OM MEDELFELETS HÄRLEDNING VID LINJE- OCH PROVYTE-TAXERING

ON COMPUTING THE STANDARD ERROR IN LINE AND SAMPLE PLOT SURVEYING

AV

MANFRED NÄSLUND

MEDDELANDEN FRÅN STATENS SKOGSFÖRSÖKSANSTALT
HÄFTE 31 · N:r 7

Centraltr., Esselte, Sthlm 39
942857

MEDDELANDEN

FRÅN

STATENS SKOGSFÖRSÖKSANSTALT

HÄFTE 31. 1938—39

MITTEILUNGEN AUS DER FORSTLICHEN VERSUCHS-ANSTALT SCHWEDENS

31. HEFT

REPORTS OF THE SWEDISH INSTITUTE OF EXPERIMENTAL FORESTRY

N:o 31

BULLETIN DE L'INSTITUT D'EXPÉRIMENTATION FORESTIÈRE DE SUÈDE

N:o 31

CENTRALTRYCKERIET, ESSELTE, STOCKHOLM 1939

REDAKTÖR:
PROFESSOR DR HENRIK HESSELMAN

INNEHÅLL:

MANFRED NÄSLUND

OM MEDELFELETS HÄRLEDNING VID LINJE- OCH PROVYTETAXERING.

Inledning.

Sedan vid värmlandstaxeringen den tanken blivit väckt att använda sannolikhetskalkylens hjälpmedel för bestämningen av noggrannheten hos en objektiv taxering (*Kommissionen för försökstaxering* etc. 1914), ha ett flertal metoder för härledning av en taxerings medelfel framlagts. Tillförlitliga sådana beräkningsmetoder utgöra en viktig förutsättning för en rationell planläggning av våra skogstaxeringar. De framkomna metoderna ha tidigare underkastats kritiska granskningar (LINDEBERG 1923, 1926, LANGSAETER 1926, 1927, 1932, NÄSLUND 1930), varav bl. a. framgått, att ett starkt behov av metodikens vidare utveckling förelåg. Avsikten är här att lämna ett bidrag till denna fråga, som i samband med de förnyade riksskogstaxeringarna i de nordiska länderna erhållit ökad aktualitet.

Den tidigare diskussionen av metodiken har huvudsakligen skett i anslutning till linjetaxeringen, men vi skola nu även beröra metodfrågan för den regelbundna provytetaxeringen, som tilldragit sig ett stort intresse, och för vars objektiva bedömande en tillfredsställande metod vid medelfelets beräkning erfordras.

Den i det efterföljande framlagda metoden för medelfelets härledning har använts av mig vid vissa utredningar på uppdrag av *1937 års riksskogstaxeringsnämnd*. Med nämndens benägna tillstånd illustreras framställningen av exempel hämtade från denna undersökning. För detta tillmötesgående ber jag att få uttala mitt tack.

Jägmästare E. ÖSTLIN har omhänderhaft den närmaste ledningen av det omfattande räknearbetet för nämndens räkning, och förste skogsbiträdet K. SVENSON har utfört räknearbetet vid det förberedande metodstudiet. Till dessa medhjälpare vill jag rikta ett hjärtligt tack.

KAP. I. DEN SENASTE UTVECKLINGEN AV METODIKEN VID MEDELFELETS HÄRLEDNING.

Vi skola här endast uppehålla oss vid den utveckling, föreliggande fråga haft, sedan jag förra gången behandlade detta ämne (NÄSLUND 1930).

LANGSAETER har utfört en omfattande undersökning, vars syfte varit att genom bearbetning av empiriskt taxeringsmaterial lämna ett bidrag till frågan om den noggrannhet, som uppnås vid linjetaxering av skogar av olika storlek och sammansättning med användande av olika taxeringsprocenter (LANGSAETER 1932). I denna avhandling diskuterar LANGSAETER åter ingående metodiken vid medelfelets härledning.

LANGSAETER framlägger i avhandlingens teoretiska del bl. a. en ny metod för medelfelets beräkning (LANGSAETER 1932, s. 461, formel 21). I denna formel ingå differenser av den observerade storheten för linjeavstånd ända ned till 10 m. Utan ett extrapoleringsförfarande kan formeln ej användas i praktiken, och som stöd för denna extrapolering saknas ett entydigt, funktionellt samband. Metoden har därför huvudsakligen ett teoretiskt intresse, vilket LANGSAETER själv framhåller.

I samband med en allmän prövning av de vanligaste formlerna för medelfelets beräkning behandlar LANGSAETER även en av mig framlagd metod (jfr NÄSLUND 1930, s. 330, metod D: 2). Vid tillämpningen av densamma använder LANGSAETER delvis ett annat förfaringssätt, än det jag anvisat. Tillvägagångssättet överensstämmer ej med den ursprungliga tankegång, varpå metoden vilar. Jag skall därför något uppehålla mig härvid.

Medeldifferensens utjämning avser i princip endast att eliminera den genomsnittliga effekten av en ev. systematisk variation (tendens) från linje till linje hos den taxerade storheten. Någon annan innebörd har ej givits åt differenskurvans extrapolering till noll. Till grund för utjämningen skall därför endast läggas de medeldifferenser, som kunna bildas av i taxeringen ingående linjer.

När LANGSAETER för mycket lågprocentiga taxeringar härleder den aktuella taxeringens medeldifferens med ledning av medeldifferenserna för 10 och 20 m:s linjeavstånd, innebär detta i princip något helt annat och tyder på en felaktig tolkning av metodens teoretiska underlag (LANGSAETER 1932, s. 483). För vissa jämförelser har jag (NÄSLUND 1930, tab. 11, s. 335) beräknat 5- och 2½-procents taxeringarnas medelfel med stöd av 10-procents taxeringens medeldifferens. Detta är emellertid ej någon tillämpning av metoden utan ett approximationsförfarande, som tillgripits i brist på andra utvägar och har karaktären av en relativt måttlig extrapolering.

LANGSAETER visar vidare, att min metod under vissa förutsättningar ger samma resultat som hans metod C: 2 (jfr LANGSAETER 1932, s. 471), vilken då har fördelen att vara mindre arbetskrävande. Härtill skall blott anmärkas, att detta bl. a. förutsätter medeldifferenskurvans utjämning med stöd av endast två punkter, vilket givetvis ej varit avsikten (jfr NÄSLUND 1930, s. 332, fig. 6).

Den ovan diskuterade metoden har framlagts av mig i anslutning till en granskning av metodiken vid medelfelets härledning och var i detta sammanhang klarläggande och av teoretiskt intresse. Metoden har emellertid en begränsad praktisk användbarhet, vilket jag tidigare framhållit (NÄSLUND 1930, s. 330), varför vi här ej närmare skola uppehålla oss vid densamma.

I Sverige har ÖSTLIND för Domänstyrelsens räkning utfört en omfattande utredning med samma syfte som LANGSAETERS ovannämnda undersökning (ÖSTLIND 1932). Här diskuterar ÖSTLIND teorien för medelfelets härledning vid linjetaxering och deducerar i anslutning till LANGSAETER en formel för medelfelet, vilken kan betraktas som en modifikation eller variant av LANGSAETERS på andra ordningens differenser uppbyggda formel (LANGSAETER 1926, 1927 och 1930).

I U. S. A. har den regelbundna provytetaxeringen varit föremål för studier i här avsett syfte (SCHUMACHER och BULL 1932, MUDGETT och GEVORKIANTZ 1934). Härvid har medelfelets härledning diskuterats i anslutning till BERNOULLIS, POISSONS och LEXIS teorem, vartill vi återkomma i kap. III, s. 320. Vi kalla i det efterföljande den regelbundna provytetaxeringen enbart för provytetaxering, och avse därmed en taxering bestående av ett antal mindre provytor, fördelade i ett visst, regelbundet förband.

Härmed ha omnämnts de viktigaste, kända arbeten, som efter år 1930 behandlat frågan om metodiken för medelfelets härledning vid regelbunden såväl linje- som provytetaxering. Det torde ha framgått, att vi, då det gäller beräkningen av medelfelet hos en aktuell linjetaxering, fortfarande äro hänvisade till de metoder, som förelågo vid min tidigare behandling av detta ämne, vartill hänvisas för undvikande av upprepning (NÄSLUND 1930). Däremot ha våra erfarenheter om linjetaxeringens noggrannhet väsentligt ökats genom omfattande bearbetningar av empiriskt taxeringsmaterial (LANGSAETER 1932, *Riksskogstaxeringsnämnden* 1932, ÖSTLIND 1932).

KAP. II. ALLMÄNNA SYNPUNKTER PÅ METODFRÅGAN.

Av de framlagda metoderna är det huvudsakligen två, som hos oss kommit till allmännare, praktisk användning för beräkningen av medelfelet hos en aktuell taxering. Dessa metoder ha i min föregående uppsats benämnts resp. metod A och metod C (NÄSLUND 1930), vilka beteckningar bibehållas i det efterföljande.

Metod A. Taxeringslinjerna eller delar därav indelas i grupper (deltaxeringar), varefter medelfelet (ε_M) beräknas enligt formeln:

$$\varepsilon_M = \sqrt{\frac{n \cdot \sum_{i=1}^{n} (p_i)^2 (x_i - M)^2}{(n-1)\left(\sum_{i=1}^{n} p_i\right)^2}}, \dots\dots\dots\dots\dots (1)$$

där $M = \dfrac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}$;

x_i betecknar i:te gruppens taxeringsresultat och p_i dess linjelängd samt n antalet grupper.

Metod C. Medelfelet beräknas med stöd av differenserna mellan taxeringsresultaten från på varandra följande taxeringslinjer. Medelfelsformler enligt denna princip äro framlagda av ÖSTLIND-HAGSTRÖM (1922), LINDEBERG (1923, 1926) och LANGSAETER (1926, 1927, 1932). Ofta har härvid LANGSAETERS på andra ordningens differenser uppbyggda formel föredragits (LANGSAETER 1932, metod C: 2). Medelfelet erhålles då enligt följande formel:

$$\varepsilon_M = \sqrt{\frac{n \cdot \sum_{i=2}^{n-1} (p_i)^2 (x_{i-1} + x_{i+1} - 2\,x_i)^2}{6\,(n-2)\left(\sum_{i=1}^{n} p_i\right)^2}}; \dots\dots\dots (2)$$

där x_i är taxeringsresultatet och p_i taxerad längd för i: te taxeringslinjen samt n antalet taxeringslinjer. i anger linjernas ordningsföljd.

Stundom har en kombination av dessa metoder kommit till användning därigenom, att linjerna först indelats i grupper, varefter formel (2) tillämpats på gruppresultaten. Vi skola med dessa metoder som bakgrund närmare diskutera den föreliggande problemställningen.

De beräknade medelfelen äro i sin tur behäftade med medelfel på grund av det begränsade antalet varianter (deltaxeringar, resp. taxeringslinjer). D. v. s. om man gör om taxeringen med samma linjeavstånd, men förlägger taxeringslinjerna mellan de ursprungliga linjerna, erhålles icke endast ett avvikande taxeringsresultat utan även ett annat medelfel.

Vid tillämpning av metod A och formel (1) kan medelfelet på medelfelet ($\varepsilon_{\varepsilon_M}$) erhållas ur formeln:

$$\varepsilon_{\varepsilon_M} = \frac{\varepsilon_M}{\sqrt{2\,(n-1)}}, \dots\dots\dots\dots\dots\dots\dots (3)$$

där n är antalet grupper (deltaxeringar). Formel (3) förutsätter, att avvikelserna ($x_i - M$) fördela sig enligt den normala sannolikhetsfunktionen (HELMERT 1924, s. 75). Medelfelets relativa medelfel vid olika gruppantal (n) framgår av nedanstående sammanställning:

$n =$	2	3	4	5	6	7	8	9
$\varepsilon_{\varepsilon_M} =$	70,7 %,	50,0 %,	40,8 %,	35,4 %,	31,6 %,	28,9 %,	26,7 %,	25,0 %,

$n =$	10	15	20	25	30	40	50
$\varepsilon_{\varepsilon_M} =$	23,6 %,	18,9 %,	16,2 %,	14,4 %,	13,1 %,	11,3 %,	10,1 %.

Osäkerheten i de beräknade medelfelen är betydande för de gruppantal, som hittills vanligen kommit till användning (8—20 st.). Gäller det att beräkna medelfelet för en enskild taxering, måste denna således uppdelas i ett relativt stort antal deltaxeringar (grupper). Här komma vi till en allvarlig begränsning hos metod A.

Vid en uppdelning av taxeringen i ett större antal grupper riskerar man ofta, såvida antalet linjer eller delsträckor ej är mycket stort, att en systematisk gång gör sig gällande i gruppresultaten (jfr LANGSAETER 1926, NÄSLUND 1930). Under sådana förhållanden ger formel (1) för höga medelfel. Minskningen av det tillfälliga felet köpes sålunda med ökad risk för systematiskt fel.

Det systematiska felet kan emellertid nedbringas genom att på deltaxeringarna tillämpa formel (2), men inflytandet av den systematiska gången från grupp till grupp kvarstår alltid. Risken för ovannämnda systematiska tendens är dock betydligt mindre, om taxeringslinjerna uppdelas i delsträckor, som kombineras efter ett visst system (jfr NÄSLUND 1930, metod A: 2, s. 321). Bortsett från ovannämnda risk för ett systematiskt för högt medelfel är metoden även ur en annan synpunkt otillfredsställande.

Gemensamt för metod A är, att man ur deltaxeringarnas medelavvikelse beräknar den aktuella taxeringens medelfel (ε_M). En sådan extrapolation stöter alltid på svårigheter och måste betraktas som en svaghet hos metoden. Vi återkomma härtill i det efterföljande (s. 315).

Metod C torde ha framtvingats av behovet att vid metod A uppdela den aktuella taxeringen i ett flertal deltaxeringar. De i en taxering ingående linjerna uppgå nämligen ej sällan till ett så ringa antal, att man ej ens kan uppdela dem i ett mindre antal deltaxeringar.

För metod C och formel (2) kan medelfelets relativa medelfel vid ett linjeantal av omkring 10 och däröver approximativt uppskattas till samma storlek som för metod A och formel (1), varvid n (formel 3) nu betyder antalet taxeringslinjer. Metod C lämnar därför ofta ett med hänsyn till de tillfälliga felen relativt säkert medelfel. För taxeringar av småskogar föreligger dock ett behov att kunna dela upp taxeringslinjerna i mindre enheter (jfr s. 320). Metoden har emellertid en annan svaghet.

I regel kan man spåra en systematisk gång i taxeringsresultaten för de enskilda linjerna (LANGSAETER 1926, NÄSLUND 1930, ÖSTLIND 1932, jfr även fig. 1—6, s. 308—13). Under sådana förhållanden ger formel (2) för högt resultat, emedan den systematiska tendensen från linje till linje ej borteliminer as. Enligt ÖSTLIND torde formel (2) beträffande kubikmassan per hektar skogsproduktiv mark i genomsnitt ge omkring 20 procent för stora medelfel (*Riksskogstaxeringsnämnden* 1932, ÖSTLIND 1932). LANGSAETER uppskattar detta systematiska fel till 10 à 20 procent (*Landsskogtakseringen* 1938).

För vissa behov kan man givetvis bortse ifrån en överskattning av denna storleksordning och nöja sig med att vara på den säkra sidan. Gäller det däremot att beräkna medelfelet på skillnaden mellan två taxeringsresultat, vilket f. n. i samband med den förnyade riksskogstaxeringen är en aktuell fråga, ökas kravet på bestämningen av den enskilda taxeringens medelfel. Vi skola närmare exemplifiera detta.

Antaga vi, att exempelvis virkesförrådet vid första taxeringen är M och vid den andra $k \cdot M$, samt medelfelet vid båda uppskattningarna ε_M procent, erhålles ett enkelt uttryck för differensens relativa medelfel (ε_D). Härvid skilja vi på om förrådet minskar eller ökar.

Tab. 1. **Medelfelet på differenser (ε_D).**
The standard error of differences.

k	ε_M i procent. in percentage.								
	1	1,5	2	2,5	3	3,5	4	4,5	5
	ε_D i procent. Minskning. in percentage. Decrease.								
0,9	13,4	20,2	26,9	33,6	40,4	47,1	53,8	60,5	67,3
0,8	6,4	9,6	12,8	16,0	19,2	22,4	25,6	28,8	32,0
0,7	4,1	6,1	8,1	10,2	12,2	14,2	16,3	18,3	20,3
0,6	2,9	4,4	5,8	7,3	8,8	10,2	11,7	13,2	14,6
0,5	2,2	3,4	4,5	5,6	6,7	7,8	9,0	10,1	11,2
	ε_D i procent. Ökning. in percentage. Increase.								
1,1	14,9	22,3	29,7	37,2	44,6	52,0	59,5	66,9	74,3
1,2	7,8	11,7	15,6	19,5	23,4	27,3	31,2	35,1	39,0
1,3	5,5	8,2	10,9	13,7	16,4	19,1	21,9	24,6	27,3
1,4	4,3	6,4	8,6	10,8	12,9	15,1	17,2	19,4	21,5
1,5	3,6	5,4	7,2	9,0	10,8	12,6	14,4	16,2	18,0

Virkesförrådet minskar ($k < 1$):

$$\varepsilon_D = \frac{\varepsilon_M \sqrt{1 + k^2}}{1 - k}, \qquad (4)$$

Virkesförrådet ökar ($k > 1$):

$$\varepsilon_D = \frac{\varepsilon_M \sqrt{1 + k^2}}{k - 1}, \qquad (5)$$

där ε_D och ε_M äro uttryckta i procent.

Formlerna (4) och (5) ha tabellerats för vissa värden på k och ε_M (tab. 1). Härav framgår med skärpa svårigheten att uppskatta små förändringar samt betydelsen av en god bestämning av de enskilda taxeringarnas medelfel.

Den förda diskussionen torde ha visat, att man vid ett sökande efter en lämplig metod för medelfelets härledning måste uppställa som krav, att medelfelets såväl systematiska som tillfälliga fel vid behov skola kunna nedbringas så långt som möjligt.

KAP. III. TEORI OCH METODIK.

Linjetaxering.

De svårigheter, som äro förbundna med beräkningen av medelfelet för en regelbunden linje- eller provytetaxering, bero huvudsakligen på att taxeringslinjerna utläggas systematiskt och ej uttagas på slump. När linjeriktningen, taxeringsprocenten och läget av första taxeringslinjen bestämts, är även belägenheten av de övriga taxeringslinjerna angiven. Linjernas systematiska utläggande sker uppenbarligen i syfte att öka taxeringsresultatens säkerhet. Att så även är fallet framgår av ÖSTLINDS ovannämnda utredning (ÖSTLIND 1932, s. 451).

Det har tidigare framhållits, att taxeringsresultatet ofta varierar systematiskt från linje till linje (jfr. s. 306), och vi skola här i några exempel närmare analysera taxeringens struktur. För detta syfte har valts förra riksskogstaxeringen av Västernorrlands och Norrbottens län, vilka ur taxeringssynpunkt representera olika svårighetsgrader.

Vid riksskogstaxeringen ha taxeringslinjerna indelats i 2 km långa sträckor. Denna indelning är utförd så, att sträckor med samma 2-km nummer bilda ett mot taxeringsriktningen vinkelrätt fält (block). I det efterföljande äro 2-kmna sammanslagna till 1-mil sträckor. För 1-mil sträckor med samma nummer införa vi benämningen taxeringsblock eller enbart block.

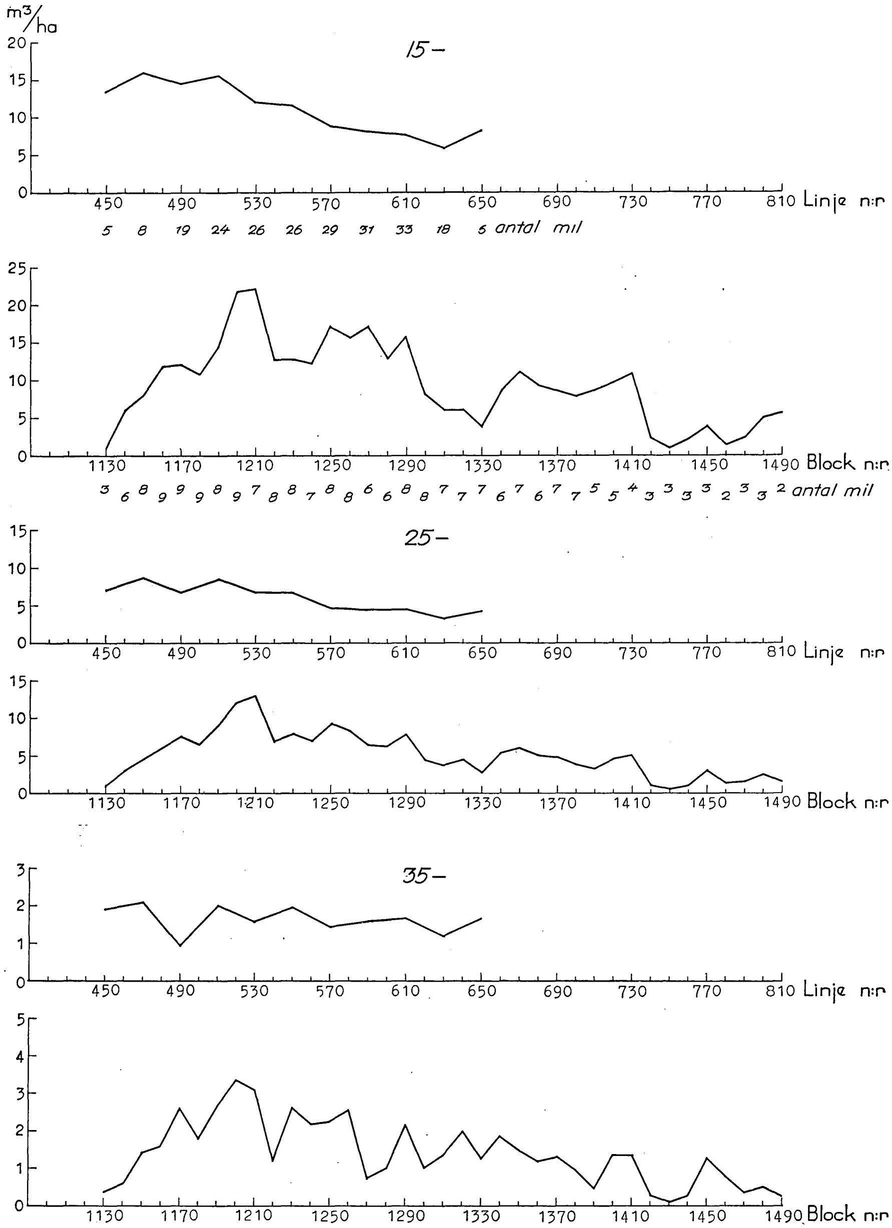

Fig. 1. Kubikmassa per hektar för taxeringslinjer och block. Antal mil avser antalet taxerade 1-mil sträckor, som ingå i resp. linje eller block. Tall: dimensionsgrupperna 15,— 25— och 35—. Norrbottens lappmark.

Cubic volume per hectare (m^3/ha) of the lines and blocks. Number of miles refers to the number of surveyed 10 km stretches, of which respective line or block consists. Pine: the diameter-classes 15—, 25— and 35—. The Province of Norrbotten: Lappmark.

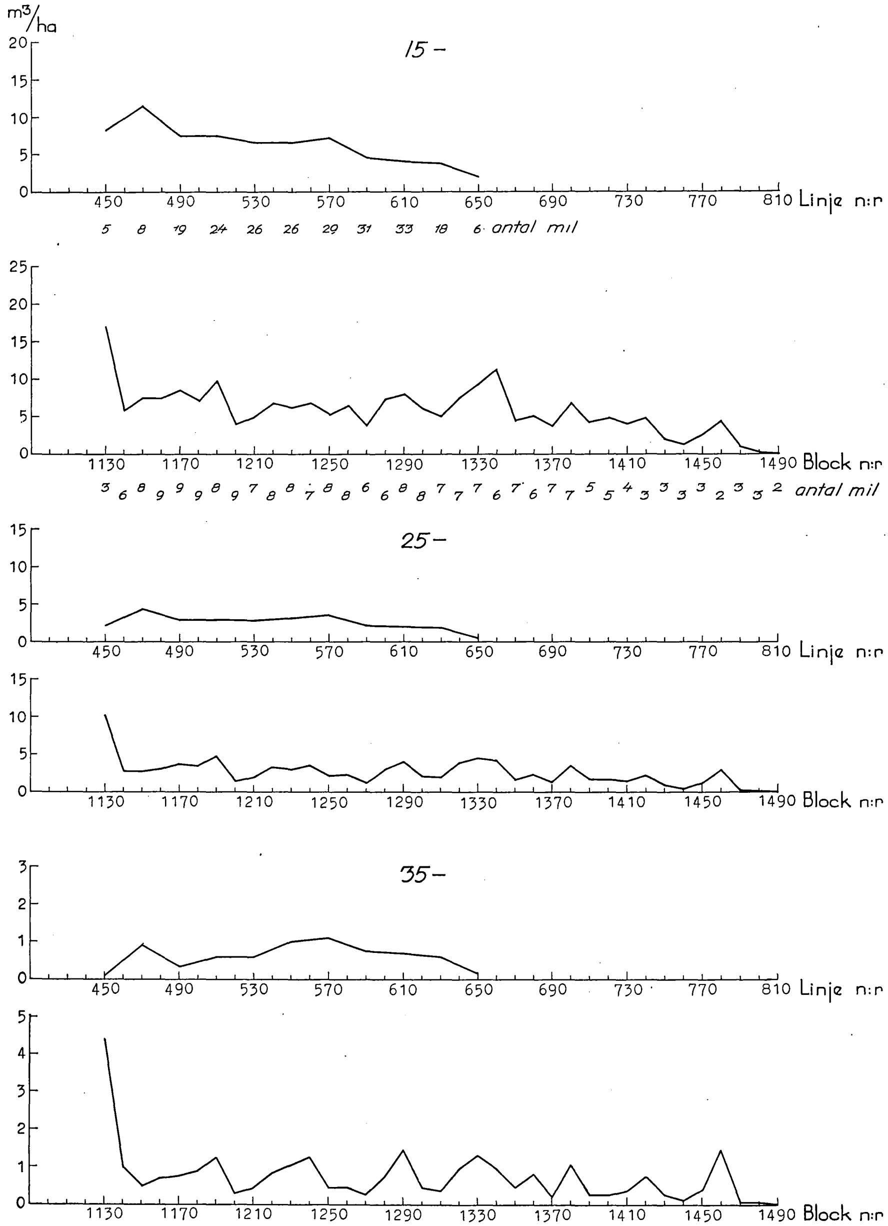

Fig. 2. Se texten till fig. 1. Gran: dimensionsgrupperna 15—, 25— och 35—. Norrbottens lappmark.

Cfr. caption to fig. 1. Spruce: the diameter-classes 15—, 25— and 35—. The Province of Norrbotten: Lappmark.

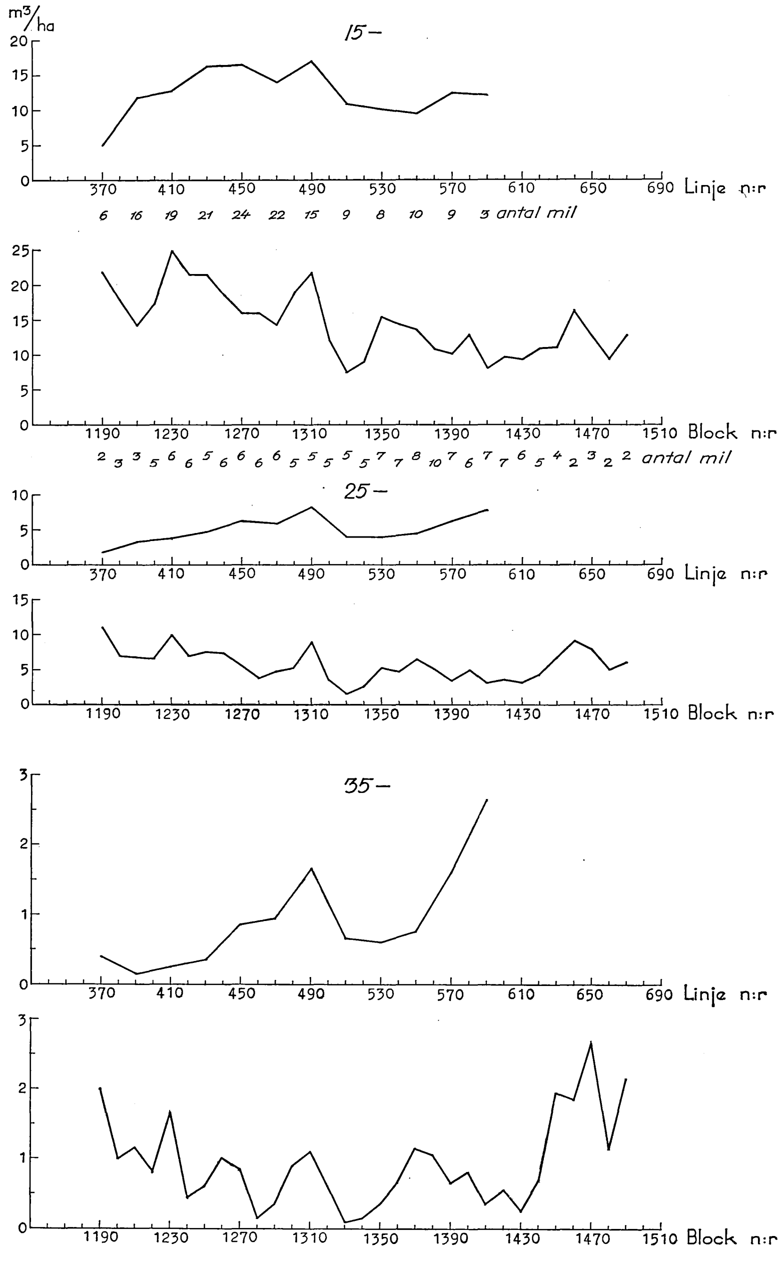

Fig. 3. Se texten till fig. 1. Tall: dimensionsgrupperna 15—, 25— och 35—. Norrbottens kustland.

Cfr. caption to fig. 1. Pine: the diameter-classes 15—, 25— and 35—. The Province of Norrbotten: Kustland.

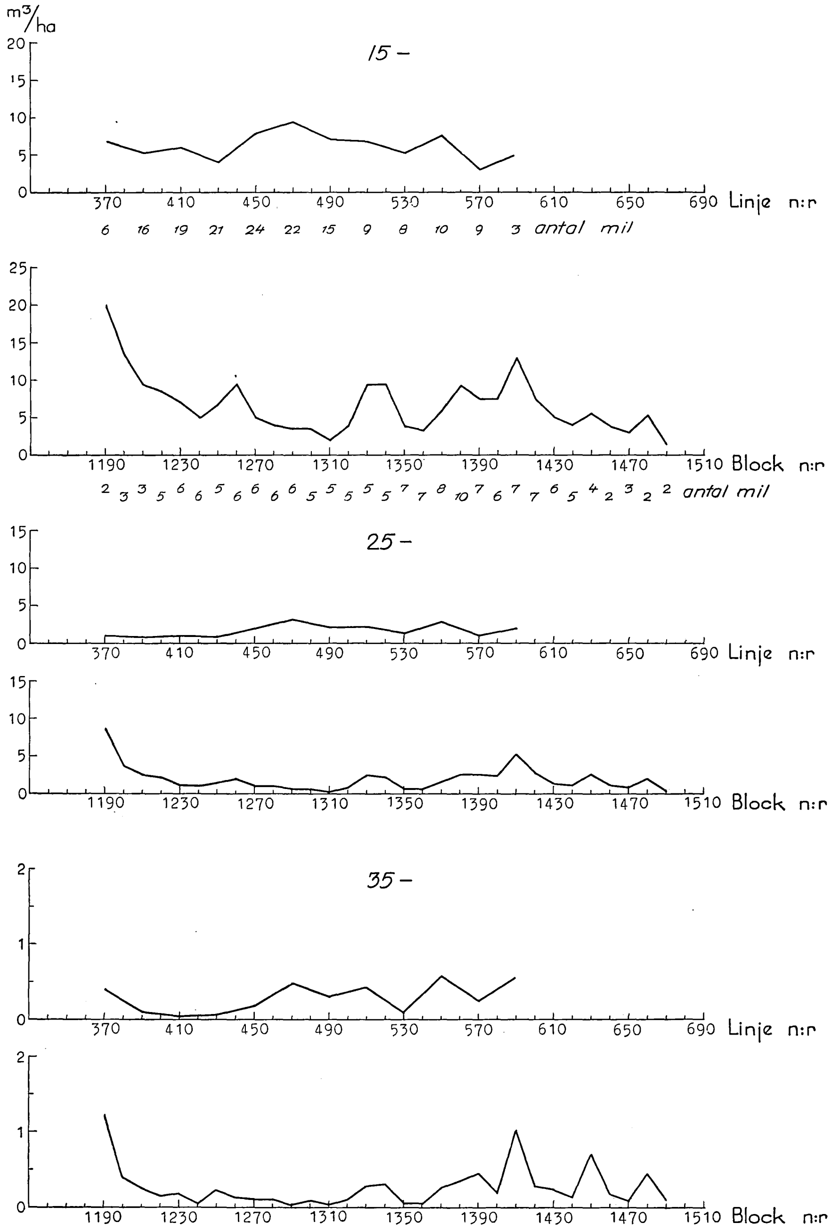

Fig. 4. Se texten till fig. 1. Gran: dimensionsgrupperna 15—, 25— och 35—. Norrbottens kustland.

Cfr. caption to fig. 1. Spruce: the diameter-classes 15—, 25— and 35—. The Province of Norrbotten: Kustland.

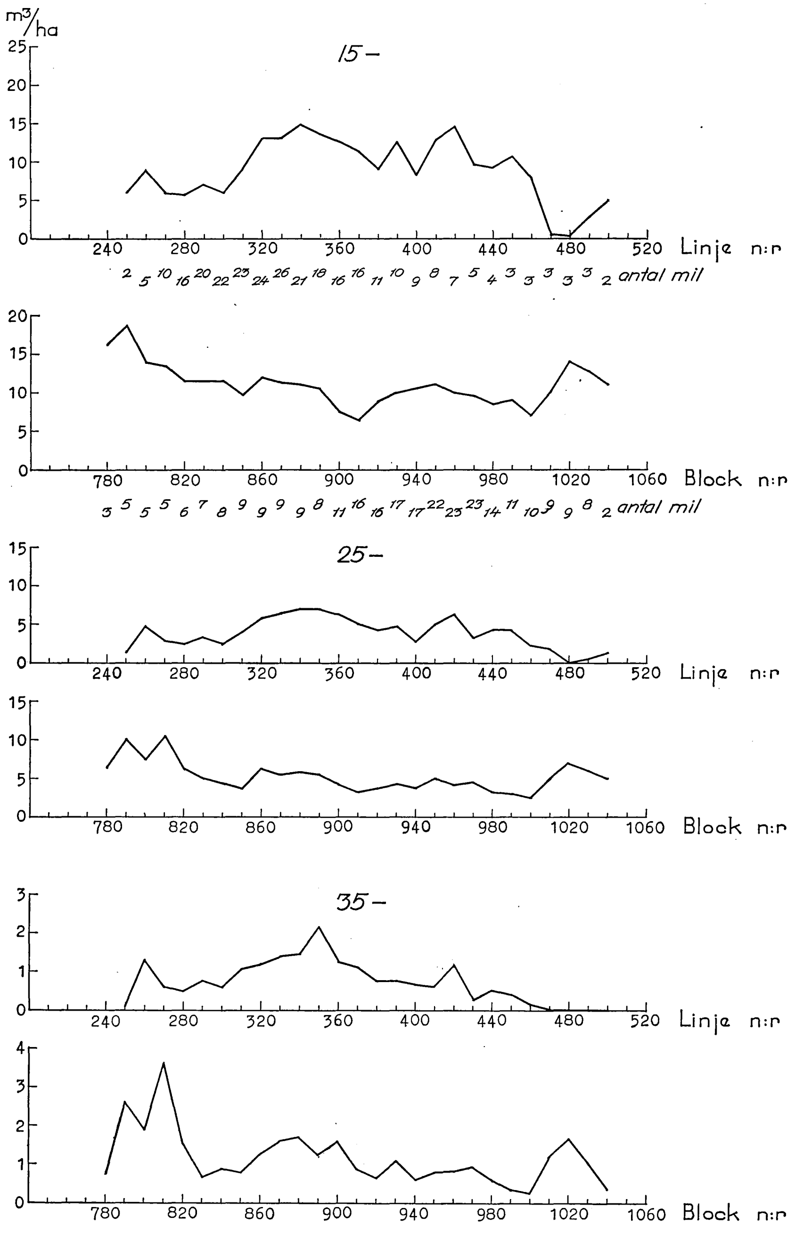

Fig. 5. Se texten till fig. 1. Tall: dimensionsgrupperna 15—, 25— och 35—. Västernorrlands län.

Cfr. caption to fig. 1. Pine: the diameter-classes 15—, 25— and 35—. The Province of Västernorrland.

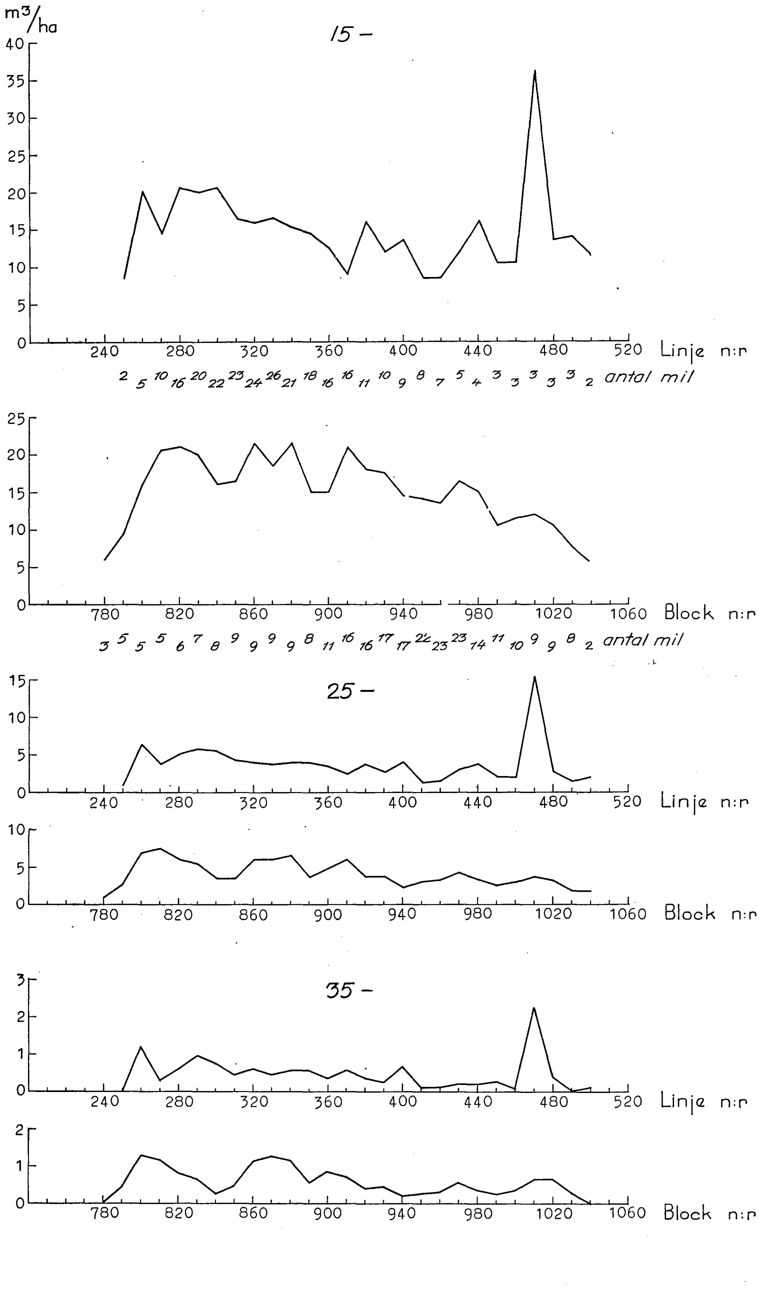

Fig. 6. Se texten till fig. 1. Gran: dimensionsgrupperna 15—, 25— och 35—. Västernorrlands län.

Cfr. caption to fig. 1. Spruce: the diameter-classes 15—, 25— and 35. The Province of Västernorrland.

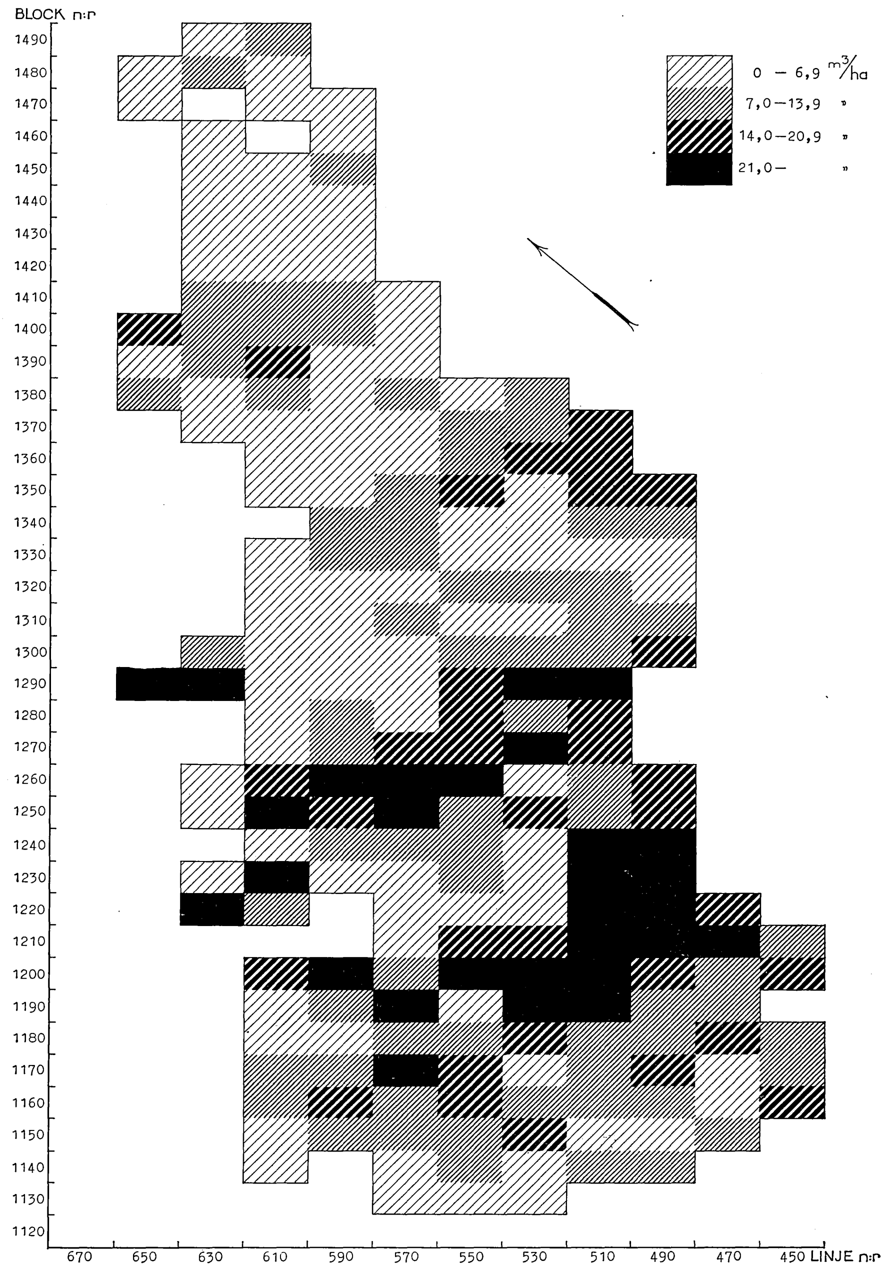

Fig. 7. Norrbottens lappmark. Fördelningen av granens kubikmassa i dimensionsgrupp 15— enligt taxeringsresultaten från 1-mil sträckorna.
The Province of Norrbotten: Lappmark. The distribution of cubic volume of spruce in diameter-class 15— according to the results of the survey from the 10 km stretches.

Kubikmassan inom bark per hektar landareal (Västernorrland) eller landareal exkl. inägor (Norrbotten) har uträknats för varje linje och block inom de båda länen, varvid Norrbotten uppdelats på kustland och lappmark. Kubikmassan har redovisats på tall och gran samt dimensionsgrupperna: 15 cm vid brösthöjd och däröver, 25 cm och däröver samt 35 cm och däröver. I fig. 1—6 har kubikmassan per hektar upplagts grafiskt med linje-, resp. blocknumret som oberoende variabel. Linjerna äro numrerade från sydost till nordväst och blocken från sydväst till nordost. Linjeavståndet utgör för Västernorrlands län en mil och för Norrbottens två mil.

Figurerna visa tydligt, att för såväl linjer som block en markerad systematisk gång ofta gör sig gällande i taxeringsresultaten. Denna tendens är ej sällan mycket starkt utpräglad. Härtill bidrar givetvis att tallens och granens kubikmassor angivits skilda åt, vilket är motiverat därav, att det föreligger ett behov att kunna beräkna säkerheten hos dessa uppskattningar var för sig.

Beträffande arealbestämningar har LANGSAETER påvisat samma tendenser (LANGSAETER 1926).

För Norrbottens lappmark har i fig. 7 fördelningen av granens kubikmassa över 15 cm närmare åskådliggjorts med ledning av taxeringsresultaten från 1-mil sträckorna. Av figuren framgår, att kubikmassan å en linjesträcka i viss grad är en funktion av sträckans avstånd till två koordinataxlar. Det är uppenbart, att vid medelfelsmetodikens utformning stor hänsyn måste tagas till förekomsten av här illustrerade fördelningstendenser.

Innan vi övergå till att diskutera medelfelsberäkningens teori, skola vi något dröja vid de metoder, som grunda sig på att taxeringslinjerna eller delar därav sammanföras i grupper (jfr NÄSLUND 1930, metod A: 1 och 2).

Vi förutsätta, att den utförda linjetaxeringen är en 10-procents taxering, och att den är uppdelad i 10 grupper, som var och en omfattar en 1-procents taxering. Det efter denna metod beräknade medelfelet är då approximativt det riktiga medelfelet för den taxering, som erhålles, om man av de 100 stycken möjliga 1-procents taxeringarna av skogen på slump plockar ut 10 stycken och därav bildar en 10-procents taxering (jfr LANGSAETER 1932, s. 446). Det är emellertid ej detta medelfel, vi önska veta, utan medelfelet på den systematiskt utlagda 10-procentiga linjetaxeringen.

Med taxeringslinjernas systematiska utläggande avses, att varje linje eller del därav (exv. 1-mil sträcka) skall representera den parcell av skogen, som ligger inom halva linjeavståndet på vardera sidan om linjen (delsträckan). Kände vi resultatet ($\bar{x}_i$) av en totaltaxering för varje sådan avdelning, skulle linjetaxeringens medelfel (ε_M) erhållas enligt formlerna:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}{n}}, \quad \ldots \ldots \ldots \ldots \ldots \ldots \quad (6)$$

$$\varepsilon_M = \frac{\sigma}{\sqrt{n}}, \quad \ldots \ldots \ldots \ldots \ldots \ldots \quad (7)$$

där x_i är linjetaxeringens resultat för i:te linjen eller delsträckan och n är antalet linjer eller delsträckor, vilka alla förutsättas vara lika långa.

Nu känna vi ej $\bar{x}_i$ utan måste ersätta denna storhet med ett närmevärde. Vid den fortsatta diskussionen tänka vi oss linjerna uppdelade i delsträckor på sätt, som tidigare nämnts, varigenom linjetaxeringen blir omförd till ett kvadratiskt eller rektangulärt rutnät (jfr fig. 7). Samtidigt införa vi beteckningarna $\bar{x}_{uv}$ och x_{uv} och avse därmed den undersökta storhetens verkliga, resp. taxerade värde för den ruta, som ligger i linje u och block v.

Vi ha i det föregående visat, att exv. kubikmassan för en ruta i stora drag kan betraktas som en funktion av ytans belägenhet. Det enklaste antagande vi kunna göra angående detta samband är, att den undersökta storheten på en ruta kan betraktas som summan av tre storheter: 1) en del α, som är lika för samtliga rutor, 2) en del β_u, som är konstant för varje linje samt 3) en del γ_v, som är konstant för varje block. Härvid äro β_u och γ_v så valda, att summan av alla β-värden är lika med noll, likaså summan av samtliga γ-värden, Under dessa förutsättningar blir kubikmassan för en viss ruta: $\bar{x}_{uv} = \alpha + \beta_u + \gamma_v$. Det gäller att ur taxeringsmaterialet härleda de bästa värdena på α, β_u och γ_v. Detta erhålles tydligen, om man sätter: $\alpha = M$, $\beta_u = m_u - M$, $\gamma_v = m_v - M$, där M betecknar medeltalet av samtliga rutor, m_u medeltalet för linjen u och m_v medeltalet för blocket v. Härav erhålles:

$$\bar{x}_{uv} = m_u + m_v - M, \quad \ldots \ldots \ldots \ldots \ldots \ldots \quad (8)$$

och alltså avvikelsen mellan den undersökta storhetens taxerade och verkliga värde: $x_{uv} - m_u - m_v + M$ samt avvikelsernas kvadratsumma:

$$\sum_{u,v} (x_{uv} - \bar{x}_{uv})^2 = \sum_{u,v} (x_{uv} - m_u - m_v + M)^2 \quad \ldots \ldots \ldots \quad (9)$$

Summationstecknet $\sum_{u,v}$ betyder summering över samtliga förekommande värden på u och v. Denna kvadratsumma är således lika med de enskilda rutornas avvikelser från »stormedeltalet» (M), sedan dessa korrigerats för de systematiska avvikelser, som härleda från belägenheten i olika linjer och block.

Det problem, som här förelegat, är analogt med jordbruksforskningens behov att vid fältförsöken kunna borteliminera inverkan av olikheter i markens bördighet. Den tankegång, som utvecklats i det ovanstående har också gemensamma beröringspunkter med vid jordbruksforskningen använda betraktelsesätt (jfr BACHÉR 1933). Det härledda uttrycket för avvikelsernas kvadratsumma (ekv. 9) är identiskt lika med »felens kvadratsumma» vid tillämpningen av FISHERS variationsanalys på jordbruksförsök. Beträffande variationsanalysens teori och tekniska utförande hänvisas till FISHERS bok *Statistical methods for research workers* (1932) eller elementära framställningar häröver av bl. a. BACHÉR (1933), SNEDECOR (1934) och TIRÉN (1934).

Vi återvända till ekv. (9). Härur erhålles rutornas medelavvikelse enligt formeln:

$$\sigma = \sqrt{\frac{\sum\limits_{u,v} (x_{uv} - m_u - m_v + M)^2}{(n - 1) - (l - 1) - (b - 1)}}, \quad \ldots\ldots\ldots\ldots \quad (10)$$

där n är totala antalet rutor och l betecknar antalet linjer samt b antalet block. Uttrycket i nämnaren anger antalet frihetsgrader enligt FISHERS terminologi (degrees of freedom). Antalet frihetsgrader är lika med antalet termer, som ingå i kvadratsumman minskat med antalet av varandra oberoende relationer mellan dessa termer. Vi gå här ej närmare in på den matematiska betydelsen av detta begrepp, utan hänvisa till FISHERS ovannämnda bok.

Medelfelet på medeltalet (ε_M) eller det sökta taxeringsfelet beräknas sedan enligt formeln:

$$\varepsilon_M = \frac{\sigma}{\sqrt{n}}$$

och alltså:

$$\varepsilon_M = \sqrt{\frac{\sum\limits_{u,v} (x_{uv} - m_u - m_v + M)^2}{n\,[(n - 1) - (l - 1) - (b - 1)]}}, \quad \ldots\ldots\ldots \quad (11)$$

Vi ha hittills för enkelhetens skull antagit, att linjelängden i rutorna är konstant. Varierar linjelängden (p_{uv}) måste avvikelserna vägas, samt vägda linje-, block- och stormedeltal införas i formel (11). I analogi med ett av LINDEBERG (1923) använt enkelt viktsystem erhålles följande formel för medelfelet:

$$\varepsilon_M = \sqrt{\frac{n}{(n-1)-(l-1)-(b-1)} \cdot \frac{\sum\limits_{u,v} (p_{uv})^2 (x_{uv} - m_u - m_v + M)^2}{\left(\sum\limits_{u,v} p_{uv}\right)^2}}, \ldots \quad (12)$$

där

$$M = \frac{\sum\limits_{u,v} p_{uv}\, x_{uv}}{\sum\limits_{u,v} p_{uv}}; \qquad m_u = \frac{\sum p_{u\cdot}\, x_{u\cdot}}{\sum p_{u\cdot}}; \qquad m_v = \frac{\sum p_{\cdot v}\, x_{\cdot v}}{\sum p_{\cdot v}}$$

I formel (12) betecknar x_{uv} taxeringsresultatet och p_{uv} linjelängden för den ruta, som ligger i linje u och block v, m_u och m_v äro motsvarande vägda linje- och blockmedeltal samt M det vägda medeltalet för samtliga rutor (stormedeltalet). n anger totala antalet rutor, l antalet linjer samt b antalet block.

Av den förda diskussionen (s. 316) och formel (11) eller (12) inses rent aritmetiskt, att möjligheten till elimination av systematiska avvikelser på grund av rutornas belägenhet i olika linjer och block blir större, om linjerna och blocken sinsemellan äro ungefär lika långa, än om stor variation föreligger, d. v. s. en mera regelbunden skogsfigur är gynnsam.

Vid beräkningen av medelfelet enligt denna metod för en skog med en mycket oregelbunden figur, bör skogen därför uppdelas i mera regelbundna delar, som felberäknas var för sig, varefter medelfelet för hela skogen (ε_M) erhålles enligt formeln:

$$\varepsilon_M = \sqrt{\left(\frac{P_1}{P}\right)^2 \cdot \varepsilon_1^2 + \left(\frac{P_2}{P}\right)^2 \cdot \varepsilon_2^2 + \left(\frac{P_3}{P}\right)^2 \cdot \varepsilon_3^2 \ldots}, \quad \ldots\ldots \quad (13)$$

där P betecknar totala linjelängden för hela skogen och P_1, P_2... etc. motsvarande längder för delarna samt ε_1, ε_2... etc. delarnas medelfel.

På grund av ovannämnda förhållande har i formel (12) någon hänsyn ej tagits till linje- och blockmedeltalens (m_u och m_v) olika vikt.

Ett kriterium på att det förefinnes en systematisk gång i x_{uv}-värdena erhålles, om medelfelet dessutom beräknas enligt följande formel, där beteckningarna betyda detsamma, som i formel (12).

$$\varepsilon_M = \sqrt{\frac{n}{n-1} \cdot \frac{\sum\limits_{u,v} (p_{uv})^2 (x_{uv} - M)^2}{\left(\sum\limits_{u,v} p_{uv}\right)^2}}, \quad \ldots\ldots\ldots\ldots \quad (14)$$

Lämnar formel (14) ett större medelfel än formel (12), föreligger det en systematisk gång i x_{uv}-värdena, som i motsvarande grad eliminerats vid medelfelsberäkningen. I annat fall har antingen någon systematisk tendens ej förefunnits, eller också har tillämpningen av metoden ej förmått eliminera densamma. Är skogens figur mycket oregelbunden och erfordras en större noggrannhet i medelfelsbestämningen, bör i detta fall en uppdelning av figuren prövas, såvida man ej på annat sätt, exv. genom grafiska uppläggningar (jfr fig. 1—6), kan styrka, att taxeringsmaterialet ej företer några markerade fördelningstendenser.

Medelfelet enligt ovan angivna formler avser populationen, och är således ett mått på medelavvikelsen mellan upprepade taxeringar, lika omfattande, som den aktuella taxeringen. Vill man däremot betrakta medelfelet som ett mått på avvikelsen mellan taxeringsresultatet och den undersökta storhetens verkliga värde, kan en uppskattning härav erhållas, om medelfelsformlerna multipliceras med faktorn $\sqrt{\frac{100-\alpha}{100}}$ (jfr CZUBER 1924), där α är taxeringsprocenten. Denna korrektionsfaktor har endast för högprocentiga taxeringar någon praktisk betydelse.

Vi skola något beröra frågan om rutornas storlek och form. Ju mindre rutorna göras, ju mindre blir i princip på grund av det ökade rutantalet medelfelets relativa medelfel. En uppfattning härom lämnar sammanställningen å s. 305, där n nu betyder antalet frihetsgrader. Lätt inses även, att den mindre rutstorleken ökar möjligheten att eliminera de systematiska tendenserna.

Vid den praktiska tillämpningen blir emellertid räknearbetet betungande med alltför stora rutantal, emedan det taxeringsresultat, som skall medelfelsberäknas, måste uträknas för varje ruta. Särskilt beträffande kubikmassan är detta arbetskrävande, varför rutantalet bör noga avvägas mot den önskade precisionen i medelfelsbestämningen. Härvid kan det ofta vara lämpligt att sammanföra taxeringslinjerna i grupper så, att exv. linjerna n:ris 1—3 bilda första gruppen och n:ris 4—6 den andra etc. Rutorna komma således att innehålla flera linjer. Antalet block eller linjegrupper bör dock ej understiga tre utan helst vara betydligt större.

Beträffande arealbestämningar tillkommer en annan synpunkt, nämligen att rutorna ej få tagas så små, att variationen blir alternativ, då metoden ej bör användas (jfr LANGSAETER 1926, s. 18.). Man får här av detta skäl nöja sig med den noggrannhet, som erhålles med relativt stora rutor. Vid den praktiska tillämpningen torde i regel samma rutor kunna användas för såväl arealbestämningar som kubikmassor och tillväxt etc., vilket ur arbetssynpunkt är en stor fördel.

Med hänsyn till rutornas form inses lätt, att dessa böra ha sin minsta utsträckning i den riktning, som den systematiska tendensen starkast gör sig gällande. För undvikandet av alternativ variation vid arealbestämningar, kan dock avsteg få göras från denna regel. Vet man ingenting om ev. fördelningstendenser i materialet, ligger det givetvis nära till hands, att göra rutorna kvadratiska.

Av det föregående torde ha framgått, att den här utvecklade metoden för medelfelets härledning innesluter stora möjligheter att beräkna ett med hänsyn till såväl systematiska som tillfälliga fel tillförlitligt medelfel. För härledningen av arealbestämningars medelfel bör metoden i regel endast användas vid taxering av större skogar.

I anslutning till den i det föregående behandlade metoden för medelfelets bestämning torde böra omnämnas en annan beräkningsmetod, som kan betraktas som en variant av metod C (jfr. s. 304). Metoden bygger på samma block- och rutindelning som i det föregående. Medelfelet beräknas med ledning av differenserna mellan taxeringsresultaten från på varandra följande rutor i samma block, varvid såväl första (formel 15) som andra ordningens differenser (formel 16) kunna användas. Medelfelet för v:te blocket (ε_v) erhålles enligt följande formler (LINDEBERG 1923, LANGSÆTER 1927):

$$\varepsilon_v = \sqrt{\frac{r \cdot \sum_{u=1}^{r-1} (p_{uv} + p_{(u+1)\,v})^2 \,(x_{uv} - x_{(u+1)\,v})^2}{2\,(r-1)\left(2 \sum_{u=1}^{r} p_{uv}\right)^2}}, \quad \ldots\ldots \quad (15)$$

$$\varepsilon_v = \sqrt{\frac{r \cdot \sum_{u=2}^{r-1} (p_{uv})^2 \,(x_{(u-1)\,v} + x_{(u+1)\,v} - 2\,x_{uv})^2}{6\,(r-2)\left(\sum_{u=1}^{r} p_{uv}\right)^2}} \quad \ldots. \quad (16)$$

I dessa formler betecknar x_{uv} taxeringsresultatet och p_{uv} linjelängden för den ruta, som ligger i linje u och block v, samt r antalet rutor (linjer) i v:te blocket. Härefter beräknas medelfelet för hela skogen (ε_M) i analogi med formel (13) s. 318.

Av dessa formler är den första i princip fördelaktigare med avseende på det tillfälliga felet å medelfelet, emedan antalet differenser blir större, och den senare med hänsyn till det systematiska felet. Föreligger en systematisk gång i taxeringsresultaten ge båda formlerna för höga medelfel, och formel (15) högre värde än formel (16). Metoden förmår ej borteliminera den genomsnittliga tendensen från ruta till ruta (jfr. s. 306).

Denna metod synes tillämplig, när större noggrannhet eftersträvas, och den förra huvudmetoden (formel 12) ej kan komma till användning samt vid mindre fordrande beräkningar, då linjeantalet är för ringa för en direkt tillämpning av LANGSAETERS formel (formel 2, s. 304). Beträffande rutornas storlek och form gäller, vad som sagts under huvudmetoden.

Provytetaxering.

Den i det föregående angivna metodiken för medelfelets härledning vid linjetaxering gäller även för den regelbundna provytetaxeringen. Härvid möta inga särskilda svårigheter. Provytetaxeringens resultat uträknas för varje ruta, varefter medelfelet erhålles ur formel (12). För arealbestämningar synes emellertid ej sällan en annan beräkningsmetod vara att föredraga. Vi skola närmare diskutera denna metod.

Vid den regelbundna provytetaxeringen göres i regel den enskilda provytan relativt liten, för att den endast i undantagsfall skall behöva fördelas på olika markslag eller andra arealbestämningar. Det på sina håll tillämpade förfaringssättet att i sådana fall flytta ytan, så att den endast omfattar ett markslag, är ur statistisk synpunkt oriktigt och bör ej komma till användning. Under ovan nämnda förutsättningar blir variationen i princip alternativ, d. v. s. antingen finns egenskapen å ytan, eller också finns den ej. Det ligger då nära till hands att söka härleda arealbestämningarnas medelfel med stöd av den alternativa eller homograda statistikens klassiska satser.

Sådana försök ha gjorts av SCHUMACHER och BULL (1932) samt av MUDGETT och GEVORKIANTZ (1934). Härvid har principiellt hela taxeringen hänförts till ett visst urnschema, varvid de förra tillämpat BERNOULLIS och de senare POISSONS eller LEXIS schema (CHARLIER 1931). LANGSAETER har visat, att fördelningen av olika markslag har samma tendenser, som vi här funnit hos kubikmassan (jfr. s. 315). Det är för övrigt helt naturligt, att så skall vara fallet. En tillämpning av det enklaste urnschemat (BERNOULLIS) på hela taxeringsobjektet måste därför betraktas som alltför schematisk, vilket även framgått av de ovannämnda undersökningarna. MUDGETTS och GEVORKIANTZS användning av POISSONS eller LEXIS teorem är därför en utveckling i rätt riktning mot verklighetens komplexa urnschemata.

En närmare anpassning av beräkningsmetodiken till de faktiska förhållandena kan ske på två sätt. Antingen genom kombinationer av POISSONS och LEXIS schemata eller också genom uppdelning av taxeringen i mindre enheter, för vilka BERNOULLIS schema approximativt är tillämpligt. Jag har i anslutning till metodiken för härledningen av kubikmassans medelfel valt den senare vägen.

Vi återvända till den i det föregående (s. 316) använda indelningen av skogen i rutor. Dessa rutor betrakta vi som urnor med approximativt regellöst (slumpvis) fördelade svarta och vita kulor. Kulorna utgöras av de provytor, vari den enskilda rutan kan tänkas uppdelad. Härvid betecknar exv. de svarta kulorna det markslag, vars medelfel skall härledas. Taxeringen av rutan sker visserligen strängt systematiskt, emedan provytorna utläggas i ett visst förband. Men om arealens fördelning inom rutan ej visar några systematiska tendenser, kunna vi approximativt betrakta provytetaxeringen som dragningar ur en urna, varvid den för varje gång dragna kulan ej lägges tillbaka i urnan. Antalet dragningar motsvaras härvid av antalet provytor, som vid taxeringen utfalla i rutan. Under sådana förutsättningar kan provytetaxeringens medelfel för den enskilda rutan lätt beräknas.

Låt N beteckna totala antalet provytor, vari i:te rutan kan tänkas uppdelad (rutans areal : provytans areal) och p_i antalet vid taxeringen utfallande provytor i denna ruta, varav x_i procent ligga på det markslag, vars medelfel skall bestämmas. Detta medelfel (ε_i) erhålles då enligt följande formel (CRAMÉR 1926):

$$\varepsilon_i = \sqrt{\frac{(N - p_i)}{(N - 1)} \cdot \frac{x_i\,(100 - x_i)}{p_i}} \quad \ldots\ldots\ldots\ldots \quad (17)$$

Men $N = \frac{100\,p_i}{a}$, där a betyder taxeringsprocenten. Formel (17) kan därför även skrivas:

$$\varepsilon_i = \sqrt{\frac{(100 - a)\,x_i\,(100 - x_i)}{100\,p_i - a}} \quad \ldots\ldots\ldots\ldots \quad (18)$$

På detta sätt (formel 18) kan medelfelet beräknas för varje ruta, varefter medelfelet på taxeringsresultatet för hela skogen (ε_M) erhålles enligt formeln:

$$\varepsilon_M = \sqrt{\left(\frac{p_1}{P}\right)^2 \cdot \varepsilon_1{}^2 + \left(\frac{p_2}{P}\right)^2 \cdot \varepsilon_2{}^2 + \left(\frac{p_3}{P}\right)^2 \cdot \varepsilon_3{}^2, \ldots \left(\frac{p_n}{P}\right)^2 \cdot \varepsilon_n{}^2} \quad (19)$$

där P anger totala antalet ytor (Σp_i) och n betecknar antalet rutor.

Faktorn $\frac{N - p_i}{N - 1}$ uppgår ofta till så obetydliga värden, att den kan försummas. Härvid får formel (19) följande enkla form:

$$\varepsilon_M = \sqrt{\frac{\sum\limits_{i=1}^{n} p_i x_i\,(100 - x_i)}{(P)^2}} \quad \ldots\ldots\ldots\ldots \quad (20)$$

Den ovan angivna metoden för härledningen av medelfelet på arealbestämningar är föga arbetskrävande och synes ej sällan kunna ersätta huvudmetoden (formel 12). Förutsättningen härför är, att det ej förekommer någon mera markerad, systematisk tendens inom rutorna. Denna risk minskas givetvis ju mindre rutorna göras, men å andra sidan bör inom rutan provyteantalet för det arealslag, vars medelfel skall beräknas, ej vara alltför ringa (jfr YULE 1937).

KAP. III. METODIKENS TILLÄMPNING VID LINJETAXERING.

Jag har, som tidigare nämnts (s. 301), använt den här utvecklade metoden för medelfelets beräkning vid viss utredning för riksskogstaxeringsnämndens räkning. Vi skola i det följande diskutera några ur metodsynpunkt intressantare resultat från denna undersökning.

Materialet utgöres av den förra riksskogstaxeringen av Västernorrlands och Norrbottens län samt Ljungans flodområde. Taxeringsmetoden var som bekant linjetaxering med en mils linjeavstånd i Västernorrlands län och Ljungans flodområde samt två mils avstånd i Norrbottens län. Bältesbredden utgjorde 10 m. Den utförda bearbetningen begränsades till att gälla medelfelet å kubikmassan inom bark med hänsyn till grundytans osäkerhet, varvid kubikmassan i rutorna beräknades med hjälp av för taxeringsområdet genomsnittliga kuberingstal i de olika diameterklasserna. I tab. 2 lämnas några uppgifter för att ur taxatorisk synpunkt karakterisera materialet.

Tab. 2. **Areal- och kukikmasseuppgifter för taxeringsområdena.**
Area and cubic volume of the survey districts.

Taxeringsområde Survey district	Landareal Land area km²	Härav fjäll Thereof land above the coniferous timber line km²	Linjeavstånd Distances between the survey lines km	Taxerad areal Surveyed area km²	Antal rutor Number of squares	Kubikmassa i tusental m³ / Volume in thousands of cubic metres 15— Tall Pine	15— Gran Spruce	25— Tall Pine	25— Gran Spruce	35— Tall Pine	35— Gran Spruce
Norrbottens lappmark....	72 232	34 646	20	19,15	225	41 679	24 488	22 941	10 596	6 367	2 895
Norrbottens kustland....	26 428	—	20	12,34	162	34 130	16 260	13 049	4 421	1 952	607
Västernorrlands län	24 128	—	10	23,94	291	24 801	37 527	11 324	9 716	2 446	1 280
Ljungans flodområde: Västernorrlandsdelen...	5 388	—	10	5,01	74	6 700	9 697	3 150	2 823	780	465
Ljungans flodområde: Jämtlandsdelen	7 433	1 295	10	6,02	89	9 302	9 869	4 501	3 037	961	384
Hela Ljungans flodområde	12 821	1 295	10	11,03	163	16 002	19 566	7 651	5 860	1 741	849

Vid medelfelets beräkning enligt formel (12) bör lämpligen användas ett räkneschema av den form, som framgår av tab. 3. Denna tabell utgör endast ett brottstycke, omfattande tre taxeringslinjer, ur en av de utförda medelfelsberäkningarna.

Medelfelet har beräknats för tall och gran skilda åt samt dimensionsgrupperna: 15 cm vid brösthöjd och däröver, 25 cm och däröver, 35 cm och däröver. För Västernorrlands län och Ljungans flodområde gäller medelfelet kubikmassan per hektar landareal samt för Norrbotten kubikmassan per hektar landareal exklusive inägor. Resultaten av beräkningarna äro sammanförda i tab. 4, s. 326.

F o r m e l (14) h a r g e n o m g å e n d e l ä m n a t h ö g r e m e d e l f e l ä n f o r m e l (12). I g e n o m s n i t t u p p g å r m e d e l f e l e t e n l i g t f o r m e l (14) f ö r d e o l i k a d i m e n s i o n s g r u p p e r n a t i l l r e s p. 116, 112 o c h 109 p r o c e n t a v m e d e l f e l e t e n l i g t f o r m e l (12). Detta visar, att systematiska tendenser förelegat i materialet och i motsvarande grad eliminerats vid medelfelets beräk-

Tab. 3. **Räkneschema**
Calculating-scheme

$x = (m^3/ha) \cdot 100$; p = linjelängd
length of line

Block n:r Block no	Linje n:r 1 Line no. $M - m_u = -44$							Linje n:r 2 Line no. $M - m_u = -194$						
	p	x	px *	$x - m_v$	$x - m_v + M - m_u = w$	pw **	$(pw)^2$ ***	p	x	px *	$x - m_v$	$x - m_v + M - m_u = w$	pw **	$(pw)^2$ ***
1	2 800	2 010	563	— 114	— 158	— 44	194	1 400	2 351	329	+ 227	+ 33	+ 5	3
2	9 990	2 122	2 120	— 515	— 559	— 558	31 136	9 745	3 639	3 546	+ 1 002	+ 808	+787	61 937
3	10 000	3 148	3 148	+1 688	+ 1 644	+1 644	270 274	8 840	759	671	— 701	— 895	—791	62 568
4	9 800	996	976	— 2	— 46	— 45	203	9 795	1 155	1 131	+ 157	— 37	— 36	130
5	8 915	632	563	— 28	— 72	— 64	410	9 915	477	473	— 183	— 377	—374	13 988
6	8 930	336	300	— 248	— 292	— 261	6 812	9 405	644	606	+ 60	— 134	—126	1 588
7	9 330	696	649	— 25	— 69	— 64	410	6 400	895	573	+ 174	— 20	— 13	17
8	9 990	518	517	— 932	— 976	— 975	95 063	9 960	2 276	2 267	+ 826	+ 632	+629	39 564
9	9 830	679	667	+ 13	— 31	— 30	90	9 240	462	427	— 204	— 398	—368	13 542
10	—	—	—	—	—	—	—	7 315	1 363	997	+ 118	— 76	— 56	314
Σ	79 585	—	9 503					82 015	—	11 020				
Medeltal Average		1 194 m_u							1 344 m_u					

Sifferreduktioner: * $px \cdot 0{,}0001$
Digitreductions ** $pw \cdot 0{,}0001$
*** $(pw \cdot 0{,}0001)^2 \cdot 0{,}1$

ning enligt formel (12) (jfr. s. 318). På grund av det stora rutantalet (tab. 2) äro medelfelen enligt formlerna (12) och (14) härledda med en betydande säkerhet.

För Norrbottens och Västernorrlands län har en jämförelse gjorts mellan Langsaeters formel (formel 2 s. 304) och formel (12). De enskilda medelfelen enligt Langsaeters formel äro med hänsyn till det ringa linjeantalet behäftade med avsevärda medelfel. Vi kunna därför endast diskutera den genomsnittliga tendensen för hela jämförelsematerialet. Tab. 5 visar, att medelfelet enligt Langsaeters formel i genomsnitt är 114 procent av det enligt formel (12) beräknade medelfelet.

Av fig. 1—7 och tab. 4 framgår, att systematiska tendenser förefinnas i materialet. Under sådana förhållanden har tidigare visats (jfr. s. 306), att Langsaeters formel ger för högt resultat. Enligt andra erfarenheter (*Riksskogstaxeringsnämnden* 1932, Östlind 1932, *Landsskogtakseringen* 1938), skulle denna överskattning i genomsnitt uppgå till 10 à 20 procent, vilket ganska väl motsvarar den här funna skillnaden mellan Langsaeters formel och formel (12). Den här framlagda metoden synes således möjliggöra en

för formel (12).
to formula (12).

i m; $n = 28$; $l = 3$; $b = 10$.
in m.

Linje n:r 3 Line no.						$M - m_u = + 305$	Summa och medeltal Sum and average			
p	x	px *	$x - m_v$	$x - m_v + M - m_u = w$	pw **	$(pw)^2$ ***	$\sum px$	$\sum p$	m_v	$\sum (pw)^2$
—	—	—	—	—	—	—	892	4 200	2 124	197
2 045	383	78	— 2 254	— 1 949	— 399	15 920	5 744	21 780	2 637	108 993
9 750	364	355	— 1 096	— 791	— 771	59 444	4 174	28 590	1 460	392 286
9 945	845	840	— 153	+ 152	+ 151	2 280	2 947	29 540	998	2 613
9 700	874	848	+ 214	+ 519	+ 503	25 301	1 884	28 530	660	39 699
9 540	757	722	+ 173	+ 478	+ 456	20 794	1 628	27 875	584	29 194
9 530	629	599	— 92	+ 213	+ 203	4 121	1 821	25 260	721	4 548
9 960	1 559	1 553	+ 109	+ 414	+ 412	16 974	4 337	29 910	1 450	151 601
1 900	1 590	302	+ 924	+ 1 229	+ 234	5 476	1 396	20 970	666	19 108
1 450	645	94	— 600	— 295	— 43	185	1 091	8 765	1 245	499
63 820	—	5 391					25 914	225 420	—	748 738
	845								1 150	
	m_u								M	

$$\varepsilon_M = \sqrt{\frac{28}{(27-2-9)} \cdot \frac{748\,738 \cdot (10)^9}{(225\,420)^2}}$$

$\varepsilon_M = 160{,}6$ eller i % av $M = 14{,}0$

noggrann uppskattning av medelfelet, även när systematiska fördelningstendenser föreligga i taxeringsmaterialet.

Relativa medelfelet, då taxeringslinjerna omfatta 100 ha land ($\varepsilon_{100\ \text{ha}}$), kan betraktas som ett jämförelsemått på variationen och därmed också på taxeringssvårigheten. I fig. 8 har detta medelfel upplagts grafiskt över kubikmassan per ha. Härav framgår, att medelfelet sjunker med stigande kubikmassa, vilket även andra undersökningar visa (jfr LANGSAETER 1932, ÖSTLIND 1932). Vidare synes, att granen i Norrbotten utgör en från de övriga grupperna skild variationstyp (population) med större taxeringssvårighet. Av HESSELMANS karta (HESSELMAN 1935) framgår också, att granen har en mycket oregelbunden förekomst i Norrbotten. De noggranna medelfelsbestämningarna i tab. 4 (formel 12) utgöra ett empiriskt material av betydande värde. En mera taxatoriskt betonad diskussion av detsamma har dock ansetts falla utom ramen för denna uppsats.

Tab. 4. **Medelfelet å kubikmassan per hektar**
The standard error of cubic volume per hectare

Taxeringsområde Survey district	Trädslag Tree species	Dimensions- Diameter-				
		15—				
		Medelfel i procent Standard error in percentage			Kol. 4 i % av kol. 3 Col. 4 in % of col. 3	Kubikmassa per ha Cubic volume per hectare
		å taxeringsresultatet enligt The actual survey accord. to		vid taxering av 100 hektar enligt formel 12 Surveying 100 hectares accord. to formula 12		
		formel 12 formula 12	formel 14 formula 14			
1	2	3	4	5	6	7
Norrbottens lappmark	Tall Pine	4,15	5,34	18,2	129	10,58
	Gran Spruce	5,54	6,07	24,3	110	6,15
Norrbottens kustland	Tall Pine	3,28	4,21	11,5	128	13,77
	Gran Spruce	7,28	7,69	25,6	106	6,55
Västernorrlands län	Tall Pine	3,17	3,63	15,5	115	10,32
	Gran Spruce	2,65	3,08	13,0	116	15,64
Ljungans flodområde: Västernorrlansdelen	Tall Pine	5,31	6,53	11,9	123	13,20
	Gran Spruce	5,41	5,97	12,1	110	18,07
Ljungans flodområde: Jämtlandsdelen	Tall Pine	5,73	5,77	14,1	101	13,35
	Gran Spruce	5,99	7,46	14,7	125	14,31
Hela Ljungans flodområde	Tall Pine	3,96	—	13,2	—	13,28
	Gran Spruce	4,02	—	13,4	—	16,02
Medeltal Average	—	—	—	—	116	—

med hänsyn till grundytans osäkerhet.
accord. to uncertainly of the basal area.

grupp class									
25—					35—				
Medelfel i procent Standard error in percentage			Kol. 9 i % av kol. 8 Col. 9 in % of col. 8	Kubikmassa per ha Cubic volume per hectare	Medelfel i procent Standard error in percentage			Kol. 14 i % av kol. 13 Col. 14 in % of col. 13	Kubikmassa per ha Cubic volume per hectare
å taxeringsresultatet enligt The actual survey accord. to		vid taxering av 100 hektar enligt formel 12 Surveying 100 hectares accord. to formula 12			å taxeringsresultatet enligt The actual survey accord. to		vid taxering av 100 hektar enligt formel 12 Surveying 100 hectares accord. to formula 12		
formel 12 formula 12	formel 14 formula 14				formel 12 formula 12	formel 14 formula 14			
8	9	10	11	12	13	14	15	16	17
4,41	5,46	19,3	124	5,81	6,77	7,04	29,6	104	1,61
7,06	7,46	30,9	106	2,67	8,95	9,82	39,2	110	0,73
5,55	6,17	19,5	111	5,27	8,92	10,78	31,4	121	0,79
11,55	12,10	40,6	105	1,79	17,63	18,00	61,9	102	0,25
4,29	4,74	21,0	110	4,71	6,80	7,43	33,3	109	1,02
4,15	4,62	20,3	111	4,04	7,72	8,29	37,8	107	0,53
7,18	8,34	16,1	116	6,24	9,75	12,05	21,8	124	1,54
9,22	9,58	20,6	104	5,06	14,30	15,38	32,0	108	0,82
7,16	7,22	17,6	101	6,47	12,14	12,12	29,8	100	1,37
8,27	11,01	20,3	133	4,24	14,60	16,65	35,8	114	0,55
5,10	—	16,9	—	6,37	7,83	—	26,0	—	1,45
6,19	—	20,6	—	4,61	10,30	—	34,2	—	0,67
—	—	—	112	—	—	—	—	110	—

Tab. 5. **Jämförelse mellan enligt formel (2) och enligt formel (12) beräknade medelfel.**
Comparison between standard error computed accord. to formula (2) and formula (12).

Taxeringsområde Survey district	Antal Number of: tax-erings-linjer survey lines	rutor squares	Träd-slag Tree species	Dimen-sions-grupp Diam.-class	Medelfel i procent enligt Standard error in percentage accord. to: formel 2 formula 2	formel 12 formula 12	Kol. 6 i % av kol. 7 Col. 6 in % of col. 7
1	2	3	4	5	6	7	8
Norrbottens lappmark	11	225	Tall Pine	15—	3,77	4,15	91
				25—	5,84	4,41	132
				35—	8,68	6,77	128
			Gran Spruce	15—	5,11	5,54	92
				25—	5,81	7,06	82
				35—	8,44	8,95	94
Norrbottens kustland	12	162	Tall Pine	15—	5,43	3,28	166
				25—	7,85	5,55	141
				35—	13,34	8,92	150
			Gran Spruce	15—	8,96	7,28	123
				25—	12,57	11,55	109
				35—	23,03	17,63	131
Västernorrlands län ..	27	291	Tall Line	15—	3,19	3,17	101
				25—	3,34	4,29	78
				35—	6,77	6,80	100
			Gran Spruce	15—	3,19	2,65	120
				25—	4,37	4,15	105
				35—	8,77	7,72	114
				Medeltal Average	—	—	114

KAP. IV. OM VALET AV METOD FÖR MEDELFELETS HÄRLEDNING.

Vid valet av metod för medelfelets härledning gäller den allmänna regeln att anpassa metoden efter det föreliggande behovet. I vissa fall är det nödvändigt att arbeta med den största noggrannhet, som kan åstadkommas, i andra fall äro enklare metoder fullt tillfredsställande, varför de för kostnadernas nedbringande böra äga företräde.

Det är givetvis omöjligt att ange några generella regler, när den ena eller andra metoden bör användas. Detta beror dels på det behov, medelfelsberäkningen skall tillfredsställa, dels på taxeringsmaterialets inre struktur. Vi kunna därför endast anlägga några allmänna synpunkter på frågan. Härvid behandla vi linje- och provytetaxeringen var för sig.

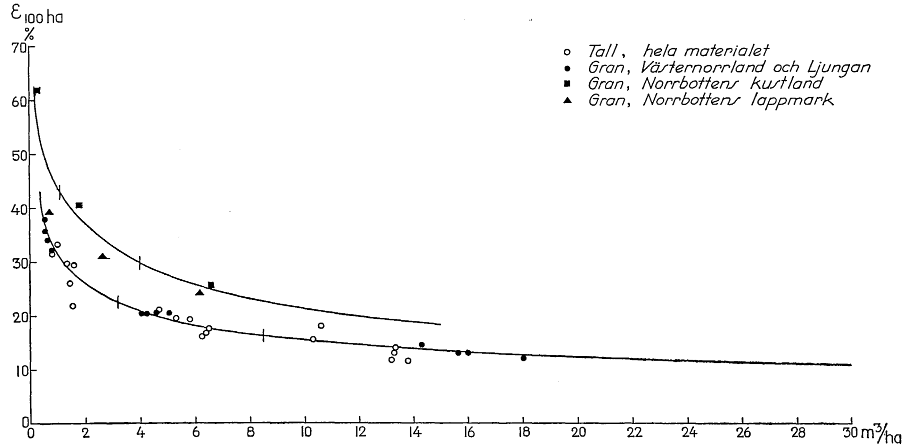

Fig. 8. Sambandet mellan relativa medelfelet ($\varepsilon_{100\ \mathrm{ha}}$), då taxeringslinjerna omfatta 100 hektar, och kubikmassan per ha. Tvärstrecken på kurvorna avskilja de olika dimensionsgrupperna.

Relation between the relative standard error, when the survey lines contain 100 hectares ($\varepsilon_{100\ \mathrm{ha}}$) and cubic volume per hectare. The cross-lines on the curves separate the different diameter-classes.

Linjetaxering. Erfordras större precision i medelfelsbestämningen, bör den i det föregående utvecklade huvudmetoden (formel 12, s. 317) komma till användning, om systematiska tendenser med avseende på taxeringsobjektets fördelning förefinnas. Detta torde ofta vara fallet. Saknas sådana tendenser, eller äro de mindre utpräglade, kan LANGSAETERS formel (formel 2, s. 304) tillämpas, såvida ej linjeantalet är ringa, då den här angivna varianten av den linjevisa differensmetoden (formlerna 15 och 16, s. 320) bör tillgripas. Erfordras ej någon större precision i medelfelsbestämningen, kunna som allmänna beräkningsmetoder användas de sistnämnda, mindre arbetskrävande formlerna (2, 15 och 16). Härvid erhålles systematiskt ett för högt medelfel, om tendenser föreligga i taxeringsmaterialet.

Provytetaxering. Med undantag för arealbestämningar gälla samma regler som för linjetaxeringen. Vid beräkningen av medelfelet å arealuppgifter torde, om den enskilda provytan är så liten, att variationen blir alternativ, den i det föregående härledda specialmetoden ej sällan vara att föredraga (jfr. s. 322).

Under diskussionen av metodiken vid medelfelets härledning har det framgått, att vi med de metoder, vilka f. n. stå oss till buds, ej alltid kunna beräkna ett medelfel, som har något verkligt värde för den aktuella frågeställningen. Taxeringsmaterialet kan vara av för ringa omfattning eller i andra avseenden så extremt, att de grundläggande förutsättningar, varpå metoderna vila, ej föreligga i tillräcklig grad. I sådana fall böra vi givetvis avstå ifrån medelfelsberäkningen och söka bedöma taxeringens tillförlitlighet med stöd av andra indicier.

KAP. V. SAMMANFATTNING.

Vid sannolikhetskalkylens tillämpning på praktiska frågeställningar är det ofta svårt att bevisa tillämpningens berättigande. Sannolikhetskalkylen i här avsedd bemärkelse vilar på rent matematisk grund och bygger således på vissa bestämda förutsättningar, som endast med större eller mindre approximation kunna återfinnas i ett konkret material. Graden av approximation blir avgörande för kalkylens värde. För sannolikhetskalkylens praktiska användning är det således av väsentlig betydelse, att tillämpningsmetoder utvecklas, som ej bygga på några speciella förutsättningar utan endast på mera allmängiltiga sådana. Varje framträngande mot detta mål måste betraktas som ett framsteg.

Avsikten med denna uppsats har varit att lämna ett bidrag till en sådan utveckling av metodiken för medelfelets härledning vid linje- och provytetaxering. Den här framlagda huvudmetoden (formel 12, s. 317) innebär ett nytt grepp på denna fråga och gäller i princip även om systematiska tendenser föreligga med avseende på taxeringsobjektets fördelning, vilket är ett ofta förekommande fall.

Metoden har använts av mig för beräkningen av medelfelet å virkesförrådet vid den förra riksskogstaxeringen av Västernorrlands och Norrbottens län samt Ljungans flodområde (tab. 4, s. 326). Denna bearbetning utgör ett empiriskt material av betydande värde för belysandet av linjetaxeringens noggrannhet. Undersökningen har dock här endast utnyttjats i syfte att jämföra olika metoder för medelfelets härledning, under det att en mera taxatoriskt betonad diskussion av detsamma ansetts falla utom ramen för denna uppsats.

För härledning av medelfelet å arealbestämningar vid provytetaxering har dessutom utvecklats en specialmetod (s. 322). Metoden är föga arbetskrävande och synes ej sällan vara användbar.

Anförd litteratur.

BACHÉR, I., 1933, Moderna synpunkter på fältförsökets metodik och den statistiska analysen av försöksresultatet. Nord. Jordbr. Forskning.

CHARLIER, C. V. L., 1931, Vorlesungen über die Grundzüge der mathematischen Statistik. Lund.

CRAMÉR, HARALD, 1926, Sannolikhetskalkylen och några av dess användningar. Stockholm.

CZUBER, E., 1924, Wahrscheinlichkeitsrechnung. Erster Band. Leipzig und Berlin.

FISHER, R. A., 1932. Statistical Methods for Research Workers. Edinburg.

HELMERT, F. R., 1924, Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. Leipzig und Berlin.

HESSELMAN, H., 1935, Barrskogens arealfördelning på tall-, gran- och barrblandbestånd i Norrland och Dalarna. Medd. från Statens skogsförsöksanstalt. H. 28, nr 8.

Kommissionen för försökstaxering etc., 1914, Värmlands läns skogar. Betänkande. Stockholm.

Landsskogtakseringen, 1938, Taksering av Norges skoger. Østfold fylke. Oslo.

LANGSAETER, A., 1926 och 1927, Om beregning av middelfeilen ved regelmessige linjetakseringer. Medd. fra det norske skogforsøksvesen, H. 7 och 8.

—— 1932, Nøiaktigheten ved linjetaksering av skog I. Medd. fra det norske skogforsøksvesen, H 15.

LINDEBERG, J. W., 1923, Über die Berechnung des Mittelfehlers des Resultates einer Linientaxierung. Acta forestalia fennica, H. 25.

—— 1926, Zur Theorie der Linientaxierung. Acta forestalia fennica, H. 31.

MUDGETT, BRUCE D. and GEVORKIANTZ, S. R., 1934, Reliability of forest surveys. Journal of the American Statistical Association.

Riksskogstaxeringsnämnden, 1932, Uppskattningen av Sveriges skogstillgångar verkställd åren 1923—29. Stockholm.

SCHUMACHER, FRANCIS X. and BULL, HENRY, 1932, Determination of the errors of estimate of a forest survey, with special reference to the bottom-land hardwood forest region. Journal of Agricultural Research.

SNEDECOR, GEORGE W., 1934, Calculation and interpretation of analysis of variance and covariance. Ames. Iowa.

TIRÉN, LARS, 1934, Nyare fältförsöksmetodik, belyst genom några skogsodlingar på Kulbäckslidens försökspark. Medd. från Statens skogsförsöksanstalt. H. 27, nr 6.

YULE, G. UDNY and KENDALL, M. G., 1937, An introduction to the theory of statistics London.

ÖSTLIND, J. och HAGSTRÖM, H. G., 1922, Felrisken vid provtaxering (manuskript).

ÖSTLIND, J., 1932, Erforderlig taxeringsprocent vid linjetaxering av skog. Skogsvårdsför. tidskrift.

RESUMÉ.

On computing the standard error in line and sample plot surveying.

Introduction.

Since the Värmlandian survey has proved the application and suitability of the probability-computation for determining the degree of accuracy in an objective survey (*the Commission for an experimental survey* etc., 1914), several methods have been suggested for computing the standard error in a survey. Reliable methods of calculation of this kind are an important condition for planning our forest surveys on rational lines. The methods proposed have previously been subjected to critical scrutiny (LINDEBERG 1923, 1926, LANGSAETER 1926, 1927, 1932, NÄSLUND 1930), which revealed, inter alia, the fact that it was highly necessary to develop the methodics of the subject still further. The object here is to make a contribution to this question, which has become an increasingly urgent one in conjunction with the renewal of the national forest surveys in the Northern Countries.

The method proposed in the following article of computing the standard error has been employed by me in certain inquiries made on behalf of *The National Forest Survey Commission* of 1937.

The most recent development of the methods of computing the standard error.

We shall here merely confine ourselves to the development which the question under discussion has undergone since I last dealt with the subject (NÄSLUND 1930).

LANGSAETER has carried out a far-reaching investigation, the aim of which was by compiling empirical survey material to contribute towards solving the question of what degree of accuracy is attained in the linesurveying of forests of various size and composition (evenness) with the use of different survey percentages (LANGSAETER 1932). In this paper LANGSAETER again discusses at length the methodics of computing the standard error.

In the theoretical section of his paper LANGSAETER gives, among other things, a new formula for calculating the standard error (LANGSAETER 1932, p. 461, formula 21). This formula contains differences of the second order in the observed quantity for distances between the lines right down to 10 m. The formula cannot be used without a process of extrapolation, and there is no definitly known function in support of this extrapolation. The method is therefore, as LANGSAETER himself points out, only of a certain theoretical interest.

In making a general trial of the most common formulae for calculating the mean error, LANGSAETER discusses, among others, a method suggested by me (cf. NÄSLUND 1930, p. 330, method D: 2). When employing it LANGSAETER uses, in part, a method of procedure different from the one given by me. His procedure is not in accord with the original train of thought on which my method is based. I shall therefore dwell on this point briefly here.

In principle, the purpose of smoothing out the mean difference is merely to eliminate the average effect of any systematic tendency there might be from line to line in the actual survey. No other meaning has been given to the extrapolation of the difference-curve to zero. As the basis of the smoothing out process, therefore, we shall only take those mean differences which may arise from lines included in the survey.

When for very low-percentage surveys LANGSAETER calculates the mean difference of the actual survey with the guidance of the mean differences for 10 and 20 m line distances, this implies in principle something quite different and indicates an erroneous interpretation of the theoretical grounds on which the method is based (LANGSAETER 1932, p. 483). For certain comparisons I have calculated (NÄSLUND 1930, tab. 11, p. 335) the 5 and 2 ½ % surveys' standard error on the basis of the 10 % survey's mean difference. This is however not an application of the method but a process of approximation resorted to for lack of other means, and has the character of a relatively moderate extrapolation.

LANGSAETER further shows that under certain conditions my method gives the same results as his method C: 2 (cf. LANGSAETER 1932, p. 471), which has the advantage of requiring less labour. We need only add the observation that this presupposes, inter alia, a smoothing-out of the mean-difference curve with the aid of only 2 points, which of course was not the intention (cf. NÄSLUND 1930, p. 332, fig. 6).

The method discussed above has been proposed by me in connection with an examination of the methodics used in computing the standard error and was then explanatory and of mainly theoretical interest. However, it has a limited practical utility, as I have pointed out previously (NÄSLUND 1930, p. 330), so that it is not necessary to dwell on that point here.

In Sweden, ÖSTLIND has carried out a comprehensive investigation on behalf of the Royal Board of Domains with the same object in view as LANGSAETER's above-mentioned enquiry (ÖSTLIND 1932). ÖSTLIND discusses here the theory of computing the standard error in line surveying and deduces a formula for the standard error, which formula may be regarding as a modification or variant of LANGSAETER's formula based on differences of the second order (LANGSAETER 1926, 1927 and 1930).

In the U. S. A. an opportunity has been afforded by the regular sample plot survey to make a close study of the object aimed at here (SCHUMACHER and BULL 1932, MUDGETT and GEVORKIANTZ 1934). For this purpose the question of computing the standard error has been discussed in conjunction with the theorem of BERNOULLIS, POISSON and LEXI, to which we shall revert in p. 341. In the following pages we shall call the regular sample plot survey merely the sample plot survey, meaning thereby a survey consisting of a number of fairly small sample plots in a certain regular formation.

The above observations cover the most important known works which have been published since 1930 and deal with the question of the methodics used in computing the standard error in regular line and sample plot surveying. It will have been gathered that for calculating the standard error in an actual line survey we still have to take recourse to those methods which were employed when I dealt with this subject in my former paper, to which the reader is referred in order to avoid unnecessary repetition (NÄSLUND 1930). On the other hand, our experiences in regard to the accuracy of the line survey have become greatly widened as a result of the considerable amount of work expended on compiling empirical survey material (LANGSAETER 1932, ÖSTLIND 1932, *The National Forest Survey Commission* 1932).

General views on the question of method.

Of the methods proposed there are principally two that in Sweden have been in practical use for calculating the standard error in an actual survey. In my previous paper on the subject these methods were denoted by method A. and method C. respectively (NÄSLUND 1930), and these letters will serve the same purpose in the following discussion.

Method A. The survey lines or parts thereof are divided into groups (partial surveys), after which the standard error (ε_M) is calculated according to the formula (1), p. 304, in which x_i denotes the survey results of the i-th group, p_i its total length of line and n the number of groups.

Method C. The standard error is calculated with the aid of differences between the survey results derived from succeeding survey lines. Standard error formulae according to this principle are given by ÖSTLIND—HAGSTRÖM (1922), LINDEBERG (1923, 1926) and LANGSAETER (1924, 1932). For this purpose LANGSAETER's formula based on differences of the second order has often been preferred (LANGSAETER 1932, method C: 2). The standard error is then obtained by means of the following (2), p. 304, in which x_i is the result survey and p_i the surveyed length for the i-th survey line, where i indicates the sequence of the lines. n denotes the number of survey lines.

Sometimes a combination of these methods has been used, the lines being first divided into groups and formula (2) being then applied to the group results. Taking these methods as our background we shall discuss in detail the problem as it now presents itself.

The calculated standard errors are in their turn liable to standard errors owing to the limited number of variants (partial surveys, survey lines, as the case may be).

By employing method A and formula (1) the standard error in the standard error can be obtained from formula (3) p. 304 where n is the number of groups. Formula (3) assumes that the deviations (x_i—M) are distributed in accordance with the normal function of probability (HELMERT 1924, p. 75). The relative standard error in the standard error ($\varepsilon_{\varepsilon_M}$) by using different numbers of groups (n) will be seen from the table on p. 305.

The uncertainty in the calculated standard errors is considerable in the case of those numbers of groups which have hitherto been customarily used (8—20). If, then, the standard error is to be calculated for an individual survey, it must be divided up into a relatively large number of partial surveys (groups). This brings us to a serious limitation inherent in method A.

In dividing up the survey into a considerable number of groups there is often the risk, provided the number of lines is not very large, of a systematic variation (tendency) appearing in the group results (cf. LANGSAETER 1926, NÄSLUND 1930). Under such circumstances formula (1) produces too high standard errors. Consequently the reduction of the accidental error is made at the cost of an increased risk of a systematic error.

The systematic error can however be reduced by applying formula (2) to the partial surveys, but the influence of the systematic variation from group to group is always present. Nevertheless the risk of a systematic variation, as mentioned above, is far less if the survey lines are divided up into partial sections, which are combined according to a certain system (cf. NÄSLUND 1930, method A: 2, p. 321). Apart from the above-mentioned risk of a systematically too high standard error, the method is also unsatisfactory from another point of view.

A common feature of method A is that the standard error (ε_M) of the actual survey is calculated from the standard deviation of the partial surveys. Such an extrapolation is invariably fraught with difficulties and must be regarded as a weakness in the method. We shall revert to this point later on.

Method C has apparently arisen out of the necessity, when applying method A, for dividing up the actual survey into a number of partial surveys; for not infrequently the lines included in a survey are so few in number that one cannot even divide them up into a smaller number of partial surveys.

In respect of method C and formula (2) the relative standard error in the standard error if the line-number is about 10 and over can be approximately estimated at the same size as in respect of method A and formula (1), *n* now meaning the number of survey lines. Frequently, therefore, method C produces a standard error which, if account is taken of the accidental error, is relatively certain (cf. p. 305). In surveying small forest areas, however, one must be able to divide up the survey lines into smaller units (cf. p. 340). However, the method possesses a further weakness.

It is generally possible to trace a systematic variation in the survey results of the individual lines (LANGSAETER 1926, NÄSLUND 1930, ÖSTLIND 1932; cf. also figs. 1—6, p. 308). Under such circumstances formula (2) produces too high a result because the systematic tendency from line to line is not eliminated. According to ÖSTLIND, in regard to the cubic volume per hectare of forestland, formula (2) should yield on an average about 20 % for large standard errors (*The National Forest Survey Commission* 1932, ÖSTLIND 1932). LANGSAETER estimates this systematic error at 10—20 % (*The Norwegian Forest Survey* 1938).

The discussion so far will no doubt have brought out the fact that in our search for a suitable method of computing the standard error we must make it an essential condition that we shall be able, if necessary, to reduce not only the systematic but also the accidental error of the standard error as far as possible.

Theory and Methodics.

Line surveying.

The difficulties associated with the calculation of the standard error in a regular line or sample plot survey are chiefly due to the fact that the survey lines are laid out systematically and not taken at random. When the direction of the line, the survey percentage and the position of the first survey line have been decided upon, the position of the other survey lines is indicated. The object of laying out the lines systematically is obviously to enhance the certainty of the results of the survey; and ÖSTLIND's abovementioned investigation clearly shows that it does so (ÖSTLIND 1932, p. 451).

It has previously been pointed out that the result of the survey often varies systematically from line to line (cf. p. 334), and we propose here to take a few examples and make a detailed analysis of the structure of the survey. We have chosen for this purpose the earlier national forest survey of the provinces of Västernorrland and Norrbotten, which represent different degrees of difficulty from the survey point of view.

In the national forest survey the survey lines have been divided up into 2-kilometre stretches (sections). The method of dividing them up is as follows: stretches having the same 2-km numbers form a field perpendicularly to the line-direction (block). In the following account the 2-km stretches are combined into 10-km stretches. We shall introduce here the term »survey blocks», or simply »blocks», for 10-km stretches bearing the same numbers.

The cubic volume per hectare of land area has been worked out for every line and block within the two provinces, Norrbotten being divided up into »Kustland» (coast district) and »Lappmark» (lapp district). Records have been taken of the cubic volume of pine and spruce and of the diameter-groups: 15 cm at breast height and over, 25 cm and over, and finally 35 cm and over. In figs. 1—6 the cubic volume per hectare has been expressed in graphs, the line-, respectively the block-number being taken as independent variables (p. 308—313).

The figures clearly demonstrate that in respect of both lines and blocks a marked systematic variation is often observable in the survey results. Not infrequently this tendency is strongly emphasized. A contributory factor is, of course, the cubic volumes of the pine and of the spruce being given separately, this being justified by the fact that it is necessary to be able to judge of the degree of certainty in these estimates each severally.

LANGSAETER has shown the existence of the same tendencies in regard to area determinations (LANGSAETER 1926).

As far as the »Lappmark» of Norrbotten province is concerned, the distribution shown in fig. 7 p. 314 of the cubic volume of spruce in diameter-group 15 cm and over has been further illustrated with the aid of the survey results taken from the 10-km stretches. It is clear from the figure that the cubic volume is in a certain degree a function of the distances to two coordinate axes. Manifestly, when working out the method of calculating the standard error, due account must be taken of the existence of tendencies of distribution which are illustrated here.

Before passing on to discussing the theory of standard error calculation we shall dwell a little on those methods which are based on the idea of combining the survey lines or parts thereof into groups (cf. NÄSLUND 1930, method A: 1 & 2).

Let us assume that the line survey that has been carried out is a 10 % survey and that it is divided up into 10 groups, each comprising a 1 % survey. The standard error calculated by this method is then approximately the correct standard error for that survey which is obtained if, out of the 100 possible 1 % surveys of the forest, we pick out at random 10, and form of them a 10 % survey (cf. LANGSAETER 1932, p. 446). It is however not this standard error we want to ascertain, but the standard error in the systematically laid-out ten-percentage line survey.

The idea of systematically laying-out the survey lines is in order that each line or part thereof (e. g. 10-km. stretch) shall represent that parcel of the forest which lies within half the line-distance on either side of the line (the partial stretch).

If we knew the result (x_i) of a total survey in respect of each one of these parcels the standard error of the line survey (ε_M) would be obtained by means of the formulae:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}{n}} \qquad (6)$$

$$\varepsilon_M = \frac{\sigma}{\sqrt{n}} \qquad (7)$$

where x_i is the result for the i-th line or partial stretch and n is the number of lines or partical stretches, all of which are assumed to be of equal length.

Now, we do not know $\bar{x}_i$, but must substitute for this quantity an approximate value. In our future discussion we shall imagine the lines to be divided up into partial stretches in the manner previously indicated, with the result that the line survey will be converted into a quadratic or rectangular system of squares (cf. fig. 7). At the same time we introduce the terms $\bar{x}_{uv}$ and x_{uv}, by which we shall denote respectively the true and the surveyed value of the quantity under investigation in respect of that square which lies in line u and block v.

We have shown in the foregoing that e. g. the cubic volume in respect of a square may be regarded roughly as a function of the position of the plot. The simplest assumption we can make regarding this association is that the investigated quantity of a square may be considered as the sum of three quantites: 1) one part α which is the same for all squares, 2) one part β_u, which is constant for each line, and 3) one part γ_v, which is constant for each block. For this purpose β_u and γ_v are chosen in such a way that the sum of all the β values is equal to 0, as is also the sum of all the γ values. On these hypotheses, the cubic volume of a given square works out at: $\bar{x}_{uv} = \alpha + \beta_u + \gamma_v$. The problem is how to get from the survey material the best values of α, β_u and γ_v. This is clearly obtained if we make $\alpha = M$, $\beta_u = m_u - M$, $\gamma_v = m_v - M$, where M denotes the average for all the squares, m_u the average for the line u, and m_v the average for the block v. From this we obtain:

$$\bar{x}_{uv} = m_u + m_v - M, \qquad (8)$$

and consequently the difference between the surveyed and the true value of the quantity investigated: $x_{uv} - m_u - m_v + M$; further, the sum of the squared deviations:

$$\sum_{u,v} (x_{uv} - \bar{x}_{uv})^2 = \sum_{u,v} (x_{uv} - m_u - m_v + M)^2 \qquad (9)$$

This sum of the squared deviations is thus equal to the individual squares' deviations from the »great mean» (M) after they have been corrected for the systematic deviations that originate in the different lines and blocks.

The problem confronting us here is analogous to the need, when carrying out field experiments in the sphere of agricultural research, to be able to eliminate the effect of differences in the fertility of the soil. The trend of thought developed in the above discussion has, moreover, common points of contact with the mode of thinking employed in agricultural research (cf. BACHÉR 1933). The derivative expression for the sum of the squared deviations (equ. 9) is identically the same as »the sum of the squares of the errors» in the application of FISCHER's analysis of variation in agricultural experiments. For the theory and technical carrying-out of the analysis of variation the reader is referred to FISCHER's book *»Statistical Methods for research workers»* (1932).

To return to equation 9; This gives us the standard deviation of the squares according to the formula:

$$\sigma = \sqrt{\frac{\sum\limits_{u,v} (x_{uv} - m_u - m_v + M)^2}{(n - 1) - (l - 1) - (b - 1)}} \quad \ldots \ldots \ldots \quad (10)$$

where n is the total number of squares and l denotes the number of lines and b the number of blocks. The expression in the denominator indicates the number of degrees of freedom (to use FISHER's terminology). We shall not dwell further upon the mathematical significance of this idea but refer the reader to FISHERS' book.

The standard error of the mean (ε_M), or the survey error we are seeking, is then calculated according to the formula: $\varepsilon_M = \frac{\sigma}{\sqrt{n}}$

thus we get:

$$\varepsilon_M = \sqrt{\frac{\sum\limits_{u,v} (x_{uv} - m_u - m_v + M)^2}{n\,[(n - 1) - (l - 1) - (b - 1)]}} \quad \ldots \ldots \ldots \quad (11)$$

So far we have for simplicity's sake assumed that the length of the lines in the squares is constant. If the line-length varies (p_{uv}), the deviations must be weighted, and weighted line, block and great mean values must be introduced in formula 11. On the analogy of a simple weighting system used by LINDEBERG (1923), we get the following formula for the standard error.

$$\varepsilon_M = \sqrt{\frac{n}{(n - 1) - (l - 1) - (b - 1)} \cdot \frac{\sum\limits_{u,v} (p_{uv})^2 \, (x_{uv} - m_u - m_v + M)^2}{\left(\sum\limits_{u,v} p_{uv}\right)^2}} \quad (12)$$

$$M = \frac{\sum\limits_{u,v} p_{uv}\, x_{uv}}{\sum\limits_{u,v} p_{uv}}; \quad m_u = \frac{\sum p_{u.}\, x_{u.}}{\sum p_{u.}}; \quad m_v = \frac{\sum p_{\cdot v}\, x_{\cdot v}}{\sum p_{\cdot v}}$$

In formula (12) the sign x_{uv} denotes the result of the survey and p_{uv} the length of line for that square which lies in line u and block v; m_u and m_v are weighted line and block mean values, and M the weighted mean value for all squares (the great mean). n indicates the total number of squares, l the number of lines and b the number of blocks.

The discussion on p. 337 and formula 11 (or 12) demonstrate that the possibility of eliminating systematic deviations on account of the position of the squares in different lines and blocks becomes greater if the lines and the blocks are of approximately equal length than if there is wide variation, For the purpose, therefore, of calculating by this method the standard error in regard to a forest with a very irregular figure, the forest should be divided up into more regular sectors (sub-areas) and the error calculated for each one separately, after which the standard error in respect of the entire forest (ε_M) is obtained from the formula:

$$\varepsilon_M = \sqrt{\left(\frac{P_1}{P}\right)^2 \cdot \varepsilon_1{}^2 + \left(\frac{P_2}{P}\right)^2 \cdot \varepsilon_2{}^2 + \left(\frac{P_3}{P}\right)^2 \cdot \varepsilon_3{}^2 \; . \; . \; . \; ,} \quad . \; . \; . \; (13)$$

where P denotes the total length of line for the entire forest, P_1, $P_2 \ldots$ etc. the corresponding lengths for the sectors, and ε_1, $\varepsilon_2 \ldots$ etc. the standard error of the sectors.

In view of the circumstances mentioned above, no account has been taken in formula (12) of the different weighting of the line and block mean values (m_u and m_v).

Evidence that there is a systematic variation in the x_{uv} values is obtained if the standard error is, besides, calculated according to the following formula, where the signs mean the same as in formula (12):

$$\varepsilon_M = \sqrt{\frac{n}{n - 1} \cdot \frac{\sum\limits_{u,v} (p_{uv})^2 (x_{uv} - M)^2}{\left(\sum\limits_{u,v} p_{uv}\right)^2}} \quad . \; . \; . \; . \; . \; . \; . \; . \; (14)$$

If formula (14) produces a greater standard error than formula (12), there exists a systematic variation in the x_{uv} values, which have been eliminated to a corresponding degree in the calculation of the standard error. If the contrary is the case, either no systematic tendency has existed or else the application of the method has not succeeded in eliminating it. If the figure of the forest is very irregular and a greater degree of accuracy is required in determining the standard error, then dividing up the figure should be tried, provided it is not possible by some other means, e. g. by graphs (cf. figs. 1—6), to confirm that the survey material does not exhibit any marked distribution-tendencies.

According to the formulae given above, the standard error relates to the population and is thus a measure of the standard deviation between repeated surveys as comprehensive as the actual survey. If, on the other hand, we should desire to regard the standard error as a measure of the deviation between the result of the survey and the true value of the quantity under investigation, we can obtain an estimate thereof if the standard-error formulae are multiplied by the factor $\sqrt{\frac{100 - \alpha}{100}}$ (cf. Czuber 1924), where α is the survey percentage. This correction factor is only of any practical significance in the case of high-percentage surveys.

We propose to say something about the size and form of the squares. The smaller the squares, the smaller will also be, in principle, the relative standard error in the standard error owing to the increased number of squares. Some idea of this is given by the summary on p. 305, where n now means the number of degrees of freedom. It is moreover easy to realize that the smaller the size of square the greater the possibility of eliminating the systematic tendencies.

In practice, however, the task of working out the calculations becomes arduous if there are too many squares, as the result of the survey for which the standard error is to be computed has to be worked out for each square. This is particularly laborious in regard to the cubic volume, so that the number of squares should be carefully weighed against the desired degree of precision in determining the standard error. For this purpose it may often be expedient to combine the survey lines into groups, so that e. g. the lines Nos. 1—3 form the first group, Nos. 4—6 the second, and so on. The squares will thus contain several lines.

In regard to determinations of area, there is a further point of view, namely, that the size of the squares may not be so small that the variation becomes alternative, in which case the method should not be used (cf. LANGSAETER 1926, p. 18). For this reason one must here be content with the degree of accuracy obtained with relatively large squares. In the practical application of the method it is generally possible to use the same squares for determinations of both area and cubic volumes as well as growth, etc., which is a great advantage from the point of view of labour.

As to the form of the squares, it will readily be understood that they should have their smallest extent in the direction in which the systematic tendency is strongest. In order, however, to avoid alternate variation in determining the area, it is possible to make deviations from this rule. If nothing is known of the existence of any systematic tendencies in the material, the simplest way, of course, is to make the squares quadratic.

It will no doubt have been gathered from the above that the method expounded here for calculating the standard error affords considerable possibilities of computing a standard error that is reliable in regard to both systematic and accidental errors. For deducing standard errors in determinations of the area, the method should, as a rule, only be used when the survey comprises extensive forests.

In connection with the method dealt with in the foregoing it should perhaps be mentioned that there is another method of calculation which may be regarded as a variant of method C (cf. p. 334). This method is based on a division into blocks and squares similar to that in the preceding method. The standard error is computed with the aid of the differences between the survey results taken from successive squares in the same block, it being possible to make use of differences both of the first order (formula 15) and of the second order (formula 16). The standard error of the v-th block (ε_v) are obtained by means of the following two formulae (LINDEBERG 1923, LANGSAETER 1927):

$$\varepsilon_v = \sqrt{\frac{r \cdot \sum_{u=1}^{r-1} (p_{uv} + p_{(u+1)v})^2 \; (x_{uv} - x_{(u+1)v})^2}{2\,(r-1)\left(2 \sum_{u=1}^{r} p_{uv}\right)^2}} \quad \ldots\ldots\ldots \quad (15)$$

$$\varepsilon_v = \sqrt{\frac{r \cdot \sum_{u=2}^{r-1} (p_{uv})^2 \left(x_{(u-1)v} + x_{(u+1)v} - 2x_{uv}\right)^2}{6\,(r-2)\left(\sum_{u=1}^{r} p_{uv}\right)^2}} \quad \ldots\ldots \ (16)$$

In these formulae x_{uv} denotes the result of the survey and p_{uv} the line-length of the square lying in line u and block v, and r the number of squares (lines) in the v-th block. Then the standard error of the entire forest (ε_M) is calculated analogous to the formula (13) p. 339.

Of these two formulae the first is in principle more favourable in regard to the accidental error in the standard error because the number of differences becomes greater, and the second formula in regard to the systematic error. If a systematic tendency runs through the results of the survey, both formulae give excessively high standard errors and formula (15) a higher value than formula (16). The method is not able to eliminate the average tendency from square to square (cf. p. 334).

This method is apparently applicable when a greater degree of accuracy is required and the former method (formula 12) cannot be employed, as well as in calculations that need not be so exact, when the number of lines is too small to permit of LANGSAETER's formula (formula 2, p. 304) being directly applied. In regard to the size and form of the squares, the same observations apply as those made on the preceding method.

Sample plot surveying.

The method indicated in the foregoing section for computing the standard error in line surveying likewise applies to regular sample plot surveying. There are no peculiar difficulties in the way of its employment. The results of the sample plot survey are worked out in respect of each square, after which the standard error is obtained from formula (12). For determining the area, however, it is not infrequently preferable to employ another method of calculation. We shall now discuss this method in detail.

In regular sample plot surveying the individual sample plot is, as a rule, made relatively small so that only in exceptional cases will it be necessary to divide it up into different kinds of soil or other areal determinations. The method of procedure employed in certain quarters, namely that of moving the plot in such cases so that it only comprises one type of ground, is from a statistical point of view incorrect and should not be used. Under the conditions mentioned above the variation will in principle be alternate, i. e. the characteristic is either to be found in the plot or else it is not. The obvious thing, then, is to try to trace the standard error in the area determinations by means of the classic principles of alternate statistics.

Attempts along those lines have been made by SCHUMACHER and BULL (1932) and also by MUDGETT and GEVORKIANTZ (1934). For this purpose the entire survey has in principle been refered to a certain urn scheme, the former authors having applied BERNOULLI's scheme and the latter POISSON's or LEXI's. LANGSAETER has demonstrated that different types of area determinations have the same distribution-tendencies as we have found here in the cubic volume (cf. p.

336). Indeed it is only natural that this should be so. MUDGETT's and GEVORKIANTZ' use of POISSON's or LEXI's theorem may therefore be described as a development in the right direction towards the complex urn schemes of reality.

It is possible to adapt the method of calculation more or less exactly to the actual circumstances in two different ways; either by combinations of POISSON's and LEXI's schemes or else by dividing up the survey into smaller units for which the simplest urn scheme (BERNOULLI's) is approximately applicable. Following the method employed in computing the standard error of the cubic volume, I have chosen the latter alternative.

We shall now revert to the method which we have employed in the foregoing (p. 337), of dividing up the forest into squares. These squares we regard as urns containing black and white balls spread about more or less irregularly (at random). The balls represent those sample plots into which the individual square is imaginarily divided. For this purpose e. g. the black balls denote the type of ground in respect of which the standard error is to be computed. The survey of the square is carried out on strictly methodically, but provided the distribution of the area within the square does not exhibit any systematic tendencies, we can approximately regard the sample plot survey as if it were drawings from an urn, none of the balls when drawn being put back into the urn. The number of drawings here correspond to the number of sample plots in the square that fall out in the course of the survey. Under such circumstances it is easy to estimate the standard error in the sample plot survey for the individual square.

Let N denote the total number of sample plots, in which the i-th square can be imagined to have been divided up (the area of the square: the area of the sample plot) and p_i the number of sample plots in this square that fall out in the course of the survey, and of which x_i per cent lie on the type of ground in respect of which the standard error is to be determined. This standard error (ε_i) is then obtained by means of the following formula (CRAMÉR 1926):

$$\varepsilon_i = \sqrt{\frac{(N - p_i)}{(N - 1)} \cdot \frac{x_i\,(100 - x_i)}{p_i}} \quad \ldots\ldots\ldots\ldots \quad (17)$$

But $N = \frac{100\,p_i}{\alpha}$, where α means the survey percentage. Formula (17) can also, therefore, be written:

$$\varepsilon_i = \sqrt{\frac{(100 - \alpha)\,x_i\,(100 - x_i)}{100\,p_i - \alpha}} \quad \ldots\ldots\ldots\ldots \quad (18)$$

By this means (formula 18) the standard error can be calculated for each square, after which the standard error in the survey result for the whole forest (M) is obtained from the formula:

$$\varepsilon_M = \sqrt{\left(\frac{p_1}{P}\right)^2 \cdot \varepsilon_1^2 + \left(\frac{p_2}{P}\right)^2 \cdot \varepsilon_2^2 + \left(\frac{p_3}{P}\right)^2 \cdot \varepsilon_3^2 \ldots\ldots \left(\frac{p_n}{P}\right)^2 \cdot \varepsilon_n^2} \quad (19)$$

where P denotes the total number of plots (Σp_i), and n the number of squares.

The factor $\frac{N - p_i}{N - 1}$ frequently amounts to such trifling values that it may be disregarded. Formula (19) then acquires the following simple form:

$$\varepsilon_M = \sqrt{\frac{\sum_{i=1}^{n} p_i x_i \, (100 - x_i)}{(P)^2}} \quad \ldots \ldots \ldots \ldots \quad (20)$$

The method outlined above of finding the standard error in area-determinations does not involve much work and can not unfrequently be employed, but only provided that there is no very marked systematic tendency in the squares. The risk of this is of course less the smaller the squares are made; on the other hand, the number of sample plots within the square should not be too small.

The employment of methodics in Line surveying.

As previously mentioned (p. 332), I have employed the method developed here for calculating the standard error in the course of a certain enquiry made on behalf of the National Forest Survey Commission. In the following pages we shall discuss some of the results of this investigation that are of particular interest from the point of view of method.

The material consists of the former National Forest Survey of the provinces of Västernorrland and Norrbotten and the watershed of the river Ljungan. The width of the belt was 10 metres. The compiling of the results was confined to calculating the standard error in the cubic volume per hectare according to the uncertainty of the basal area. Table 2 (p. 323) gives some data intended to illustrate the nature of the material from the survey point of view.

When the standard error is worked out by means of formula (12), a calculation table of the form given in Tab. 3 p. 324 may suitably be used.

The standard error has been calculated for pine and spruce separately and for the diameter groups: 15 cm. at breast height and over, 25 cm. and over, 35 cm. and over. The results of the calculations are combined in Table 4 p. 326.

Formula (14) has throughout given higher standard errors than formula (12). The standard error accord. to formula (14) in respect of the different diameter groups on an average comprise 116, 112 and 109 per cent respectively of the standard error accord. to formula (12). This shows that systematic tendencies have occurred in the material and have been eliminated in a corresponding degree when the standard error has been calculated according to formula (12) (cf. p. 339). In view of the large number of squares, the standard errors calculated according to formulae (12) and (14) have been found with a considerable measure of certainty.

For Norrbotten and Västernorrland provinces a comparison has been made between LANGSAETER's formula (formula 2), and formula (12), Tab. 5 p. 328). The uncertainty of the calculated individual standard errors according to LANGSAETER's formula are considerable in view of the small number of lines. We can therefore only discuss the average tendency in respect of the whole comparative material. Table 5 shows that the standard error according to LANG-

SAETER's formula on an average comprise 114 per cent of the accord. to formula (12) computed standard error.

It will be seen from figs. 1—7 and tab. 5 that systematic tendencies occur in the material. Under such circumstances it has previously been shown (cf. p. 334) that LANGSAETER's formula gives too high results. According to the experiences of others (*The National Forest Survey Commission 1932*, ÖSTLIND 1932, *The Norwegian Forest Survey 1938*), this over-estimation amounts on an average to 10—20 per cent., which quite fairly corresponds to the difference found here between LANGSAETER's formula and formula 12. It would appear, therefore, that the method expounded here enables an accurate estimate of the standard error to be made even when the survey material exhibits systematic distribution-tendencies.

Summary.

When applying the calculation of the laws of probability to practical problems it is often difficult to produce evidence to justify its application. The calculation of probabilities in the sense intended here is based on purely mathematical grounds and consequently depends on certain definite hypotheses, which can only exist in a concrete body of material with a greater or less degree of approximation. It is this degree of approximation that determines the value of the calculation. It is therefore essential, if the calculation of probabilities is to be of any practical use, that methods of application should be worked out which are not dependent upon any special hypotheses but only on assumptions of more or less universal applicability. Any step made in this direction must be regarded as an achievement.

The aim of this paper has been to make some contribution towards the development along these lines of the methodics used in finding the standard error in line and sample plot surveying.
