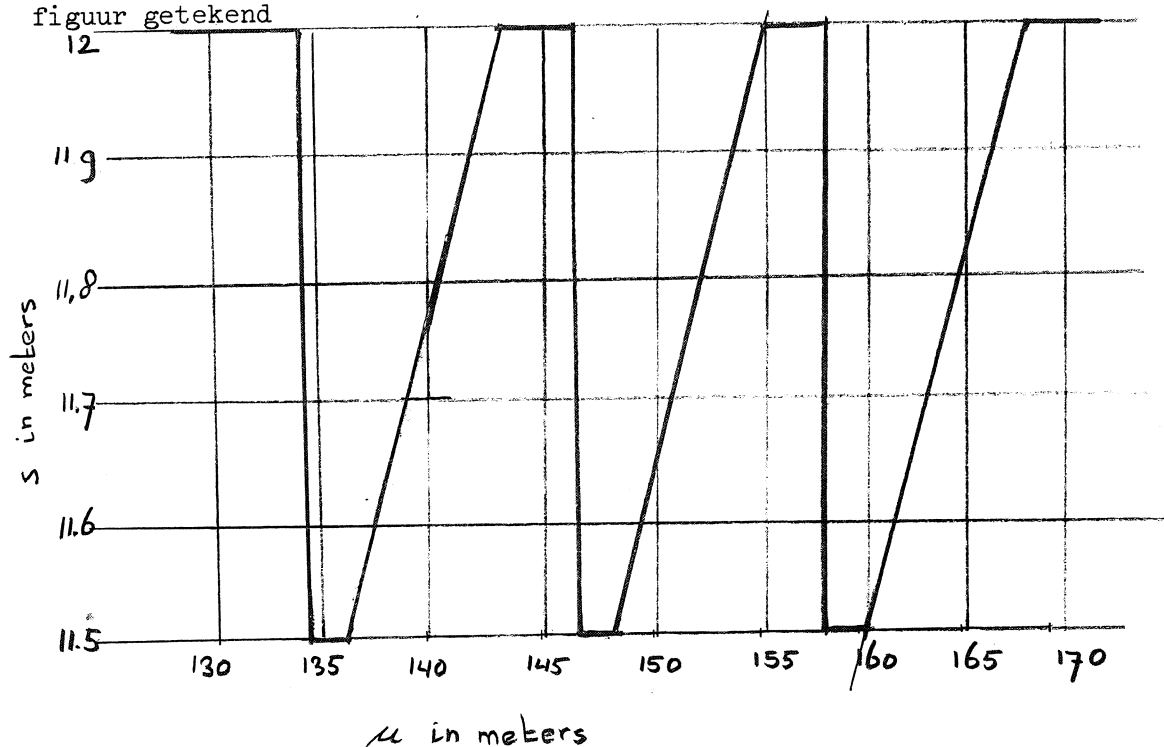
In de simulatie is voor c_2 en c_{33} de waarde 3 gekozen

c_1 , c_{32} en c_{31} hebben de waarden gekregen $c_1 = 60$, $c_{32} = 1$ en $c_{31} = 3$, waarbij moet worden opgemerkt dat dit de kosten zijn die gelden per meter ijzer in de eindlengte (C_1 , C_{32} en C_{31} golden per ton ijzer).

Voor de verdeling van $\underline{x}|\mu$ is een normale verdeling gekozen met $\sigma = 40$ cm. (en verwachting μ).

Onder deze voorwaarden bleek het polynoom ongeveer een lineaire functie te zijn. Daarom is, nadat eerst voor een aantal waarden van μ de optimale strategie werd berekend, aan deze punten een lineaire functie aangepast met behulp van de kleinste kwadraten methode.

De gevonden strategie bij afwezigheid van DY is in de onderstaande figuur getekend



§5 De invloed van a op de strategie.

In §8 zal langs analytische weg worden aangetoond dat a een duidelijke invloed heeft op het verlies. Dit wordt bevestigd door de simulatie uitkomsten in §7.

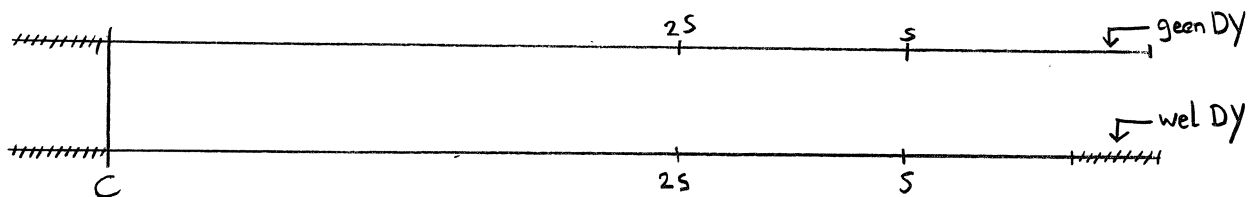
In deze paragraaf wordt echter aangegeven waarom de grootte van a geen invloed heeft op de strategie $s(\mu)$, die bij iedere verwachte totale lengte μ een kniplengte s voorschrijft.

De bewering van deze paragraaf is volledig bevestigd door een simulatie.

De vraag of de grootte van a invloed heeft op de strategie komt op hetzelfde neer als de vraag of de hoeveelheid DY aan het begin invloed heeft op de strategie. (Uiteraard wordt wanneer men a verandert de verwachte totale lengte μ anders, de bewering is echter dat de functie $s(\mu)$ hetzelfde blijft).

Het probleem is terug te brengen tot de vraag of men op de volgende 2 willekeurig gekozen even grote lengtes, een verschillende kniplengte s moet toepassen omdat de ene lengte wel DY aan het begin (rechts) bevat en de andere niet.

In paragraaf 1 van dit hoofdstuk werd gesteld dat s gekozen mag worden tussen 11,5 en 12 meter terwijl eindlengtes die tussen 10 en 12 meter GY bevatten een maximale opbrengst waarde aan GY hebben. In de praktijk blijkt de eerste knuppel altijd minstens 10 meter GY te bevatten.



De tweede staaf zal bij gelijke s een kleinere opbrengst hebben dan de eerste staaf daar de eerste knuppel van de tweede staaf DY bevat.

Als restloos delen mogelijk is zal men in beide gevallen een maximale opbrengst hebben als men s zo kiest dat de laatste knip bij c valt.

Als restloos delen niet mogelijk is zal men s zo kiezen dat de restlengte zo klein mogelijk is.

s is in beide gevallen - waarbij de totale lengte willekeurig is gekozen - dus even groot.

Er is een theoretische kleine kans dat $s-p$ een proef 13 knuppel is (in de praktijk komen proef 13 knuppels aan het begin niet voor) maar $s - p + \Delta s$ net niet. Door dus de strategie $s + \Delta s$ te kiezen in plaats van s wordt de opbrengst van de beginknuppel groter. Het verlies aan de eindknuppel zal echter ook weer groter worden.

De ervaringen in praktijk en simulatie wijze dan ook niet in de richting van een beïnvloeding van de strategie door a .

§6 Simulatie van de optimale strategie bij aanwezigheid van DY.

In de vorige paragraaf werd de veronderstelling dat DY in het begin van de totale lengte geen invloed op de strategie zou hebben uitgewerkt.

Een positieve kans op aanwezigheid van DY aan het einde van de totale lengte zal echter tot gevolg hebben dat men de kniplengte s iets kleiner moet kiezen dan wanneer geen DY aanwezig is; men kiest dus een s die hoort bij een kleinere waarde van μ .

Hopelijk blijft echter de strategiefunctie die in de paragraaf 4 berekend is bruikbaar door als argument $\mu - \epsilon$ in te vullen in plaats van μ .

Door simulatie kan men nu trachten bij verschillende waarden van b , dus bij verschillende kansen op DY aan het einde van $\underline{x}$, de functie $\epsilon(b)$ te vinden waardoor de strategiefunctie $\mathcal{T}(\mu - \epsilon(b))$ een algemenere geldigheid zou krijgen.

Bij de simulatie, met de in deze scriptie gekozen waarden voor de constanten, bleek dat door de verschuiving ϵ de verlieskosten bij aanwezigheid van DY aan het eind inderdaad iets verminderden om bij grote ϵ weer sterk te stijgen. Het minimum werd bereikt op een traject dat veelal 30 cm. breed was zodat er ruimte was om voor $\epsilon(b)$ een eenvoudige functie te kiezen.

De gekozen formule is

$$\epsilon(b) = \begin{cases} 0 & \text{als } b \geq .83 \\ 4.15 - 5b & \text{als } b < .83 \end{cases}$$

Mocht echter blijken dat voor andere waarden van de constanten een lineaire functie niet voldoet dan kan men een hogere graads functie in b kiezen of zelfs - als blijkt dat de correctie ook van de verwachte lengte μ afhangt - een functie $\epsilon(\mu, b)$.

TABEL I
(simulatie resultaten)Percentages afkeurknuppels
proef 13 knuppels
afbrand knuppelsbegin - eind
begin - eind
begin - eind

b	.40	.45	.50	.55	.60	.65	.70	.75
a	0 - 53	0 - 53	0 - 53	0 - 52	0 - 50	0 - 47	0 - 10	0 - 8
.20	42 - 0 58 - 47	42 - 0 58 - 47	42 - 0 58 - 47	42 - 0 58 - 48	42 - 0 58 - 50	42 - 0 58 - 52	42 - 0 58 - 33	42 - 1 58 - 15
.25	0 - 51 25 - 0 75 - 49	0 - 51 25 - 0 75 - 49	0 - 51 25 - 0 75 - 49	0 - 50 25 - 0 75 - 50	0 - 48 25 - 0 75 - 52	0 - 44 25 - 0 75 - 54	0 - 7 25 - 0 75 - 34	0 - 4 25 - 1 75 - 17
.30	0 - 49 15 - 0 83 - 51	0 - 49 15 - 0 83 - 51	0 - 49 15 - 0 83 - 51	0 - 48 15 - 0 83 - 52	0 - 46 15 - 0 83 - 54	0 - 44 15 - 0 83 - 55	0 - 4 15 - 0 83 - 36	0 - 3 15 - 2 83 - 17
.35	0 - 49 7 - 1 89 - 50	0 - 49 7 - 1 89 - 50	0 - 49 7 - 1 89 - 50	0 - 48 7 - 1 89 - 51	0 - 46 7 - 1 89 - 53	0 - 42 7 - 1 89 - 54	0 - 3 7 - 1 89 - 35	0 - 3 7 - 3 89 - 16
.40	0 - 48 3 - 1 93 - 51	0 - 48 3 - 1 93 - 51	0 - 48 3 - 1 93 - 51	0 - 47 3 - 1 93 - 52	0 - 44 3 - 1 93 - 53	0 - 40 3 - 1 93 - 56	0 - 3 3 - 2 93 - 34	0 - 2 3 - 3 93 - 16
.45	0 - 47 0 - 1 82 - 52	0 - 47 0 - 1 82 - 52	0 - 47 0 - 1 82 - 52	0 - 45 0 - 1 82 - 53	0 - 42 0 - 1 82 - 55	0 - 39 0 - 2 82 - 54	0 - 2 0 - 2 82 - 34	0 - 1 0 - 5 82 - 15
.50	0 - 44 0 - 1 67 - 55	0 - 44 0 - 1 67 - 55	0 - 44 0 - 1 67 - 55	0 - 42 0 - 1 67 - 56	0 - 40 0 - 2 67 - 56	0 - 38 0 - 2 67 - 55	0 - 1 0 - 3 67 - 33	0 - 0 0 - 5 67 - 15
.55	0 - 43 0 - 3 46 - 54	0 - 43 0 - 3 46 - 54	0 - 42 0 - 3 46 - 55	0 - 41 0 - 4 46 - 54	0 - 39 0 - 4 46 - 55	0 - 36 0 - 4 46 - 56	0 - 0 0 - 4 46 - 32	0 - 0 0 - 7 46 - 14
.60	0 - 43 0 - 6 28 - 51	0 - 42 0 - 6 28 - 52	0 - 42 0 - 6 28 - 51	0 - 41 0 - 6 28 - 52	0 - 38 0 - 6 28 - 54	0 - 36 0 - 6 28 - 54	0 - 0 0 - 6 28 - 31	0 - 0 0 - 12 28 - 13

TABEL II
(simulatie resultaten)

onder elkaar

verlies aan het begin
verlies aan het eind
t o t a a l verlies

b	.40	.45	.50	.55	.60	.65	.70	.75	.80
a									
.20	461 2575 3036	461 2575 3036	461 2575 3036	461 2585 3047	461 2651 3113	461 2850 3311	461 2581 3042	461 3595 4057	461 4851 5312
.25	287 2403 2690	287 2403 2690	287 2403 2690	287 2413 2700	287 2478 2766	287 2684 2972	287 2431 2718	287 3446 3733	287 4806 5093
.30	204 2295 2500	204 2295 2500	204 2295 2500	204 2306 2510	204 2371 2576	204 2581 2786	204 2296 2501	204 3402 3607	204 4771 4976
.35	159 2208 2366	159 2208 2366	159 2208 2366	159 2218 2376	159 2283 2442	159 2485 2643	159 2252 2411	159 3367 3526	159 4749 4908
.40	177 2142 2319	177 2142 2319	177 2142 2318	177 2151 2328	177 2201 2378	177 2439 2616	177 2217 2394	177 3335 3512	177 4727 4904
.45	307 2053 2360	307 2053 2360	307 2053 2360	307 2059 2366	307 2133 2440	307 2395 2702	307 2185 2492	307 3312 3620	307 4692 4999
.50	674 1977 2650	674 1977 2650	674 1976 2650	674 1991 2665	674 2085 2759	674 2357 3030	674 2153 2826	674 3278 3952	674 4701 5375
.55	1334 1938 3272	1334 1937 3272	1334 1943 3277	1334 1977 3312	1334 2060 3395	1334 2306 3640	1334 2128 3463	1334 3298 4632	1334 4751 6085
.60	2291 1930 4220	2291 1935 4226	2291 1949 4240	2291 1974 4264	2291 2045 4336	2291 2316 4607	2291 2147 4438	2291 3347 5638	2291 4758 7049

§7 Simulatie resultaten.

De simulatie resultaten in tabel I en tabel II spreken vrijwel voor zichzelf.

De genoemde percentage in tabel I zijn percentages van het totaal aantal begin (eind) knuppels, dus niet percentages genomen over de som van begin en eind knuppels.

Een nadere verklaring is voorts ook nodig voor het constant zijn van het verlies voor kleinere waarden van b bij constante a (Tabel II): Dit laat zich als volgt verklaren:

De verdeling van het DY $\underline{y}$ in z is $\mathcal{N}(13 \text{ cm}, 2 \text{ cm})$.

Onderstel eenvoudigheidshalve de verlengingsfactor n constant dan zou de verdeling van het DY in $\underline{y}$ ook normaal zijn.

Als $n = 5$ dan zou de verwachting van het DY in $\underline{y}$ 65 cm zijn en de spreiding 10 cm. Dit betekent dat men in het simulatie model bijna altijd een stuk DY van minstens 45 cm aan het eind van de totale lengte vindt.

Het afkappen van b.v. 40 cm. in plaats van 20 cm. DY verandert het verlies aan het eind niet.

Uiteraard is het niet zeker dat dit verschijnsel voor kleinere waarden van b ook in de praktijk of in een nauwkeuriger gekozen simulatie model optreedt.

Opmerking: Er zij op gewezen dat hoewel de gemeten lengte $\underline{l}$ kleiner is naarmate het afgekapte stuk b groter wordt, dit de knipstrategie hier nauwelijks beïnvloed daar deze strategie gecorrigeerd wordt voor de aanwezigheid van DY aan het eind met een correctieterm $\epsilon(b)$ die in dit geval de verandering van b vrijwel precies te niet doet. (zie §6).

Voor andere verdelingen en andere waarden voor de kostenconstanten kunnen de tabellen I en II wel ingrijpend zijn gewijzigd.

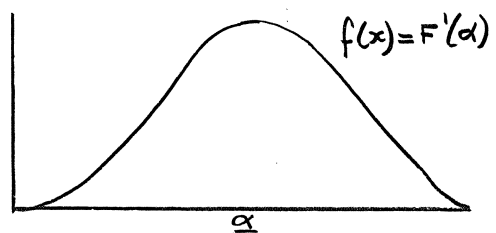
Uit deze simulatie resultaten kunnen daarom geen conclusies voor de praktijk worden getrokken.

§8 Mogelijke analytische bepaling van het verlies in de beginknuppel als functie van a.

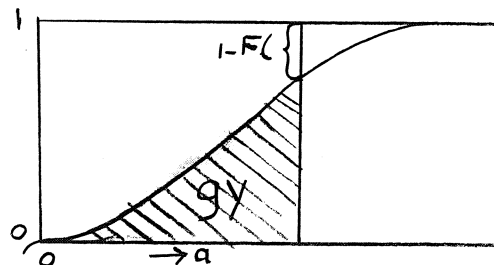
Het toepassen van hetgeen in deze paragraaf staat is zeker niet ondoenlijk, maar daar er voor het verlies aan het eind toch gesimuleerd moest worden en de oplossingsmethode er overzichtelijker door werd is ook het verlies in de beginknuppel als functie van a door simulatie onderzocht.

Strikt genomen is dit echter een onnodige vereenvoudiging van het model. Vandaar dat deze paragraaf aan de scriptie is toegevoegd α is de lengte van het DY in de bloem.

Deze α heeft bijvoorbeeld de nevenstaande verdelingsdichtheid $f(\alpha)$.



$F(a)$ is de kans op het aantreffen van een waarde van α met $\alpha \leq a$ dus op het aantreffen van GY op een afstand a van het uiteinde van de bloem.



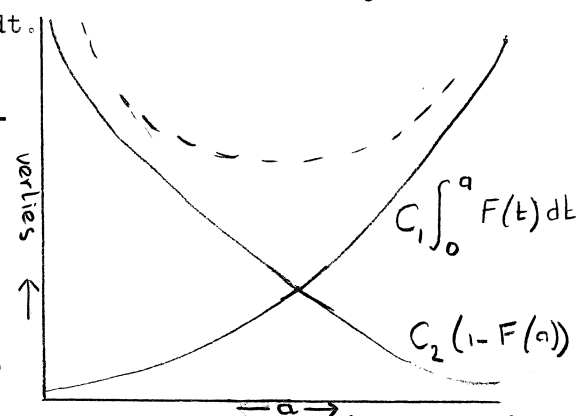
De kans op DY is daar dan $1 - F(a)$.

De verwachte afbrandkosten per

knuppel zijn dus $C_2(1 - F(a))$. C_1 zij de verlieskosten bij het verloren gaan van 1 ton GY. Stel c_1 zijn de verlieskosten per meter GY in de bloem. De hoeveelheid afgekapte GY bij kniplengte a is dan $\int_0^a F(t)dt$.

Het verlies aan GY is dan $c_1 \int_0^a F(t)dt$.

Wanneer alleen de bovengenoemde twee factoren meespelen - en dus geen proef 13 knuppels aan het begin verschijnen dan is de optimale waarde van a te bepalen op de in nevenstaande grafiek gesuggereerde manier.



Wanneer er wel proef 13 knuppels of zelfs afkeur knuppels in het begin kunnen optreden dan moet men ook met deze factoren rekening trachten te houden.

In de uitgevoerde simulatie vormt het eventuele optreden van proef 13 of afkeurknuppels nog minder een probleem.

490.91