634.91.5

MEDEDELINGEN VAN DE LANDBOUWHOGESCHOOL TE WAGENINGEN/NEDERLAND
57 (5) 1–60 (1957)

ESTIMATION OF STANDING TIMBER

(*MET EEN SAMENVATTING IN HET NEDERLANDS*)

by (door)

F. E. ESSED

Afdeling Houtmeetkunde en Afdeling Wiskunde van de Landbouwhogeschool

(Received/Ontvangen 19.3.'57)

CONTENTS

2C4[9]C12

INTRODUCTION

A group of trees on a certain area that can be considered a sylvicultural as well as an administrative unity will be called a stand.

Estimation of the volume of standing timber is necessary for many purposes. We will mention some of these:

1. One of the principal objects of sylviculture is the production of the largest possible quantity of valuable timber. The estimation of the volume is necessary to compare the effects of different sylvicultural treatments (establishing, thinning, rotation).
2. The long duration of the production necessitates periodical estimations of the standing timber in order to regulate the annual felling.
3. We also require an estimate to predict the price of timber to be sold.
4. The total volume of the standing timber of a country is important for the national economy and therefore periodical inventarisation is required to estimate this volume.
5. Determination of the increment of the standing timber is very important to judge whether the profit of the management is satisfactory or not. Its estimation is closely related to that of the volume.

An obvious method of estimation is to determine the volume of each separate tree, after which the total volume can be found by addition. The estimation of the volume of a single tree is usually done by measurement in sections. This method, however, takes so much time and is in consequence so expensive, that it can not be used in forestry practice. Our aim should be a simplification.

Let the trees of a certain stand possess a measurable property of interest A, whose determination takes much time. Apart from A the trees possess one or more other properties b whose determination is quick and simple. If a known functional relation exists between b and A, it is sufficient to measure only b and to determine A by the functional relation.

As a rule a functional relation does not exist. In many cases however the following assumptions hold:

1. For each fixed value of b there exists a random variable $\underline{A} = \underline{A}_b$ with expectation $E(\underline{A}_b)$. The experimental data are values which $\underline{A}$ takes.
2. The function g: $b \rightarrow E(\underline{A})$, that defines the relation between b and $E(\underline{A})$ belongs to a class of simple functions.

The problem to find the best function g, given a large number of observations, is a regression problem.

The knowledge we need about A is sufficiently covered by $E(\underline{A}_b)$ and the distribution of b. If $E(\underline{A}_b) = g(b)$ is known, then the procedure can be limited to measuring b.

As an example we mention the volume tables, in which the expectation of the volume $\underline{V}$ can be found for values of two or more easily measurable properties of the tree (e.g. diameter at breast height ($d_{1.30}$) and total height (h)). It may be recalled that a (volume) table can be considered an expression for a function.

In order to construct a volume table the measurable quantities V, $d_{1.30}$ and h of a great number of trees of a certain wood species are determined, and then a function is fitted.

The way of fitting this function, i.e. the construction of the tables, can differ. (See e.g. KUIPER (1954)). In forestry very often the so called form factor method is used, which will be briefly outlined here.

From the explanations of the concept "form" that are given in forestry literature, we may conclude that this term is not identical with "form" in the geometrical sense (geometrical multiplication). In forestry a tree is always considered a solid of revolution. The following definition of form in forest mensuration covers the usual interpretation.

Two trees B and B′ have the same form if the coordinates (x_1', x_2', x_3') of the points of B′ are found from the coordinates (x_1, x_2, x_3) of the points of B by the following transformation.

$$\begin{aligned} x_1' &= px_1 & & & + g_1 \\ x_2' &= & px_2 & & + g_2 \\ x_3' &= & & rx_3 & + g_3 \end{aligned}$$

(x_1, x_2 resp. x_1', x_2' are the horizontal coordinates and x_3 resp. x_3' the vertical coordinate).

Let g_k be the area of the horizontal cross section at height (1–k)h $0 < k < 1$.

As $f_k = V \cdot (g_k h)^{-1}$ is the same for trees of the same "form", this quantity is considered characteristic for form, and is called the normal form factor.

However, without affecting the invariance of the normal form factor we may use a definition of form more general than the preceding:

B and B′ (see above) have the same "form" if:

$$\begin{aligned} x_1' &= p_1x_1 + q_1x_2 + r_1x_3 + s_1 \\ x_2' &= p_2x_1 + q_2x_2 + r_2x_3 + s_2 \\ x_3' &= r_3x_3 + s_3 \end{aligned}$$

The proof of the equality of the form factors of B and B′ is as follows:

The volume of B′ is $V' = \begin{vmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ 0 & 0 & r_3 \end{vmatrix} . V = r_3(p_1q_2 - q_1p_2)V.$

The total height of B' is $h' = r_3h$, and the area of the cross section at height $(1-k)h'$ is: $g' = (p_1q_2 - q_1p_2)g$.

Consequently $f_k' = V'(g'h')^{-1} = V(gh)^{-1} = f_k$. Hence we see that the normal form factor is invariant under many transformations of the tree.

In case the trees of a certain wood species all have the same "form", f_k is a useful quantity. If f_k is known, the measurements can be confined to h and g_k, and then the estimated volume is $V = h \cdot g_k \cdot f_k$.

The height of the cross section (g_k) however, varies with the height of the tree. It is much simpler and quicker to measure diameters at a convenient constant height. As a consequence of this the diameter at breast height (dbh) was introduced. (In many countries dbh is measured at 1.30 m above the ground). In analogy with the procedure for the normal form factor the volume is calculated with $V = \frac{\pi}{4}(d_{1.30})^2\, h\, f_{1.30}$.

Research workers often occupied themselves with the number $f_{1.30}$. They found $f_{1.30}$ to be dependent on $d_{1.30}$ and h, and in some cases on h only.

The formula $f_{1.30} = (d_{1.30})^{-2} \cdot d_k^2 \cdot f_k$ is frequently mentioned in forestry literature.

The form factor is obtained and used as follows:

For a great number of trees, V, $d_{1.30}$ and h are determined. For every tree $f_{1.30} = V(g_{1.30}\, h)^{-1}$ is computed. Then a simple function giving the relation between the expectation of $f_{1.30}$, $d_{1.30}$ and h (or of $f_{1.30}$ and h) is fitted.

With the estimated expectations of $f_{1.30}$ for fixed values of $d_{1.30}$ and h, the volume is estimated. It is remarkable that $f_{1.30}$ is also called "form factor", although any notion of constant form is absent here. It seems to us that the value of this so called form factor method is generally overestimated.

According to another well known method, the expectation of V is considered a function of the quantities d_n, $d_{1.30}$ and h. In particular the number $q = d_n \cdot (d_{1.30})^{-1}$ is used[1]). This number is supposed to be a better characteristic for "form" and is called "formquotient". It seems that one and the same function holds approximately for many species; on the other hand a better fit can be obtained when one restricts oneself to one species.

In case a volume table is available the total volume of the stand can be found by measuring $d_{1.30}$ and h of each tree. This is still too laborious and further simplification is necessary. As a height measurement takes much more time than a diameter measurement (according to recent investigations of the Forestry Research Institute of the University of Wageningen resp. 1.01 min and 0.109 min.) we have to limit the height measurements. First of all we replace the complete measurements of all heights by a *regression function* $d_{1.30} \rightarrow h$, which is graphically obtained (height curve).

In many cases the fitting can be considerably simplified by a transformation of scales. Logarithmic transformation (single or double logarithmic graph paper) generally enables one to fit a straight line, for example

$$h = a \log d_{1.30} + b \quad \text{(Henriksen)}$$

$$\log h = a \log d_{1.30} + b \quad \text{(Stoefels)}\ .$$

[1]) n ($\neq$ 1.30) indicates the constant height above the ground at which d_n is measured.

(In case linearity can be assumed, a rather small number of points suffices to obtain a fairly good fit).

The height curve is found to be almost the same in stands of a certain wood species with the same $h_{\bar{d}}$ and average diameter $\bar{d}$. In other words $\bar{d}$ and $h_{\bar{d}}$ determine the so called standard height curve. Height measurements are then limited to 4–10 measurements of $h_{\bar{d}}$.

Instead of standard height curves also standard volume curves are used. KOPEZKY and GEHRHARDT take for example a linear relationship between $(d_{1.30})^2$ and V in a stand. SPIECKER and VON LAER, constructed relevant tables (1951).

BERKHOUT (1920) introduced the formula $V = \gamma d^{\beta}$, in which β is constant for a certain species, and γ a number that depends on the stand. STOFFELS (1953) using BERKHOUTS formula, with γ however as a function of $\bar{d}$ and $\bar{h}$, also constructed tables.

In all methods discussed so far we assumed that the diameters of all trees are measured. We can achieve a further simplification if we take sample plots over the area and estimate the volume pro ha within those plots in one of the ways described above. The average of these estimates is used as an allround estimate of the volume pro ha.

The marking out of plots however, takes much extra time. Consequently this method will only be more efficient if the cruising percentage is small enough (in practice mostly $<20\,\%$).

The method which we will consider is as follows:

Let $\mathcal{S}(N)$ and $\mathcal{S}(\bar{v})$ be estimates of N and $\bar{v}$ (the number of trees pro ha and the average volume per tree resp.). As the volume V pro ha equals $V = N\,\bar{v}$, an estimator of V is [1]):

$$\mathcal{S}(V) = \mathcal{S}(N)\,\mathcal{S}(\bar{v})$$

If one (more) tree(s) with the average volume $\bar{v}$ is (are) known, for example recognised by a particular diameter, then $\mathcal{S}(\bar{v})$ can be found by measuring the volume of this (these) tree(s) (model tree(s)).

Until now $\mathcal{S}(\bar{v})$ has been determined mainly as $v_{\bar{g}}$, the volume of the tree with the average basal area. (If a linear relation between v and d^2 (KOPEZKY-GEHRHARDT) holds $\bar{v} = v_{\bar{g}}$). This requires the determination of $d_{\bar{g}} = n^{-1}\,\Sigma d^2$ and $h_{\bar{g}}$ (the height of the tree with the mean basal area).

As $\bar{d}$ can be estimated easily, we use a function (f) of $v_{\bar{d}}$, (the volume of the tree with diameter $\bar{d}$) instead of $\bar{v}$.

In this thesis we develop a.o. some theory concerning the estimation of V, for which a formula of the kind

$$\mathcal{S}(V) = \mathcal{S}(N)\cdot f\{\mathcal{S}(\bar{d}),\ \mathcal{S}(h_{\bar{d}})\}$$

is taken as point of departure. The function f as well as the estimators $\mathcal{S}(N)$, $\mathcal{S}(\bar{d})$ and $\mathcal{S}(h_{\bar{d}})$ are studied separately.

[1]) Estimators will often be indicated by the letter $\mathcal{S}$

CHAPTER I

SOME WELL KNOWN METHODS FOR THE DETERMINATION OF THE STEM NUMBER [1]

1. THE SAMPLE PLOT METHOD

The most commonly used method for the determination of the stem number is as follows: sample plots are chosen in a stand; in each plot the trees are counted; an average is calculated and finally this number is multiplied by an appropriate factor to get the number of trees per square unit. The plots are taken either at random or systematically. (Systematic sampling is in general more efficient than random sampling, although less is known about the variance of the estimator). According to the shape of the plots the methods are called circular plot method, strip method etc. Sometimes square plots are used.

Given a total sample area, better estimates are obtained with a greater number of smaller sample plots. The size of the sample plots however, cannot be taken too small (effect of boundary trees,[2]) extra labour). A good standard for the practice is the use of plots of such a size that the expected number of trees in the plot is about twenty.

Due to the effect of boundary trees, circles and squares are preferred to narrow strips; moreover, circles are generally preferred because their establishing and the counting of the trees take less time.

In modern survey optical instruments are used instead of tapes to stake out the sample plots. This is an important improvement. The quickest method however, is the counting of the trees that occur inside a circle with a radius (r) without the use of any optical instrument but the human eye. The Dutch Forest Service introduced and developed a method by which the trees that occur inside a circle with a radius $r = 7.98$ m are counted in this way. Experiments showed one can master this method rather quickly.

2. THE DISTANCE METHOD OF BAUERSACHS AND KÖHLER

BAUERSACHS (1942) has developed a method for the determination of the stem number by the measurement of distances. After having chosen a tree at random he measured the distance from this tree to the next nearest tree. (For the sake of convenience the nearest tree will be called the first, the next nearest tree the second etc.). The average a (in meters) of these distances was used for the estimation of the stem number (N pro ha) by the empirical formula $N = 8500\ a^{-2}$.

For his experiments BAUERSACHS used "Kunstbestände", i.e. he drew dots at random on a paper and considered the result the map of a stand. The dots represented the trees.

KÖHLER (1954) introduced the formula $N = 10.000(2n)^2 a_{23}^{-2}$ in which a_{23}

[1]) The number of trees per unit area (usually ha) is called stem number.

[2]) Trees that occur in the neighborhood of the circumference of a sample plot, which makes it difficult to judge whether these trees are inside or outside the sampling area, are called boundary trees.

is the average of the distances from any chosen tree to the second and the third tree and n the number of chosen trees.

BAUERSACHS did not publish a theoretical explanation of his formula.

KÖHLER however ventured the following explanation:

"In einem Quadratverband von 1.00 m stehen die vier nächsten Stämme in einem Abstand von 1.00 m, die folgenden vier Stämme in einem Abstand von 1.41 m von dem Stamm, von dem aus die Messungen vorgenommen werden."

"Bei einer natürlichen Verteilung der Stammabstände werden von den acht nächsten Stämme nur zufälligerweise mal zwei, wohl nie mehrere Stämme in einem genau gleichen Abstand zu ein und demselben Stamm stehen. Es muss daher angenommen werden, dasz von den vier nächsten Stämmen des unterstellten Quadratverbandes bei einer natürlichen Verteilung der nächste und der zweitnächste in einem kleineren Abstand als 1.00 m stehen, die folgenden 2+2 Stämme sich auf Abstände von über 1.00 bis 1,41 m verteilen und so weiter. In einer Entfernung von 1.00 m würde danach der 2.5-nächste Stamm, in einer Entfernung von 1.41 m der 6.5 – nächste Stamm stehen usw. Verbinden wir diese zwei festsehenden Punkte in einem Koördinatensystem durch eine Kurve nach der Gleichung $y^t = px$, deren Auflösung $y^{2.78} = 0.4\,x$ ergibt, so erhalten wir für die uns interessierenden zweit-, dritt- und viertnächsten Stämme theoretische Abstände von 0.92 m, bzw. 1.07 m und 1.19 m. Eine gute Übereinstimmung mit der Messungen tritt somit klar zu Tage. Daraus kann für die Praxis gefolgert werden, dasz der zweit- und der drittkleinste Stammabstand zu messen wären, deren Mittel dem Abstand bei einem Quadratverband entspricht:

$$\tfrac{1}{2}(0.92 + 1.07) = 0.995 \approx 1.00$$ "

There are severe objections to this explanation:

1. The concept "natürliche Verteilung" is not defined.
2. a. The concept "2.5 nächste Stamm" is not defined.
   b. No model is mentioned with respect to which the number 2.5 (or 6.5) occurs in some way or other.
3. The choice of the function $y^t = px$ is not motivated.

It does not seem possible to replace these arguments by others that are logically more consistent.

J. WECK (1953) e.g. when discussing KÖHLERS article writes: "Die mathematische Beweisführung, die A. KÖHLER in seinem oben erwähnten Aufsatz zur Begründung seines Verfahrens bringt, ist recht originell, aber nicht ganz zwingend".

CHAPTER II

SOME REMARKS ON THE DISTANCE METHOD OF BAUERSACHS AND KÖHLER

Early in 1954 the author did some preliminary studies concerning the distance method of BAUERSACHS and KÖHLER, a summary of which was published in the Indian Forester of May 1956.

The basic idea was that the distance a_n from a tree A to its n^{th} neighbor, can be considered the radius of the smallest circle with centre A that contains at least (n + 1) trees.

In each sample spot is $r_n < a_n < r_{n+1}$ (r_{n+1} is the radius of a circle containing n+1 trees). I therefore assumed $\rho_n < E(\underline{a}_n) < \rho_{n+1}$ when ρ_n is the radius of a circle of such an area that n is the expected value of the number of trees in that area).

Next I suggested the following approximation:

(2,1) $$E(\underline{a}_n) = \{(n + \tfrac{1}{2})S\}^{\frac{1}{2}} \cdot \pi^{-\frac{1}{2}}$$

($S = \pi\rho_1^2$ is the size of an area of which 1 is the expected value of the number of trees).

Under the given assumptions $\bar{a}_n$ will be an estimator of $\{(n + \tfrac{1}{2})S\}^{\frac{1}{2}} \cdot \pi^{-\frac{1}{2}}$. Let S = 1, then we conclude from (2, 1) $E(\underline{a}_1) = 0.69$; $E(\underline{a}_2) = 0.89$; $E(\underline{a}_3) = 1.06$; $E(\underline{a}_4) = 1.20$. Comparing these values with the data published by Köhler we notice some resemblance. (Table 1.)

TABLE 1.

	$\underline{a}_1$	$\underline{a}_2$	$\underline{a}_3$	$\underline{a}_4$
Köhler	0.50–0.85	0.85–1.00	1.00–1.10	1.15–1.25
Bauersachs		0.81–1.00		
Bauersachs		mean 0.92		
Formula (2,1)	$E(\underline{a}_1) = 0.69$	$E(\underline{a}_2) = 0.89$	$E(\underline{a}_3) = 1.06$	$E(\underline{a}_4) = 1.20$

Encouraged by these results the author started an investigation. The Forest Research Institute of the University of Agriculture in Wageningen enabled him to use some stand maps on which the position of each tree was indicated by a dot. 6 maps of 24 years old Douglas fir (*Pseudotsuga taxifolia*) plantations in the forest range Esbeek (Netherlands) were used. On these maps trees (dots) were chosen at random, the distances a_1, a_2 and a_3 were measured and their averages determined.

From (2, 1) we find the following estimators:

(2,2) $$\underset{\cdot}{S}(N) \approx 0.48\, \bar{a}_1^{-2}; \quad \underset{\cdot}{S}(N) \approx 0.80\, \bar{a}_2^{-2}; \quad \underset{\cdot}{S}(N) \approx 1.12\, \bar{a}_3^{-2};$$

$\underset{\cdot}{S}(N)$ is an estimate of the number of trees pro square unit.

Applying (2, 2) we find estimates of N, which we compare with the known values of N. (Table 2.)

TABLE 2.

Stand number	Number of distances measured	True number of trees pro ha	Number of trees pro ha estimated with			Differences in % of true number		
			$\bar{a}_1$	$\bar{a}_2$	$\bar{a}_3$	a_1	a_2	a_3
1	50	2320	2270	2460	2395	− 2.2	+ 6	+ 3.2
6	50	1920	1870	1895	1790	− 2.6	− 1.3	− 6.8
2	30	1481	1395	1325	1375	− 5.8	− 10.7	− 6.8
4	30	1556	1395	1445	1475	−10.9	− 7.5	− 5.5
5	25	1312	1135	1445	1475	−13.5	+10.2	+12.5
9	40	1715	1870	1810	1610	+ 9	+ 5.5	− 6.2

Additionally, field work was done in some 60 years old Scotch pine (*Pinus silvestris*) stands in the Forest range Oostereng in Wageningen. From regularly spread spots in the stands the nearest tree was chosen and from this tree the distances a_1, a_2 and a_3 were measured with a tape. (In one of the stands only a_1 and a_2 were measured). The number of trees was estimated with formula (2, 2) and compared with the true number. The results are given in table 3.

TABLE 3

Stand number	Number of distances measured	True number of trees pro ha	Number of trees pro ha estimated with			Differences in % of true number		
			$\bar{a}_1$	$\bar{a}_2$	$\bar{a}_3$	a_1	a_2	a_3
2a	53	317	342	314		+ 7.9	−0.9	
18d	24	410	427	429	437	+ 4.1	+4.6	+6.6
19d	42	372	430	379	357	+15.6	+1.9	−3.8
19e	32	395	415	369	361	+ 5.1	−6.6	−8.6
20c	36	289	290	282	278	+ 0.3	−2.5	−3.8

The results are satisfactory. They seem to justify a more thorough investigation along the same lines. The theory however, is not very fundamental and not proof against criticism. The necessary tape measurements are not easily obtained in practice, and the use of optical instruments is not convenient.

In particular the way of sampling is rather complicated: we first take an arbitrary point in the stand, decide which tree is nearest and do our measurements from this tree. We therefore switched over to a different method.

CHAPTER III

A NEW DISTANCE METHOD

1. INTRODUCTION

From a point chosen at random in a forest stand we measure the distance $\underline{a}_n$ to the n^{th} tree. $\underline{a}_n$ is a random variable, i.e. a variable with a probability distribution. If the expected value of $\underline{a}_n$ is large the number of trees per square unit is small (and vice versa). In this chapter we will derive the distribution of $\underline{a}_n$ (or some parameters of the distribution) under various conditions concerning the way the trees are distributed over the area.

First the case is considered where the trees are distributed at random (Random Forest). Next we discuss the case where the trees are in the vertices of a square or equilateral triangular lattice scheme (Systematic Forests). Some parameters (median, first and second moment) are computed for all three cases, and compared. It is found that for the median the mutual differences are minimal. Moreover, there is a tendency for the differences to decrease when a farther tree is concerned.

A theoretical consideration leads to the hypothesis that for forest stands as they are found in practice, estimators of the number of trees assuming Random and Systematic Forests, will yield too high and too low results respectively, and that the differences from the Random Forest will be small if the number of trees is small. We therefore use in practice the estimator of the Random Forest with a small correction that increases with the number of trees estimated. When the mentioned differences are small in comparison to the standard error of the estimator (which is the case when m_4 is used), a rough correction will be enough to reduce the bias, and m_4 can be used to find an allmost unbiased estimator of the number of trees, which is good enough for application in practice. It will be shown that moreover in any case the estimator with the aid of m_4 is more efficient than the others.

2. THE RANDOM FOREST

In order to define a Random Forest, or in general a random set of points in a plane [1]) we proceed as follows:

Let a part of a plane be given and let a set of points in this part be obtained according to some random process. The set of points is called a random set or a POISSON-set (Random Forest) in case the random process obeys the following conditions:

We denote a subdomain of the plane by S, and the size of S by $x = x(S)$. Let the number of points in S be $\underline{k}(S)$. $\underline{k}$ is a random variable.

The assumptions are:

1. The probability $P\{\underline{k} = k\}$ that $\underline{k}$ assumes the value k, does not depend on S, but only on the size x of S. It will be denoted by $p(x, k)$.

2. If the intersection of S_1 and S_2 is void, then $\underline{k}(S_1)$ and $\underline{k}(S_2)$ are stochastically independent. From this fact we conclude:

$$p(x + y, 0) = p(x, 0) \cdot p(y, 0).$$

Take $\ln p(x, 0) = f(x)$, then: $f(x + y) = f(x) + f(y)$; $f(nx) = nf(x)$ and for $\alpha = nm^{-1}$ we find $f(\alpha) = f(nm^{-1}) = nf(m^{-1}) = nm^{-1} f(mm^{-1}) = \alpha f(1)$.

3. $p(x, 0)$ is a continuous function of x. Then $f(x)$ is also continuous.

$$f(x) = xf(1) = \ln p(x, 0) = x \ln p(1,0) \quad \text{for any} \quad x \geqslant 0.$$

Take $\ln p(1,0) = -\alpha$, then $p(x, 0) = e^{-x\alpha}$, and in case $\alpha = 1$.

$$p(x, 0) = e^{-x} \tag{3, 2,1}$$

In order to obtain a formula for $p(x, 1)$, we consider the partition of a domain of size n x in n parts of size x and observe that consequently

$$p(nx, 1) = np(x, 1)p^{n-1}(x, 0) = np(x, 1)e^{-(n-1)x} = np(x, 1)\exp\cdot\{-(n-1)x\}$$
$$p(mx, 1) = mp(x, 1)p^{m-1}(x, 0) = mp(x, 1)\exp\cdot\{-(m-1)x\}$$

Hence: $p(nx, 1) = nm^{-1} p(mx, 1)\exp\cdot\{(m-n)x\}$

Take $mx = 1$ and $nm^{-1} = \alpha$ then $p(\alpha, 1) = \alpha\, p(1, 1)\exp\cdot(1-\alpha)$ or

$$p(x, 1) = C\, x\, e^{-x} \tag{3, 2,2}$$

[1]) The assumptions and the proof are taken from lectures of Prof. Dr. N. H. KUIPER at the University of Agriculture at Wageningen.

We next define $\gamma(x, k)$ by

(3, 2,3) $$p(x, k) = (k!)^{-1}\{C^k x^k + \gamma(x, k)\}e^{-x}$$

and we intend to prove that γ vanishes, as is does in (1) and (2).

We make the inductive assumption:

(3, 2,4) $$\gamma(x, k) = 0 \quad \text{for} \quad k \leqslant N-1 \quad (N \geqslant 2)$$

From (3, 2,3) we find:

(3, 2,5) $$p(x+y, N) = (N!)^{-1}\{C^N(x+y)^N + \gamma(x+y, N)\}\cdot \exp(-x-y)$$

$$p(x+y, N) = \Sigma_{k=0}^{N} p(x, N-k)\, p(y, k)$$

(3, 2,6) $$p(x+y, N) = (N!)^{-1}\{C^N(x+y)^N + \gamma(x, N) + \gamma(y, N)\} \exp(-x-y)$$

From (3, 2, 5) and (3, 2, 6) follows:

(3, 2,7) $$\gamma(x+y, N) = \gamma(x, N) + \gamma(y, N)$$

Take the constant $\gamma(1, N) = t$ then $\gamma(x, N) = x\gamma(1, N) = xt$

From (3, 2, 3) follows:

(3, 2,8) $$p(x, N) = (N!)^{-1}(C^N x^N + tx)\, e^{-x}$$

If $t \neq 0$ then

$$\lim_{x \to 0} \frac{p(2x, N)}{p(x, N)\, p(x, 0)} = 2$$

If we divide a domain of size 2x containing N points in two parts of size x, then there are $N-k$ points in one part in case there are k in the other part. We therefore have

$$p(2x, N) = 2\, p(x, N)\, p(x, 0) + \Sigma_{k=1}^{N-1} p(x, N-k)\, p(x, k)$$

If $t \neq 0$, then (compare 8),

$$\lim_{x \to 0} \frac{p(2x, N)}{2p(x, N)\, p(x, 0)} = 1$$

Hence for 2x very small there is hardly any chance that the N points are not in the same part. The N points are "tied" together. We now give the following assumption which excludes this possibility:

4. Points are not "tied".

With this assumption $t = 0$; $\gamma = 0$ and we get:

$$p(x, k) = \frac{1}{k!} C^k x^k e^{-x}.$$

The sum of the probabilities is:

$$\Sigma_{k=0}^{\infty} p(x,k) = e^{(Cx-x)} = 1; \quad C = 1$$

In the general case $\alpha \neq 1$:

(3, 2,9) $$p(x, k) = \frac{(\alpha x)^k}{k!} e^{-\alpha x} \quad \text{(Poisson distribution)}$$

It is well known that:

$$E(\underline{k}) = \alpha x = \lambda, \quad E(\underline{k}^2) = \lambda_2 + \lambda, \quad \sigma^2 = E(\underline{k}^2) - \{E(\underline{k})\}^2 = \lambda$$

The POISSON distribution with parameter λ can also be obtained as the limit of the binomial distribution $P(\underline{k} = k) = \binom{s}{k} q^k(1-q)^{s-k}$ with parameters s and q, for $q \to 0$; $\lambda = sq = \text{constant}$.

$P(\underline{k}, \text{POISSON } \lambda) = \lim_{q=\lambda s^{-1} \to 0} P(k, \text{binomial } s, q)$ which is proved in many text books.

3. THE PROBABILITY DISTRIBUTION OF $\underline{a}_n$ IN A RANDOM FOREST

In case a number of disjoint domains S_1, S_2, ... all of the same size are given in the plane (= Random Forest) and the numbers of points (trees) in these domains are $k(S_1)$, $k(S_2)$, ... then these numbers can be considered samples from the POISSON distribution:

$$P(\underline{k} = k) = P_k = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, 2, \ldots$$

$\lambda = E(\underline{k})$ is the expectation of $\underline{k}$.

Remarks. 1. In practice distances from an arbitrary point to the centre of a tree are measured. Only the centres of the trees are considered. The discussion remains the same for trees as well as for points.

2. We will consider only circular domains.

Let the distance from P to the (i – th) tree be $\underline{a}_i$. The event $(\underline{a}_i > a)$ only occurs when the circle with centre P and radius "a" contains at most (i – 1) trees.

Let c be the radius of a circle, such that $E(\underline{k}) = 1$. If N is the number of trees per square unit then $\pi c^2 = N^{-1}$.

Instead of $\underline{a}_i$ we now consider the standardized variable

$$\underline{r}_i = \underline{a}_i \cdot c^{-1} = \underline{a}_i(\pi N)^{\frac{1}{2}}$$

$$E(\underline{r}_i) = (\pi N)^{\frac{1}{2}} E(\underline{a}_i)$$

$$N^{\frac{1}{2}} = \pi^{-\frac{1}{2}} E(\underline{r}_i) . \{E(\underline{a}_i)\}^{-1} = C_i \{E(\underline{a}_i)\}^{-1};$$

(3, 3,1) $$C_i = \pi^{-\frac{1}{2}} E(\underline{r}_i)$$

and $$\mathcal{S}(N^{\frac{1}{2}}) = C_i \, \bar{a}_i^{-1}$$

In first approximation we take:

(3, 3,2) $$\mathcal{S}(N) = C_i^2 \, \bar{a}_i^{-2}$$

(3, 3,2) gives a biased estimate of N but it can be shown that this estimate is consistent. This follows from:

$$\mathcal{S}(N^{-\frac{1}{2}}) = \bar{a}_i C_i^{-1} = \bar{a}_i N^{-\frac{1}{2}} \{ E(\bar{a}_i) \}^{-1} = (1 + u) N^{-\frac{1}{2}}; \; u = \{ \bar{a}_i - E(\underline{a}_i) \} \{ E(\underline{a}_i) \}^{-1}$$

$$E\{ \mathcal{S}(N^{-\frac{1}{2}}) \}^{-2} = E\{ (1 + u)N^{-\frac{1}{2}} \}^{-2} = N \cdot E(1 + u)^{-2}$$

$$= N \cdot E(1 - 2u + 3u^2 + \ldots)$$

(3, 3,3) $$E\{ \mathcal{S}(N^{-\frac{1}{2}}) \}^{-2} = N \cdot E(1 + 3v_i^2 n^{-1});$$

$$E(u) = 0;$$

$$E(u^2) = \{ \sigma(\bar{a}_i) \}^2 \{ E(a_i) \}^{-2} = v_i^2 n^{-1}$$

and for $n \to \infty$ we have $E\{ \mathcal{S}(N^{-\frac{1}{2}}) \}^{-2} = N$.

As we know v_i^2 (Table 4), we can calculate the bias.

If $i = 4$ then $v_i = 0.25$. For $n = 100$ we have $3\,v_i^2\,n^{-1} = 0.0018$ and for $n = 60$, $3v_i^2\,n^{-1} = 0.003$. Consequently for practical purposes we can use (3, 3,2) to estimate N.

The event $\underline{r}_n > r$ only occurs when a circle with radius r contains less than n trees.

$$P(\underline{r}_n > r) = \Sigma_{i=0}^{n-1} \; P(k = i \mid \text{a circle with radius } r)$$
$$= \sum{}' \frac{\lambda^i}{i!} e^{-\lambda}; \quad (\Sigma' = \Sigma_{i=0}^{n-1})$$

λ is the expected number of trees in a circle of size πr^2.

A circle of size π has 1 as expected number of trees, a circle of size πr^2 has r^2 as expected number, so $\lambda = r^2$.

It follows: $P(\underline{r}_n > r) = \sum{}' \frac{r^{2i}}{i!} e^{-r^2}$

(3, 3,4) $$H_n(r) \overset{\text{def.}}{=} P(\underline{r}_n > r) = \sum{}' \frac{r^{2i}}{i!} e^{-r^2}$$

$$F_n(r) = P(\underline{r}_n \leqslant r) = 1 - H_n(r)$$

The moments of the distribution.

We first compute:

$$\int_0^\infty r^j\, e^{-r^2}\, dr = \int_0^\infty u^{\frac{1}{2}j}\, e^{-u}\, d\, u^{\frac{1}{2}} = \tfrac{1}{2} \int_0^\infty u^{\frac{1}{2}(j-1)}\, e^{-u}\, d\, u = \tfrac{1}{2}\Gamma\{\tfrac{1}{2}(j+1)\}$$

$$E(\underline{r}_n^t) = \int_0^\infty r^t d\, F_n(r) = -\int_0^\infty r^t d\, H_n(r) = -\, r^t H_n(r) \Big/_0^\infty + \int_0^\infty H_n(r) dr^t$$

$$= t \int_0^\infty r^{t-1}\, H_n(r)\, dr \quad \text{and, substituting (3, 3, 4)}$$

$$= t \int_0^\infty \sum{}' \frac{r^{2i+t-1}}{i!}\, e^{-r^2}\, dr$$

$$= t \sum{}' \frac{1}{i!} \int_0^\infty r^{2i+t-1}\, e^{-r^2}\, dr$$

$$E(\underline{r}_n)^t = \tfrac{1}{2}\, t\, \Sigma_{i=0}^{n-1} \frac{1}{i!}\, \Gamma\,(i + \tfrac{1}{2}t)$$

As $\frac{1}{i!}\,\Gamma\,(i + \frac{1}{2}t) = \binom{\frac{1}{2}t + i - 1}{i}\,\Gamma\,(\frac{1}{2}t)$ we find

$$E(\underline{r}_n^t) = s\left[1 + \binom{s}{1} + \binom{s+1}{2} + \ldots \binom{s+n-2}{n-1}\right]\,\Gamma\,(s), \text{ where } s = \frac{t}{2}$$

Now we make the assumption:

(3, 3,5) $C_{k-1} = 1 + \binom{\frac{1}{2}}{1} + \binom{\frac{1}{2}+1}{2} + \ldots \binom{\frac{1}{2}+k-2}{k-1} = \binom{\frac{1}{2}+k-1}{k-1};$
$k \leqslant n.$

This holds for $n = 1, 2, 3$.

We add to (3, 3,5) for $k = n$, on both sides $\binom{\frac{1}{2}+n-1}{n}$ and find:

$$C_n = \binom{\frac{1}{2}+n-1}{n-1} + \binom{\frac{1}{2}+n-1}{n} = \binom{n+\frac{1}{2}}{n}$$

Hence if (3, 3,5) holds for $k = n$ it also holds for $k = n + 1$. Then it holds for any value of k.

As (3, 3, 5) holds for $n = 1, 2, 3$ it is true for every n.

We will next prove the formula:

(3, 3,6) $k(s, n) = 1 + \binom{s}{1} + \binom{s+1}{2} + \ldots \binom{s+n-2}{n-1} = \binom{s+n-1}{n-1}$

By substitution we observe:

$$k(s, n) = \binom{s+n-1}{n-1} \text{ for } n = 1 \text{ or } 2 \text{ and for any } s.$$

We now prove that (3, 3,6) holds in general, by induction:

Suppose $k(s, i) = \binom{s+i-1}{i-1}$ for $i < n$, then $k(s, n-1) = \binom{s+n-2}{n-2}$,

$k(s, n) = k(s, n-1) + \binom{s+n-2}{n-1} = \binom{s+n-2}{n-2} + \binom{s+n-2}{n-1} = \binom{s+n-1}{n-1}$

Hence:

(3, 3,7) $$E(\underline{r}_n^t) = \frac{t}{2} \cdot \binom{n-1+\frac{1}{2}t}{n-1} \Gamma\left(\frac{t}{2}\right)$$

The median

If $P(r \leqslant m) = F(m) = \frac{1}{2}$, then m is called the median of the distribution F. The median m of $F(r_n)$ is therefore the solution m of the equation:

$$H_n(m) = e^{-m^2} \Sigma_{i=0}^{n-1} \frac{m^{2i}}{i!} = \frac{1}{2}$$

For $n = 1$ we find $m_1^2 = \ln 2$.

In case $n \geqslant 2$ we obtain a numerical solution for m as follows: A first estimate of the root is obtained from a graph. After that, better estimates are found by approximating the function $H_n(m)$ by a few terms of a TAYLOR series. cf KUIPER (1956).

Table 4 gives some moments and the median for $n = 1$ till 4. It also contains the variance σ, the coefficient of variation $v = \sigma(\underline{r}) \{E(\underline{r})\}^{-1}$ and the skewness γ_1.

TABLE 4.

n	$E(\underline{r})$	$E(\underline{r}^2)$	$E(\underline{r}^3)$	m	σ	$\sigma(\underline{r})\{E(\underline{r})\}^{-1}$	γ_1
1	$\frac{1}{2}\pi^{\frac{1}{2}}$	1	$\frac{3}{4}\pi^{\frac{1}{2}}$	0,8325	0,46325	0,5227	0,6311
2	$\frac{3}{4}\pi^{\frac{1}{2}}$	2	$\frac{15}{8}\pi^{\frac{1}{2}}$	1,295	0,48255	0,3630	0,4057
3	$\frac{15}{16}\pi^{\frac{1}{2}}$	3	$\frac{105}{32}\pi^{\frac{1}{2}}$	1,635	0,48871	0,2941	0,3179
4	$\frac{35}{32}\pi^{\frac{1}{2}}$	4	$\frac{315}{64}\pi^{\frac{1}{2}}$	1,916	0,49168	0,2536	0,2692

Size of samples in a Random Forest. An example

Suppose we want to have such an estimate $\underset{\cdot}{S}(N)$ of N that the probability of $|\underset{\cdot}{S}(N) - N|$ being greater than 0.10 is extremely small e.g. 0.05, or

$$P\left\{0.9 \leqslant \frac{\underset{\cdot}{S}(N) - N}{N} \leqslant 1.10\right\} = 0.95 \tag{3, 3,8}$$

We take the number of measurements (q) large and use the approximation by a normal distribution.

Roughly we get:

(3, 3,8) gives: $2\,\sigma\{\underset{\cdot}{S}(N)\}N^{-1} \leqslant 0.10$ or $2\,\sigma\{\ln \underset{\cdot}{S}(N)\} \leqslant 0.10,$

$$2\,\sigma\{(\ln \bar{\underline{a}}_n^2)\} = 4\,\sigma\{(\ln \bar{\underline{a}}_n)\} \leqslant 0.10,$$

$$\sigma(\bar{\underline{a}}_n)\,\{E(\underline{a}_n)\}^{-1} \leqslant 0.025;\;\; q^{-\frac{1}{2}}\,\sigma(\underline{a}_n)\{E(\underline{a}_n)\}^{-1} \leqslant 0.025$$

As we know from table 4: $\sigma(\underline{a}_4)\,\{E(\underline{a}_4)\}^{-1} = 0.25$, we have $q \geqslant 100$.
So we need at least 100 observations.

Table 5, page 16, shows the probability distribution of r_1, r_2, r_3 and r_4 in a Random Forest.

4. COMPARISON OF THE RANDOM FOREST AND A REAL FOREST AS FOUND IN NATURE

First we want to discuss whether the four conditions defining the Random Forest can be considered valid for a real Forest. The conditions I, III and IV seem to be acceptable; Condition II however, states that the probability of occurence of k trees in one interval is independent of the probability of occurence of r trees in another interval. This independence is not strictly valid in a real forest. Take e.g. a very small circle. If we find a tree within this area the probability of occurence of one more tree is small, because there is no room available. In other words there is a minimum distance between two trees.

Consider on the other hand a forest where the position of the trees coincides with the vertices of a square (or triangle) lattice. Such forests will be called Systematic Forests. In these forests a minium distance between two trees also

TABEL 5. Probability distribution of the standardized $\underline{r}_n$.

r	$P(\underline{r}_1 < r)$	$P(\underline{r}_2 < r)$	$P(\underline{r}_3 < r)$	$P(\underline{r}_4 < r)$
0.1	0.009950	0.000050	0.000000	0.000000
0.2	0.039211	0.000779	0.000011	0.000001
0.3	0.086069	0.003815	0.000114	0.000003
0.4	0.147856	0.011513	0.000606	0.000024
0.5	0.221199	0.026499	0.002161	0.000131
0.6	0.312324	0.051161	0.005951	0.000526
0.7	0.387374	0.087187	0.013642	0.001629
0.8	0.472708	0.135241	0.027252	0.004214
0.9	0.555142	0,194807	0.048871	0.009469
1.0	0.632121	0.264242	0.080302	0.018989
1.1	0.701803	0.340985	0.122690	0.034644
1.2	0.763072	0.421896	0.176249	0.058338
1.3	0.815480	0.503641	0.240137	0.091697
1.4	0.859142	0.583060	0.312500	0.135734
1.5	0.894601	0.657453	0.390662	0.190569
1.6	0.922695	0.724794	0.471481	0.255321
1.7	0.944424	0.783809	0.551721	0.328143
1.8	0.960836	0.833945	0.628381	0.408007
1.9	0.972948	0.875290	0.699018	0.486904
2.0	0.981684	0.908420	0.761892	0.566521
2.1	0.987845	0.934241	0.816046	0.642298
2.2	0.992093	0.953823	0.861210	0.711794
2.3	0.994958	0.968286	0.897738	0.773338
2.4	0.996849	0.978699	0.926428	0.826067
2.5	0.998070	0.986008	0.948312	0.869780
2.6	0.998841	0.991006	0.964524	0.904862
2.7	0.999317	0.994338	0.976189	0.932088
2.8		0.996517	0.984408	0.952764
2.9		0.997902	0.990015	0.967918
3.0		0.998770	0.993788	0.978844
3.1		0.999289	0.996195	0.986285
3.2			0.997772	0.991508
3.3			0.998719	0.994844
3.4			0.999286	0.996968
3.5				0.998027
3.6				0.999079

exists. We notice that in case of real forests the distribution of the trees over the area shows some resemblance to the distribution in Systematic Forests.

Consider now a forest that originally is established as a Systematic Forest but in which a part (1–p) of the original number of trees is taken away in a random manner. This frequently occurs in forestry (e.g. thinning). We assume that in an arbitrary real forest the distribution of the trees is more or less the same as in the forest last mentioned. If we take sample plots (of the same size) in such a forest the random variable $\underline{k}$ (number of trees counted in a plot) will have approximately a binomial distribution. If p is small (few trees left), this distribution again can be approximated by the Poisson distribution (Random Forest).

Our assumptions are:

1. Random and Systematic forests may be regarded as limiting cases of the real forests.

2. The lesser trees we have pro square unit in a real forest, the more this forest will resemble a Random forest.

5. SYSTEMATIC FORESTS

a. Square lattice. Consider a forest where the position of the trees coincides with the vertices of a square lattice. Our purpose is to compute the first two moments and the median of $\underline{r}_1$, $\underline{r}_2$, $\underline{r}_3$ and $\underline{r}_4$, defined as before.

Assume the shortest distance between two points is 2. Take a coordinate system with axes parallel to the sides of the squares (2×2). The position of every tree can be indicated by coordinates $(x, y) = (2k, 2l)$ $(k, l = 0, 1, 2 \ldots)$.

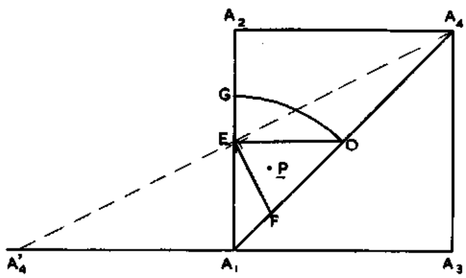


FIG. 1

Take a random point $\underline{P}$ coordinates (x, y) in a square A_1, A_2, A_3, A_4 (fig. 1) with a homogeneous probability distribution, i.e. with a constant probability density. The distance $\underline{a}_i$ from $\underline{P}$ to the i^{th} tree is a random variable. As $\underline{a}_i$ is stochastically independent of the event $\underline{a}_i$ in $\Delta\, A_1OE$ (fig. 1), only distances from $\underline{P}$ in A_1OE will be used. The density then has to be (area Δ AOE$)^{-1} = 2$. In fig. 1, $\underline{a}_1 = \underline{P}A_1$, $\underline{a}_2 = \underline{P}A_2$, $\underline{a}_3 = \underline{P}A_3$, $\underline{a}_4$ under assumption $\underline{P}$ in FEO equals $\underline{P}A_4$, $\underline{a}_4$ under assumption $\underline{P}$ in FEA_1, equals $\underline{P}A'_4 = \underline{a}'_4$.

We introduce the symbols Δ for A_1OE, Δ' for A_1FE.

The coordinates of $\underline{P}$ and A_i are (x, y) and $(-p_i, -q_i)$ respectively. Then:

$$E(\underline{a}_i) = 2 \int_{\Delta} a_i \, dxdy \quad (i = 1, 2, 3)$$

$$E(\underline{a}_4) = 2 \left(\int_{\Delta} a'_4 \, dx\, dy - \int_{\Delta-\Delta'} a'_4 \, dx\, dy + \int_{\Delta-\Delta'} a_4 \, dx\, dy \right)$$

The computation in detail is as follows:

$$E(\underline{a}_1) = 2 \int_{\Delta} a_1 \, dxdy = 2 \int_{A_1OG} r^2 dr d\varphi - \int_{EOG} r^2 dr d\varphi \quad \text{(polar coordinates)}$$

$E(\underline{a}_1) = \frac{1}{3} \cdot 2^{\frac{1}{2}} - \frac{1}{3} \ln \frac{1}{8}\pi = 0.7652$. In case $A_1A_2 = 1$ (one tree per unit) we get: 0.3826.

We obtained the other expected values with a numerical method and found $E(\underline{r}_1) = 0.3826$; $E(\underline{r}_2) = 0.6994$; $E(\underline{r}_3) = 0.9081$; $E(\underline{r}_4) = 1.0226$.

The integrals $E(\underline{a}_i^2) = 2\int a_i^2\, dx\, dy$ are also calculated. We find:

$$\int_{\Delta} a_i^2\, dx\, dy = \tfrac{1}{8} + \tfrac{1}{3}(p + 2q) + \tfrac{1}{2}(p^2 + q^2)$$

$$\int_{\Delta'} a_i^2\, dx\, dy = \tfrac{7}{162} + \tfrac{1}{27}(p + 4q) + \tfrac{1}{6}(p^2 + q^2)$$

In case $A_1A_2 = 1$ we have

$$E(\underline{r}_1^2) = \tfrac{1}{6},\quad E(\underline{r}_2^2) = \tfrac{1}{2},\quad E(\underline{r}_3^2) = \tfrac{5}{6},\quad E(\underline{r}_4^2) = \tfrac{19}{18}.$$

Median.

As the median m_n of the probability distribution of $\underline{a}_n$ is the solution of the equation $P(\underline{a}_n \leqslant m_n) = \tfrac{1}{2}$, we can find the median through a method which will be illustrated in the following example: Take $n = 2$. (fig. 2.)

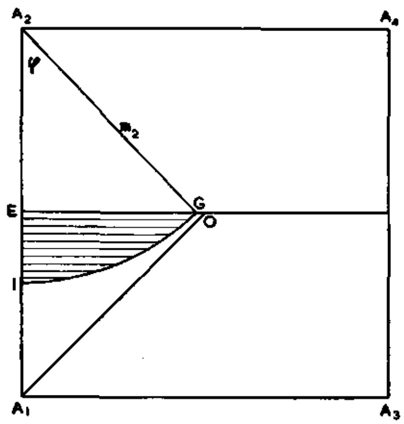


FIG. 2

Consider a circle with centre A_2 and radius m. Take m such that the area of EIG equals one half of the area of EOA_1, then m is the median of the probability distribution of $\underline{a}_2$, for $P(\underline{a}_2 \leqslant m) = P$ (a point occurs in EIG) $= \tfrac{1}{2}$. This yields the equations: $\tfrac{1}{2} m^2\varphi - \tfrac{1}{4} m \sin\varphi = \tfrac{1}{16}$ and $\varphi = \arccos(2m)^{-1}$ from which m can be solved (e.g. by numerical methods). In this way we find the equations:

I. $2\pi m_1^2 - 1 = 0$

II. $8m_2^2\varphi - 4m_2 \sin\varphi - 1 = 0;\ \ \varphi = \arccos(2m_2)^{-1}$

III. $8m_3^2(\varphi - \psi) + 4(m_3^2 - \tfrac{1}{4})^{\frac{1}{2}} + 8m_3 \sin\psi - 7 = 0;\ \ \varphi = \arcsin(2m_3)^{-1}$

$\psi = \arcsin(m_3\sqrt{2})^{-1} - \dfrac{\pi}{4}$

IV. $8m_4^2(\varphi + \psi_1 - \psi_2) - 8(m_4^2 - 1)^{\frac{1}{2}} - 4m_4(2^{\frac{1}{2}} \sin\varphi - 2\sin\psi_2) - 1 = 0$

$\varphi = \dfrac{\pi}{4} - \arcsin(2m_4)^{-1};\ \ \psi_1 = \arccos(m_4)^{-1};\ \ \psi_2 = \dfrac{\pi}{4} - \arcsin(m_4\sqrt{2})^{-1}.$

The solutions are respectively:
$m_1 = 0.3989$; $m_2 = 0.6908$; $m_3 = 0.9153$; $m_4 = 1.0500$.

b. Triangle lattice. We have for a random point in A_1OE (fig. 3.) A_1 = first tree, A_2 the second, A_3 the third and A_4 the fourth. The computations of $E(\underline{a}_i)$, $E(\underline{a}_i^2)$ and the median are analogous to those for the square lattice.

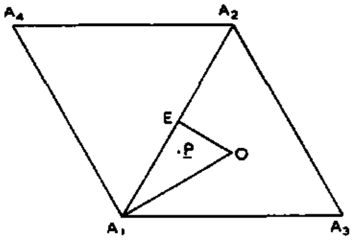


FIG. 3

The results are given in table 6.

TABLE 6.

	1st tree	2nd tree	3rd tree	4th tree
1st moment ($E\underline{r}_i$)				
Square lattice	0.3826	0.6994	0.9081	1.0226
Triangle ,, ...	0.3772	0.7286	0.8540	1.0576
Random	0.5000	0.7500	0.9375	1.0938
2nd moment ($E\underline{r}_i^2$)				
Square lattice	0.1666	0.5000	0.8333	1.0556
Triangle ,, ...	0.1604	0.5453	0.7377	1.1226
Random	0.3183	0.6366	0.9549	1.2732
Median				
Square lattice	0.3989	0.6908	0.9153	1.0500
Triangle ,, ...	0.3989	0.7071	0.8631	1.0548
Random	0.4700	0.7310	0.9230	1.0810
Coefficient of variation				
Square lattice	0.3714	0.1474	0.0959	0.0972
Triangle ,, ...	0.3566	0.1659	0.1072	0.0567
Random	0.523	0.363	0.295	0.253

6. ESTIMATION OF THE STEM NUMBER IN NATURE

Table 6 shows that the medians are more alike than the other parameters. Moreover, it shows that the differences decrease if a farther tree is concerned.

If we take the median of distances to the fourth tree as an estimator, the estimate becomes fairly good for the three types of Forest; so we will use m_4 to estimate N. Estimates with m_4 for Random Forests give a larger N than estimates with m_4 for Systematic Forest.

Suppose we obtained in a real forest in nature an estimate M_1 of the median of the probability distribution of $\underline{a}_1$, and used this parameter to estimate N (compare (3, 3, 2)). If we consider the forest a Random Forest the estimator is (take e.g. $\underline{a}_4$): $\underset{\cdot}{S}(N) = N_r' = 1.081^2 M_4^{-2}$, and if we considered it aS ystematic Forest: $\underset{\cdot}{S}(N) = N_s' = 1.0524^2 M_4^{-2}$. Let N_t be the true stem number, $E(N_r') = N_r$ and $E(N_s') = N_s$. In line with the discussions in section 4 we assume:

$$(3, 6,1) \quad \begin{cases} \text{a.} & N_t = N_r - c(N_r - N_s) \\ \text{b.} & c = (1 - e^{-kN_t}) \approx (1 - e^{-kN_r}) \end{cases}$$

An estimate of N_t is found as $N_t' = \underset{\cdot}{S}(N) = N_r' - c(N_r' - N_s')$.

Remark: As, in case of $\underline{a}_4$ the difference $N_r - N_s$ is small (compare the values of m_4 in table 6), it would be hardly necessary in practice to use such an accurate correction as in (3, 6,1). $\underset{\cdot}{S}(N) = \frac{1}{2}(N_r' + N_s')$ would also suffice for practical purposes.

To test whether our assumption (3, 6,1) holds, we used some maps as in chapter II, and also made measurements in some stands of the Forest Range Oostereng to compute N_r', N_s' and $N_t' = \underset{\cdot}{S}(N)$. If (3, 6,1^a) holds, then:

$0 \leqslant \frac{N_r' - N_t}{N_r'} \leqslant \frac{N_r - N_s}{N_r}$ in most cases. The results in table 6 were used to estimate $\frac{N_r - N_s}{N_r}$. These estimates are approximately 0.3, 0.1, 0.1 and 0.05 in case of $\underline{a}_1$, $\underline{a}_2$, $\underline{a}_3$ and $\underline{a}_4$ respectively.

Table 7 gives the values of $\frac{N_r' - N_t}{N_r'} = C$ for estimates with the median. In most of the stands about 300–400 measurements were taken.

TABLE 7.

Stand		$(N_r' - N_t)(N_r')^{-1}$, for estimates with				N_t (pro. h.a.)	Number of measurements
		m_1	m_2	m_3	m_4		
Esbeek	1	0.24	0.11	0.08	0.01	2317	320
,,	6	0.23	0.10	0.05	0.03	1920	320
,,	1	0.30	0.09	0.04	0.02	1292	360
,,	6	0.26	0.12	0.08	0.06	1226	360
,,	6	0.22	0.10	0.04	0.00	966	360
Oostereng	23g			0.10	0.07	615	199
,,	18f			0.03	0.012	475	170, and 510
,,	13a	0.01	0	0.03	0.003	300	316, 316, 312, and 394
Average		0.21	0.087	0.056	0.026		

We used the estimates with m_1 to test our hypothesis (3, 6,1^b) (the others are not convenient due to the small difference $(N_r - N_s)$ in comparison to the variance of the estimate itself).

We also used the results of m_1 to get a rough estimate of k in (3, 6, 1[b]). We roughly estimated $k = 0.001$. Fig. 4 shows the results.

The points * indicate the true number of trees N_t, • the estimates of N_t and | the confidence interval of the estimate. (All values are given in % of N_r). The distribution of the points * shows the tendency of the correction to decrease with N.

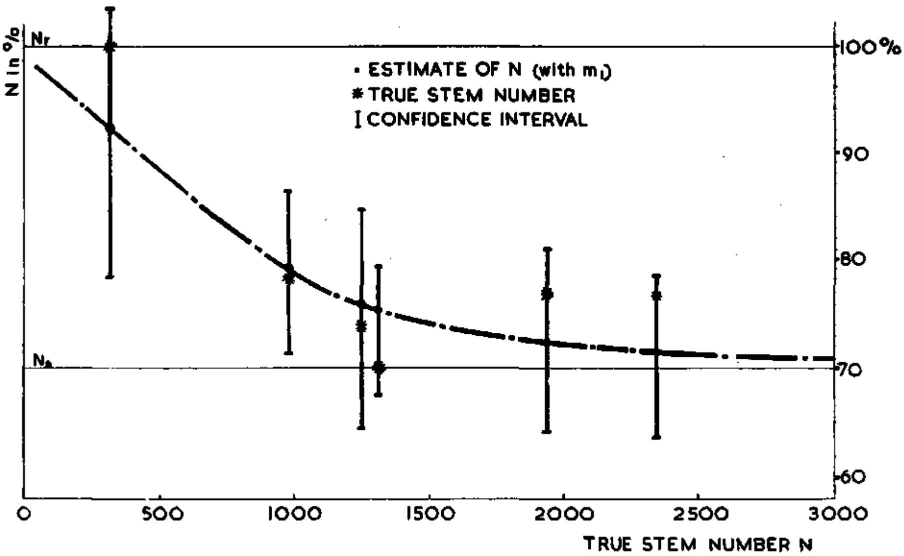


FIG. 4

As stated above confidence intervals were constructed for the stem number estimated by the median. The construction of the confidence interval was as follows:

If our sample is $(a_1 \leqslant a_2 \ldots \leqslant a_n \leqslant a_{n+1} \ldots \leqslant a_{2n-1})$, a_n gives us an estimate of the median m.

$$P(\underline{a} \leqslant m) = P(\underline{a} > m) = \tfrac{1}{2}$$

In this model the probability of i distances being smaller (c.q. larger) than the median, can be calculated.

If $(2n - 1)$ is large we can use the normal distribution approximately. The limiting rank numbers are $i = n - t\sigma$ and $j = n + t\sigma$.

σ is given by $\frac{1}{2}\sqrt{n}$, and t is the value at the 2,5 % point of the normal distribution. In order to get a conservative test we used confidence intervals of 0.90 (Probability of occurence in the critical region is 0.10). The values a_i and a_j, taken as limiting cases of the median, were calculated and used for the computation of stem numbers, which are regarded as the limits of the confidence intervals for the stem number (table 8, page 22).

The occurence of 1 true stem number out of 8, falling outside the confidence interval of 0.90 (estimate with m_4) is no reason to reject our testing hypothesis. (Use the binomial distribution with $p = 0.1$ and $n = 8$ to check this statement).

TABLE 8. Confidence interva (0.90) of the stem number pro ha estimated with the median.

Stand	1st tree	2nd tree	3rd tree	4th tree	true stem number
Esbeek 1	2372	2463	2367	2305*)	2317
„	1916	2186	2144	2108	
„ 1	1398	1402	1280	1307	1226
„	1058	1211	1174	1224	
„ 6	2020	2056	1907*)	1935	1920
„	1599	1858	1757	1832	
„ 6	1460	1357	1285*)	1297	1292
„	1243	1248	1183	1231	
„ 6	1063	1046	971	972	966
„	879	955	893	894	
Oostereng 13a	310	313	324	313	300
„	236	273	281	281	
„			505	493	475
„ 18f			439	447	
„ 23g			690	701	615
„			609	598	

7. SOME NOTES IN CONNECTION WITH THE VARIANCE

Another way to test the assumptions concerning the distribution of trees in a real Forest (section 4) is to compare the variances. In a Systematic Forest the coefficient of variation of the distribution of $\underline{a}_n$ is smaller than in a Random Forest with the same stem number. (Table 9.) In the same stands mentioned above, the variance $s(\underline{a}_n)$ of the distances $\underline{a}_n$ was calculated using: $s(\underline{a}_n) = \frac{1}{N-1} \Sigma (a_n - \bar{a}_n)^2$ (N is the number of distances) and the coefficient of variation $s(a_n) \cdot \bar{a}_n^{-1}$ was calculated. Table 9 shows these values.

TABLE 9. Estimated coefficients of variation of $\underline{a}_n$

Stand \ n	1	2	3	4
Esbeek 1	0.445	0.261	0.206	0.176
„ 1	0.426	0.225	0.166	0.149
„ 6	0.467	0.260	0.212	0.186
„ 6	0.396	0.220	0.187	0.148
„ 6	0.459	0.259	0.202	0.168
Oostereng 13a	0.460	0.255	0.218	0.192
„ 23g			0.251	0.226
„ 18f			0.256	0.236
Square lattice	0.37	0.15	0.10	0.10
Triangle lattice	0.36	0.16	0.11	0.06
Estimated average	0.45	0.25	0.22	0.19
Random	0.52	0.36	0.29	0.25

At the bottom of the table we find a weighted average of the estimated values as well as the expected values in Random and Systematic Forests.

As we might have expected the estimated values are between the expected values in Systematic and Random Forests.

In practice one would like to have an idea of the variance in order to find an estimate of the number of measurements needed to get a "sufficiently accurate" estimate of the stem number.

It is known that the variance of the median is larger than that of the mean. We also saw that the variance of the mean in Real Forests is lower than in Random Forests.

To estimate the variance $\sigma_{(N)}{}^2$ of the estimated stem number (in case the median is used) we used the confidence intervals of N given in table 8, to estimate $\sigma_{(N)}$ dividing the length of the interval by t (approximation with the normal distribution).

In table 10 the values of $S(\sigma_{(N)})N^{-1}$ are given for estimates with m_3 and m_4. (column 2 and 3).

TABLE 10.

Stand		$S(\sigma_{(N)})N^{-1}$ (est. m_3)	$S(\sigma_{(N)})N^{-1}$ (est. m_4)	$(\sigma_{(N)})N^{-1}$ in a Random Forest	
				a_3	a_4
Esbeek	1	0.030	0.026	0.033	0.028
,,	,,	0.026	0.020	0.033	0.028
,,	6	0.024	0.017	0.031	0.027
,,	,,	0.024	0.016	0.031	0.027
,,	,,	0.025	0.025	0.031	0.027
Oostereng	23[g]	0.040	0.050	0.042	0.036
,,	18[f]	0.042	0.029	0.045	0.021
,,	13[a]	0.044	0.033	0.033	0.025
Average		0.032	0.027	0.035	0.027

In column 4 and 5 the expected $\sigma_{(N)}N^{-1}$ is given in case of estimates with $\bar{a}_n$ (using the same number of measurements) in a Random Forest.

In connection with the results in table 10 we assume that for practical purposes we may use the variances of the mean in Random Forests as an estimate of the variance of the median in real forests.

8. EFFICIENCY OF THE DISTANCE METHOD

According to the preceding consideration our estimate of the stem number with the median of the distances to the fourth tree will give sufficiently accurate results for practical use.

In order to test the efficiency of measurements to the fourth tree in comparison to those to the first, second or third tree, a short time study was made.[1])

As we have seen the coefficient of variation decreases if a farther tree is

[1]) We are very greatful for the advices of Ir M. Bol, who assisted us in making the time studies.

considered. On the other hand it becomes more difficult to decide which tree is the n^{th} if n increases. These are two competing factors.

In the forest range Oostereng we chose an arbitrary Scotch pine stand to make the time study.

At the start of my experimental work a tape was used to measure distances. This way of working was rather time consuming. One measurement took on the average $\pm$ 0.75 minutes. The tape proved to be very inconvenient in practice, so we switched to optical methods. The instrument that gave the best results was the telemeter which was attached the BLUME–LEISS-hypsometer; a rod of 0.50 m length was used for the readings.

The error caused by the reading is about 0.02 $\bar{a}_4$ and has little influence on the total coefficient of variation. It only raises this from 0.254 to 0.255 in case of m_4. $((0.254^2 + 0.02^2)^{\frac{1}{2}} = 0.255)$.

Remark: The most proper instrument for the distance method is the range finder, which enables one to measure (standing in a certain point) the distance to a tree as well as its diameter. The instrument can be handled easily by one person. To our regret this instrument was not available during our research.

In a stand in "Oostereng" mentioned above we used the Blume–Leise telemeter with the rod.

The procedure of measuring was as follows. The surveyor chooses an arbitrary spot (A) in the stand, measures the distance from A to the first tree, walks 20 steps, stops at point B, measures the distance to the second tree etc. After having measured the distance to the 4^{th} tree he starts again with another first tree, etc.

This method is preferred to the measuring of all distances to the first, and then all distances to the second etc. in order to eliminate systematic errors (fatigue etc.).

Only the following times are distinguished in the time study: walking time (w) = (time used to walk from one spot to the other); recording time (c_i) = (time used to decide which tree is the i^{th}); reading time (d_i) = (time used to read the distance to the i^{th} tree and to take it down in the notebook).

During the measurement we have in succession w, c, d, w, c, d, etc.

We make the following assumptions.

1. consider w_i, c_i and d_i to be random variables.
2. c_i and d_i are normal deviates with expectation $E(c_i)$ and $E(d_i)$ and variance $\sigma_{(c_i)}$ and $\sigma_{(d_i)}$; $\sigma_{(c_i)} = \sigma_{(c_j)} = \sigma_{(c)}$; $\sigma_{(d_i)} = \sigma_{(d_j)} = \sigma_{(d)}$.
3. $E(c_1) \leqslant E(c_2) \leqslant E(c_3) \leqslant E(c_4)$; $E(d_1) \leqslant E(d_2) \leqslant E(d_3) \leqslant E(d_4)$.

If we want the maximum likelihood estimation of E(c) etc. under these conditions, the problem can be solved along the same line as VAN EEDEN (1955) discussed for ordered probabilities.[1])

In case $\bar{c}_i > \bar{c}_j$ the observations c_i and c_j must be pooled and the maximum

[1]) See also VAN EEDEN (1957).

likelihood estimation is $\bar{c}_i = \bar{c}_j = \dfrac{\Sigma (c_i + c_j)}{2n} = \bar{c}$.

In the same way if $E(i) \leqslant E(j) \leqslant E(k)$ and (e.g.) $\bar{c}_i > \bar{c}_j > \bar{c}_k$ or $\bar{c}_j > \bar{c}_i > \bar{c}_k$ we find for the maximum likelihood estimates

$$\bar{c}_i = \bar{c}_j = \bar{c}_k = \frac{\Sigma (c_i + c_j + c_k)}{3n} = \bar{c}.$$

Our experimental results are given in table 10. Column 2 gives the maximum likelihood estimate for c_i, etc.

TABLE 11.

Distance	$\bar{c}$	$\bar{d}$	$(S(\sigma_c))^2$	$(S(\sigma_d))^2$	n	Likelihood d
a_1	5.436	10.462	0.1509	0.118	39	10.33
a_2	8.475	10.725	0.4295	0.2449	40	10.33
a_3	9.675	9.800	0.4031	0.1003	40	10.33
a_4	15.50	12.725	0.6666	1.0282	40	12.41

In column 3 the three first and the last two estimates must be pooled (results in column 7).

The results show that the reading can be considered to take the same time if the average distance $\bar{a}_n < 4.50$ m (N pro h.a. $= \pm 500$). The walking time remains constant.

If we assume that the total walking time is independent of the distance we measure (this is the case in strip sampling), we find for the total working time (T_i) when measuring to the i^{th} tree.

(3, 8,1) $$T_i = C + n_i \cdot t_i = C + T_i';$$

with C = constant total walking time; n_i = number of measurements; $E(t_i) = E(c_i) + E(d_i)$.

As we have

$n_i = \{\sigma(a_i)\}^2 \{\sigma(\bar{a}_i)\}^{-2} = \{E(a_i)\}^2 \{\sigma(\bar{a}_i)\}^{-2} \{\sigma(a_i)\}^2 \{E(a_i)\}^{-2}$ we find

(3, 8,2) $$T_i' = K_i \{V(\bar{a}_i)\}^{-2};$$

with $K_i = t_i \{\sigma(a_i)\}^2 \{E(a_i)\}^{-2}$; $V(a_i)$ is the coefficient of variation $\sigma(a_i) \{E(a_i)\}^{-1}$ and analogously $V(\bar{a}_i)$.

As we know $V(a_i)$ (Table 6), and as we have an estimate of t_i, an estimate of K_i can be found using (3, 8,2). Results are given in table 12.

TABLE 12.

	$t = (\bar{c} + \bar{d})$		$\sigma^2_{(M)}$	K_i	
	N <500	N >500		N <500	N >500
1	15.77	15.77	0.2735	4.31	4.31
2	18.81	18.81	0.1323	2.51	2.51
3	20.01	20.01	0.0870	1.74	1.74
4	25.83	27.91	0.0640	1.65	1.79

Fig. (5) shows graphs in which we can find the relation between T_i' and $V(\bar{a}_i)$. The graph enables us to compare the efficiency to the different distance measurements.

Remark: If we assume that $E(t_1) : E(t_2) : E(t_3) : E(t_4)$ is independent of the surveyor and the condition of the stand, the results can be used for general conclusions (fig. 5). The more detailed conclusion concerning average distances < 4.50 m does not necessary hold in every case. It depends on special conditions (density).

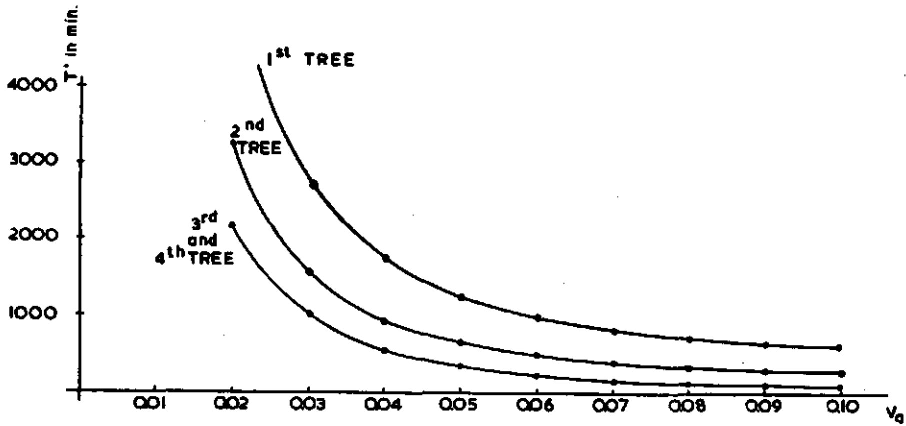


FIG. 5

From table 12 we see that in case of $N \geqslant 500$ the estimates with the distance to the 4th tree are the most efficient (less time for the same accuracy).

For $N < 500$, the estimate with the distance to the third tree is most efficient, although the difference is very small.

In the preceding discussions, the walking time is considered to be constant. If this is not the case (no strip sampling) there might be a tendency for the total walking time to decrease when a farther tree is considered.

In these cases there will be much more reasons to assume that estimates with m_4 are more efficient than those with m_1, m_2 and m_3.

Moreover, we only take the distance measurements into consideration.

As a rule the distance measurements are acompanied by diameter measurements (calliper or tree fork).

Although we do not discuss the diameter measurements thoroughly in this chapter, we mention the next investigation here to complete the discussion.

An additional time study was made to compare the effect of diameter measurements with calliper and tree fork when the distance method is applied.

The results are given below for measurements to the 3rd tree.

TABLE 13. *Total time per observation* (w + c + d)

	Distance without diameter	Distance with one diameter	Distance with two diameters
Calliper	38.98	43.62	46.52
Treefork	38.98	39.02	39.00

Conclusions:

1. The tree fork seems more efficient than the calliper.
2. When the calliper is used there is a significant effect (t-test with values of σ from table 11) of the diameter measurements.
3. There is no effect if the tree fork is used, even when we measure two diameters for each distance.

This result is due to the fact that, when the fork is used, two men can work independently and one need not wait for the measurement made by the other.

As it is shown that the estimate of the stem number with the median (m_4) is the most efficient we constructed table 14 for practical use. Additionally table 15 is given for rough approximations of N when m_3 is used. This can only be recommended in case $N \ll 500$, as was shown.

TABLE 14. Expected stem number (N pro ha) in case m_4 is estimated

m_4 in meters	2	3	4	5	6	7	8
.00	2784	1252	712	459	320	236	181
.05	2651	1212	695	450	315	233	179
.10	2528	1174	678	441	310	230	177
.15	2413	1137	662	433	305	226	175
.20	2305	1103	647	425	300	223	173
.25	2205	1070	632	417	295	220	170
.30	2121	1038	617	409	291	217	168
.35	2024	1008	604	402	286	214	166
.40	1942	979	590	394	282	211	164
.45	1864	951	577	387	278	209	162
.50	1791	925	565	380	273	206	161
.55	1723	900	553	374	269	203	159
.60	1658	875	541	367	265	201	157
.65	1597	852	529	361	261	198	155
.70	1539	829	518	354	257	195	153
.75	1485	808	508	348	254	193	152
.80	1433	787	497	342	250	190	150
.85	1384	767	487	337	246	188	148
.90	1338	748	478	331	243	186	147
.95	1284	729	468	326	239	183	145

TABLE 15. Expected stem number (N pro ha) in case m_s is estimated (only to be used for N < 500)

m_s in meters:	1	2	3	4	5	6	7
.00	–	1940	888	510	331	231	171
.05	–	1849	860	498	324	228	169
.10	–	1765	834	468	318	224	166
.15	–	1686	808	475	312	220	164
.20	–	1613	784	464	306	217	162
.25	–	1544	761	453	301	214	159
.30	–	1480	739	443	295	210	157
.35	–	1420	718	433	290	207	155
.40	–	1363	698	424	284	204	153
.45	–	1310	678	415	279	201	151
.50	3413	1260	660	406	274	198	149
.55	3198	1213	642	397	270	195	147
.60	3004	1169	625	389	265	192	145
.65	2827	1127	609	381	260	189	–
.70	2666	1087	593	373	256	186	–
.75	2518	1049	578	365	252	184	–
.80	2383	1014	563	358	247	181	–
.85	2258	980	549	351	243	178	–
.90	2144	948	535	344	239	176	–
.95	2038	917	522	337	235	173	–

9. APPLICATION IN PRACTICE

The distance method has been applied in 17 Scotch pine stands in the forest Ranges Oostereng, Ommen and Haarle. These 17 stands can be taken as a representative sample of the Dutch Forest stands with respect to the distance method.

The surveys were made by several surveyors.

An example of the recording and computing is given below.

Distance method $\underline{a}_4$
Forest Range: Haarle division 77a
Surveyor: G. H. RAETS, Agr. Cand.

distance in meters	1.50	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	3.75	4.00	
frequency. . . .	1	2	4	5	14	11	4	8	1	2	1	Total 53

Estimated median: 2.635 m
Estimated stem number: 1615

The value of the median is found by interpolation at $26\frac{1}{2}$. Class 2.75 (2.625–2.875) contains Nrs. 26 till 37.

One has at$26\frac{1}{2}$:

$$\left(2.625 + \frac{0.25}{22}\right) = 2.635.$$

The results of all observations are given in table 16.

TABLE 16.

Division	Forest Range	Area in ha	Surveyor	Numbers of measurements	m_3 in m.	m_4 in m.	Estimated stem number $\hat{S}(N)$ pro ha	True stem number N	$\frac{\hat{S}(N) - N}{N}$ m_3	$\frac{\hat{S}(N) - N}{N}$ m_4
18	Oostereng	1.60	Den Hollander	48	4.02		505	470	1.0745	
18	,,	1.60	,,	50	3.93		525	470	1.1170	
18	,,	1.60	,,	50		4.925	473	470		1.0064
18	,,	1.60	,,	50		4.915	475	470		1.0106
9a	Ommen	5.53	Raets-Thijssen	107	5.61		264	266	0.9925	
9a	,,	5.53	,,	100	5.54		271	266	1.0188	
9a	,,	5.53	,,	66	5.54		271	266	1.0188	
11cO	,,	1.93	,,	85	3.85		549	598	0.9181	
11cO	,,	1.93	Essed	60	3.685		582	598	0.9732	
11cW	,,	1.92	,,	95	3.95		522	545	0.9578	
11cW	,,	1.92	Raets	64	4.05		498	545	0.9138	
12aO	,,	2.11	Essed	52	3.74		581	538	1.0799	
12aO	,,	2.11	Thijssen	83	3.86		546	538	1.0149	
13aW	,,	2.02	Essed	54	4.45		415	467	0.8887	
13aW	,,	2.02	Raets	77	4.05		498	467	1.0664	
13b	,,	2.21	Essed	41	7.19		162	148	1.0946	
13b	,,	2.21	Thijssen	77	7.69		142	148	0.9594	
31b	,,	0.594	Raets	35	2.00		1940	2076	0.9345	
31b	,,	0.594	Thijssen	35	1.90		2144	2076	1.0328	
31b	,,	0.594	Essed	44		2.45	1864	2076		0.8979
31b	,,	0.594	,,	40		2.30	2121	2076		1.0217
31b	,,	0.594	Raets	51		2.30	2121	2076		1.0217
40d	,,	0.605	Essed	38	3.57		635	702	0.9046	
40d	,,	0.605	,,	43		3.90	748	702		1.0655
40d	,,	0.605	Thijssen	32	3.31		735	702	1.0407	
40d	,,	0.605	Raets	53	3.625		660	702	0.9402	
60a	,,	1.003	Thijssen	33	2.39		1372	1575	0.8711	
60a	,,	1.003	Essed	46	2.225		1579	1575	1.0025	
60a	,,	1.003	Raets	35	2.40		1363	1575	0.8654	
60a	,,	1.003	Thijssen	38		2.65	1597	1575		1.0140
77a	Haarle	0.64	Raets	53	2.22		1599	1596	1.0019	
77a	,,	0.64	,,	53		2.635	1615	1596		1.0119
78d	,,	0.72	,,	48	5.21		305	292	1.0445	
78d	,,	0.72	,,	48		6.125	302	292		1.0342
97a	,,	1.03	,,	65	3.29		743	741	1.0027	
97a	,,	1.03	,,	65		3.84	771	741		1.0405
99b	,,	0.76	,,	59	2.25		1544	1559	0.9904	
99b	,,	0.76	,,	59		2.72	1515	1559		0.9718
99c	,,	0.98	,,	64	2.53		1232	1038	1.1868	
99c	,,	0.98	,,	67		3.11	1167	1038		1.1243
132a	,,	2.00	,,	63	2.64		1135	1140	0.9956	
132a	,,	2.00	,,	63		3.05	1212	1140		1.0632
135a	,,	2.00	,,	63	2.72		1072	1156	0.9273	
135a	,,	2.00	,,	63		3.155	1133	1156		0.9801
					sum				29.8284	14.26380
					mean				0.9943	1.0188
					Total mean				1.0021	

Consider the numbers $\underline{x}_i = \{\underset{.}{S}(N) - N\}N^{-1}$. To test the 0-hypothesis $\underset{.}{S}(N) = N$ we can make the following assumptions:

The numbers $\underline{x}_i$ are random variables with an expectation $E(\underline{x}_i)$ and a normal distribution. If we compute $\underline{x}_i$ we may test the hypothesis $E(\underline{x}_i) = 0$ (that means $\{\underset{.}{S}(N) - N\}N^{-1} = 0$ or $\underset{.}{S}(N) = N$).

The standard error of $\underline{x}_i$ can be estimated as $\sigma(\bar{x}_i) = 2\ V_i n^{-1}$. (n is the total number of measurements and V_i is the coefficient of variation of the distribution of $\underline{a}_i$; see pg 23). We find the following results:

$\bar{x}_3 = 0.0057;\ \sigma(\bar{x}_3) = 2(0.295 \times 1780^{-1}) = 0.0132;\ \ t_3 = \bar{x}_3\ \{\sigma(\bar{x}_3)\}^{-1} = 0.432$

$\bar{x}_4 = 0.0188;\ \sigma(\bar{x}_4) = 2(0.253 \times\ 734^{-1}) = 0.0185;\ \ t_4 = \bar{x}_4\ \{\sigma(\bar{x}_4)\}^{-1} = 1{,}02.$

As the levels of significance for these values of t are 0.333 and 0.154 respectively, we have no reasons to reject the hypothesis $\underset{.}{S}(N) = N$.

We also estimated the total number of trees in all the cases considered. This is done by multiplication of the true and estimated stem numbers by the area of the stand. Table 17 gives a summary of the results. The average of the absolute values of the deviations as well as the maximum deviations are also given for both cases.

TABLE 17.

Estimator	Area in ha.	Total number of trees	Estimated Total Number	Number of measurements	deviation in %	average deviation in %	maximum deviation in %	average number of measurements. (size of sample)
m_3	54.32	32400	31933	1780	0.0144	0.059	0.187	59
m_4	14.72	15999	16228	734	0.0143	0.043	0.124	52

It was also tried to replace the optical instruments by estimates by eye only. The surveyor Mr. RAETS estimated stem numbers in 6 stands in Ommen in that way. He got an average difference of 0.05 N, but in one case a difference of 0.20 N. It seems possible for well trained persons to achieve fairly good estimates in this way.

CHAPTER IV

COMPUTATION OF THE MEAN VOLUME

1. SOME WELL KNOWN METHODS

During the development of forest mensuration much attention was paid to the dimensions of the so called "mean volume tree", i.e. a hypothetical tree that has a volume equal to the mean volume of all the trees of the stand. Especially in Germany different authors made investigations regarding this subject. As can be expected the form factor was used in all these investigations.

Ever since the beginning of forest mensuration the tree with $\bar{g}$ (the average basal area) was used. Although SPEIDEL (1893) showed that a systematic error is made if the mean basal area tree is used, it is still a common practice to use this tree as the mean volume tree.

Some methods were introduced in which use is made of the diameters of all trees and a height curve to determine the mean volume, with a volume table. The diameter and height of the mean volume tree were found by interpolation in the table. (TISCHENDORF, NEUBAUER.)

HOHENADL suggested the use of two trees with diameters $\bar{d} - s$ and $\bar{d} + s$. The average volume of these two trees was considered equal to the mean volume of the stand. It was shown that this statement holds if the following relation exists in the stand:

$$v = ad^2 + bd + c$$

v is the expected volume of a tree with diameter d.

The discussions in literature showed that the authors were aware that $d_{\bar{g}} \neq d_{\bar{v}}$, but they still used $d_{\bar{g}}$ because there was no easy method known to estimate $d_{\bar{v}}$.

2. BERKHOUT'S RELATION

The author started an investigation concerning the computation of the diameter of $d_{\bar{v}}$ in a stand. He took as point of departure the well proved formula of BERKHOUT which gives a relation between v and d.

BERKHOUT (1920) found the formula $v = ad^b$, (v is the expected volume of the tree with diameter d in even aged stands. b is a constant for the wood species and a is only constant in the stand. BERKHOUT worked with Scotch pine and found $v = ad^{2,2}$. STOFFELS introduced for the constant a the following formula: $a = p\bar{d}^q\bar{h}^r$. He stated for Scotch pine: $a = 0.501\,\bar{d}^{-0.268}\,\bar{h}^{0.865}$ and $b = 2.2$ as BERKHOUT. He used these two functions to construct standard volume tables.

The author computed the constant b from 33 surveys of 21 different Douglas fir plantations and 23 surveys of 10 different Japanese larch (*Larix leptolepus*) plantations resp. and found for:

$$\text{Larch: } b = 2.393 \pm 0.012 \text{ and Douglas fir: } b = 2.394 \pm 0.014$$

For practical purposes $b = 2.4$ can be used for both species. The constant a was also calculated for Douglas fir, from 60 surveys in 24 stands using the formula STOFFELS suggested. We found

$$a = 0.0597\,\bar{d}^{-0.54}\,h_{\bar{d}}^{0.978} \qquad (4, 2,1)$$

3. COMPUTATION OF THE MEAN VOLUME $\bar{v}$

We start from BERKHOUT's formula, which states that the expected volume E(v) of a tree with diameter d is given by the formula $E(v) = ad^b$.

The following symbols are used.

$\bar{d}$ = arithmetic mean diameter
$d_{\bar{g}}$ = diameter of mean basal area tree
$d_{\bar{v}}$ = diameter of mean volume tree
d = diameter of an arbitrary tree
$u = d - \bar{d}$
v = expected volume of an arbitrary tree

$\bar{v}$ = mean volume
$v_{\bar{d}}$ = expected volume of the tree with diameter $\bar{d}$
V = total volume
a en b = constants
n = number of trees
$s^2 = n^{-1} \Sigma u^2$ = variance of the frequency distribution of the diameters

The volume of an arbitrary tree can be expressed as follows: $v = a(\bar{d} + u)^b$. Using TAYLOR's series for $(\bar{d} + u)^b$ we find:

$$v = a(\bar{d} + u)^b = a(\bar{d}^b + b\,u\,\bar{d}^{b-1} + \tfrac{1}{2}b\,(b-1)u^2\,\bar{d}^{b-2} + R$$

After summation we find for the total volume

$$\Sigma\, v = a\,(n\bar{d}^b + b\,\Sigma\, u\bar{d}^{b-1} + \tfrac{1}{2}\, b(b-1)\,\Sigma\, u^2\bar{d}^{b-2} + R')$$

and the mean volume is:

$$\bar{v} = a\bar{d}^b\,(1 + n^{-1}\,\bar{d}^{-1}\,\Sigma\, u + n^{-1}\,\bar{d}^{-2}\,\Sigma\, u^2 + R'')$$

Furthermore we have:

1. $\Sigma\, u = 0$. This follows from $u = (d - \bar{d})$.
2. In our cases b and $s^2\,\bar{d}^{-2}$ are approximately 2, 3 and 0.06 and we can show that $|R''| < 0{,}0006$ so that R'' can be omitted for practical purposes. We find:

(4, 3,1) $$\bar{v} = a\bar{d}^b\,\{1 + \tfrac{1}{2}\, b(b-1)\cdot s^2\,\bar{d}^{-2}\}$$

(4, 3,2) $$\bar{v} = v_{\bar{d}}\cdot c \text{ with } c = \{1 + \tfrac{1}{2}\, b(b-1)\cdot s^2\,\bar{d}^{-2}\}.$$

In other words: The mean volume can be found from the volume of the tree with the arithmetic mean diameter, by multiplication with a factor that only depends on the value of b and the coefficient of variation = $s\bar{d}^{-1}$.

The computation of the correction factor c can be facilitated by the construction of a table with the form:

Douglas fir and Larch values of c

$\bar{d}$ \ s	2	2.5	3
10	1.0672	1.1050	1.1512

Scotch pine: values of c

$\bar{d}$ \ s	2	2.5	3
10	1.0528	1.0825	1.1188

4. COMPUTATION OF THE DIAMETER OF THE MEAN VOLUME TREE

As we can easily compute $\bar{d}$ we shall express $d_{\bar{v}}$ as a function of $\bar{d}$. For this purpose we use (4, 3,1)

$$a d_{\bar{v}}^{\,b} = a\,\bar{d}^{\,b}\{1 + \tfrac{1}{2}\, b(b-1)\cdot s^2\,\bar{d}^{-2}\}$$

$$d_{\bar{v}} = \bar{d}\,\{1 + \tfrac{1}{2}\, b(b-1)\, s^2\,\bar{d}^{-2}\}^{b^{-1}}$$

Using TAYLOR's series and omitting the terms of the third and higher degree we find:

(4, 3,3) $$d_{\bar{v}} = \bar{d}\,\{1 + \tfrac{1}{2}(b-1)\cdot s^2\,\bar{d}^{-2}\}$$

(4, 3,4) $$d_{\bar{v}} = \bar{d}\cdot c' \text{ with } c' = 1 + \tfrac{1}{2}(b-1)\,s^2\,\bar{d}^{-2}$$

The formula shows that the diameter of the mean volume tree can be found from the arithmetic mean diameter, the constant b and the coefficient of variation. As in (4, 3,2) we see that the factor a does not occur in the formula.

5. SOME DATA FOR PRACTICAL USE.

Using the values b we mentioned in section 1 we find, taking $s^2\,\bar{d}^{-2} = q^2$ and using formula's (4, 3,2) and (4, 3,4) for Scotch pine:

$$c = 1 + 1.32\,q^2 \quad \text{and} \quad c' = 1 + 0.6\,q^2$$

for Douglas fir and Larch:

$$c = 1 + 1.68\,q^2 \quad c' = 1 + 0.7\,q^2$$

To find the mean volume we can better use volume tables as the measuring of the volume of model trees in the stand is usually too laborious and does not guarantee better results.

The estimate of s^2 can be found with the formula $n^{-1}\,\Sigma\,u^2$. This requires some computations, but not more than in the case of $d_{\bar{g}}$.

In practice $2s = d_{84\%} - d_{16\%}$ is used, assuming a normal distribution ($d_{i\%}$ is the value of d at the i % point of the frequency distribution of the diameters). As we know that the frequency distribution of the diameters has a small positive skewness, we tried to find a simple expression by the aid of which s can be found by counting in a similar way.

For this purpose we computed s with the formula $s = (\Sigma\,u^2)^{\frac{1}{2}}\,n^{-\frac{1}{2}}$ in 65 Douglas fir stands and 36 Scotch pine stands. We assumed $b_i = d_{(100-i)\%} - d_{(i\%)} = k_i s$ and computed b_{15}, b_{16} and b_{17} using the regression formula

$$k_i = (\Sigma\,b_i\,s)\,(\Sigma\,s^2)^{-1}$$

The results were for Douglas fir as well as for Scotch pine: $b_{15} = 2.11$ s; $b_{16} = 2.03$ s; $b_{17} = 1.95$ s. By interpolation we found $2s = b_{(16.3)}$. We suggest to use $b_{15} = 2{,}1$ s for practical purposes.

To find an estimate of $\bar{d}$ and s we should take a sample of the diameters, either by callipering or by using the tree-fork. We first like to have an idea about the number of trees in our sample in order to get a certain accuracy.

It is known that if n is the size of the sample in a finite population of N individuals, we find an estimate of the variance of the sample mean by $(s_{\bar{d}})^2 = s^2\,n^{-1}\,(1 - nN^{-1})$.

Suppose we want an estimate $\bar{D}$ of $\bar{d}$ in such a way that $P\{(\bar{D} - \bar{d}) > 0.05\bar{d}\} = 0.95$. In that case we have, assuming normal distribution, $2\sigma_{\bar{d}} = 0.05\,\bar{d}$ or $\sigma_{\bar{d}} = 0.025\,\bar{d}$. If we take $s_{\bar{d}} = \sigma_{\bar{d}}$ we have $(s_{\bar{d}})^2\,\bar{d}^{-2} = \frac{1}{1600}$ or

$$s^2\,\bar{d}^{-2}\,(n^{-1} - N^{-1}) = \frac{1}{1600}$$

In case $s^2 \bar{d}^{-2} = \frac{1}{16}$ we have $n = 100\,(1-nN^{-1})$, so in practice it would hardly be necessary to measure more than 100 diameters at random. An error occurs also, since we usually take the diameters in classes. This error can be computed, assuming that every value in the class has the same probability of occurrence. We find a rectangular distribution in the class interval. The variance is $\frac{1}{12}\,a^2$ if a is the length of the class. The variance of the mean value is $\frac{1}{12}\,a^2\,n^{-1}$ if n is the size of the sample.

In practice the following method can be applied.

1. Take a sample of 100 diameters to estimate $\bar{d}$ and b_{15}. If N is small e.g. < 500, compute the size n approximately using $n = 100\,(1-nN^{-1})$.
2. Measure $h_{\bar{d}}$ and find $v_{\bar{d}}$ in a volume table using $\bar{d}$ and $h_{\bar{d}}$.
3. Compute $\bar{v}$ with (4, 3,2) or use a table for it.

It is also possible to use (4, 3,4) and compute $d_{\bar{v}}$. After measuring $h_{\bar{v}}$ the volume $\bar{v}$ can be found (volume table).

6. SOME APPLICATIONS

First we use the example of STOFFELS (1953) concerning a Scotch pine stand.

We have, $\bar{d} = 13.59$ $s^2 = 5.878$. As $b = 2.20$, for Scotch pine we have:

$$c' = 1 + 0.6\,\frac{5.878}{184.59} = 1.0191$$

$$d_{\bar{v}} = c' \times \bar{d} = 1.0191 \times 13.59 \text{ cm} = 13.85 \text{ cm}$$

$$\bar{v} = 0.1008 \text{ m}^3$$

STOFFELS gives: $\frac{61.79 \text{ m}^3}{614} = 0.1007 \text{ m}^3$

HOHENADL's method gives:

$$\bar{d} - s = 13.59 \text{ cm} - 2.44 \text{ cm} = 11.15 \text{ cm}$$

$$\bar{d} + s = 13.59 \text{ cm} + 2.44 \text{ cm} = 16.03 \text{ cm}$$

$$v_{(\bar{d}-s)} = 0.0630 \text{ m}^3$$

$$v_{(\bar{d}+s)} = 0.1386 \text{ m}^3$$

$$\bar{v} = \tfrac{1}{2}\,(0.0630 + 0.1386)\text{m}^3 = 0.1008 \text{ m}^3$$

For a second example we chose the 41 years old Douglas fir stand SS8 of the Forestry Research Institute of the University of Wageningen. The stand lies in division 30^d of the Forest Range "Speulder and Sprielderbos".

The volume was computed by a complete callipering. The diameter of each tree was callipered in two directions. A height curve was constructed from a great number of height measurements. The volumes given are calculated with BECKING's Volume table for Douglas fir. The total volume is 90.162 m^3 and the mean 0.8669 m^3. The results are given in table (18).

In the same table the calculation with our method is given.

First the arithmetic mean diameter $\bar{d} = 30.45$ cm was computed. $h_{\bar{d}} = 25.1$ m. We found $v_{\bar{d}} = 0.811 \text{ m}^3$.

We computed s using $s^2 = n^{-1}\,\Sigma\,u^2$ and found $s = 6.37$ cm.

Using (4, 3,2) we found c = 1.074 and $\bar{v} = 1.074 \times 0.811\ m^3 = 0.8710\ m^3$.

A different method gives (4, 3,4) $d_{\bar{v}} = 31.7$ cm, and consequently $h_{\bar{v}}$ 25.4 m and $\bar{v} = 0.8704\ m^3$.

TABLE 18. Douglas stand SS8, dev. 30d, Speulder en Sprielderbos, Apeldoorn. Area 0.2736 ha (41 years old)

Old method					New method				
$d_{1.30}$ in cm	h in m	n	v_i in m^3	nv_i in m^3	$d_{1.30}$ in cm	n	u	n.u	$n.u^2$
15	20.0	1	0.160	0.160	15	1	−15	−15	225
19	21.3	3	0.271	0.813	19	3	−11	−33	363
20	21.6	1	0.304	0.304	20	1	−10	−10	100
21	22.0	2	0.340	0.680	21	2	− 9	−18	162
22	22.3	1	0.380	0.380	22	1	− 8	− 8	64
23	22.6	4	0.419	1.676	23	4	− 7	−28	196
24	23.0	6	0.464	2.784	24	6	− 6	−36	216
25	23.3	9	0.510	4.590	25	9	− 5	−45	225
26	23.6	8	0.548	4.384	26	8	− 4	−32	128
27	24.0	6	0.611	3.666	27	6	− 3	−18	54
28	24.3	5	0.665	3.325	28	5	− 2	−10	20
29	24.6	4	0.722	2.888	29	4	− 1	− 4	4
30	25.0	3	0.784	2.352	30	3	0	0	0
31	25.3	3	0.845	2.535	31	3	+ 1	+ 3	3
32	25.6	7	0.910	6.370	32	7	+ 2	+14	28
33	26.0	4	0.981	3.924	33	4	+ 3	+12	36
34	26.3	7	1.050	7.350	34	7	+ 4	+28	112
35	26.6	7	1.128	7.896	35	7	+ 5	+35	175
36	26.9	4	1.207	4.828	36	4	+ 6	+24	144
37	27.2	2	1.288	2.576	37	2	+ 7	+14	98
38	27.5	5	1.372	6.860	38	5	+ 8	+40	320
39	27.8	4	1.460	5.840	39	4	+ 9	+36	324
40	28.2	4	1.547	6.188	40	4	+10	+40	400
44	28.5	3	1.900	5.700	44	3	+14	+42	588
46	28.8	1	2.093	2.093	46	1	+16	+16	256

N = 104 $V = 90.162\ m^3$

$\bar{v} = 0.8669\ m^3$

$\Sigma n.u. = +47 \quad \Sigma n.u.^2 = 4241$

$\bar{d} = 30 + \frac{47}{104} = 30.45$ cm

$h_{\bar{d}} = 25.1$ m

$c' = \left(1 + 0.7 \times \frac{40.58}{927.20}\right) = 1.0306$

$d_{\bar{v}} = 30.45 \times 1.0306\ \text{cm} = 31.4\ \text{cm}$

$h_{\bar{v}} = 25.4$ m

$\bar{v} = 0.8704\ m^3$

deviation: + 0.40 %

$s = \left(\frac{4241}{104} - 0.45^2\right)$ cm = 6.37 cm

$v_{\bar{d}} = 0.811\ m^3$

$c = \left(1 + 1.68 \times \frac{40.58}{927.20}\right) = 1.074$

$\bar{v} = 0.811 \times 1.074\ m^3 = 0.8710\ m^3$

deviation: + 0.47 %

Applying HOHENADL's method we find $\bar{v} = 0.869$ (deviation: 0.24 %).

We also tested the method, computing $\bar{v}$ for about 20 Douglas fir plantations. All differences from the complete measurements were less than 1 %.

An example can show that a great deal of the deviation is due to the roundings in the table. Let a tree have $d = 31.35$ cm; $h = 25.35$ m; its volume is 0.8662 m^3. If we had rounded to $d = 31.4$ cm; $h = 25.4$ m we should have 0.8704 m^3. The deviation is: 0.48 %.

7. DIAMETER AND VOLUME OF THE MEAN BASAL AREA TREE

As it is common in forestry practice to assume $d_{\bar{g}} = d_{\bar{v}}$ and consequently $v_{\bar{g}} = \bar{v}$, we want compare these two quantities.

From $d_{\bar{g}}^2 = n^{-1}\Sigma d^2$; $d = \bar{d} + u$ and $s^2 = n^{-1}\Sigma u^2$ we know: $d_{\bar{g}}^2 = \bar{d}^2 + s^2$.

Using TAYLORS series we find with a sufficient approximation

$$d_{\bar{g}} = (\bar{d}^2 + s^2)^{\frac{1}{2}} = \bar{d}(1 + \tfrac{1}{2}\, s^2\, \bar{d}^{-2})$$

For $v_{\bar{g}}$ we find: $v_{\bar{g}} = a\, d_{\bar{g}}^{\,b} = a\, \bar{d}^b\, (1 + \tfrac{1}{2}\, s^2\, d^{-2} ...)^b$.

(4, 7,1) $$v_{\bar{g}} = a\bar{d}^b\, (1 + \tfrac{1}{2}\, b . s^2\, \bar{d}^{-2} ...)$$

From (4, 3,1) and (4, 7, 1) we find for:

$$\bar{v} - v_{\bar{g}} = a\bar{d}^b\, \tfrac{1}{2}\, \{\, b(b-1) - b\, \}\, s^2\, \bar{d}^{-2}$$

(4, 7,2) $$\bar{v} - v_{\bar{g}} = a\bar{d}^b\, \tfrac{1}{2}\, b(b-2) \cdot s^2\, \bar{d}^{-2}$$

(4, 7,3) $$\bar{v} - v_{\bar{g}} = v_{\bar{d}} \cdot c'' \text{ with } c'' = \tfrac{1}{2}\, b(b-2) s^2\, \bar{d}^{-2}$$

From (4, 7,3) we read:

1. $\bar{v} > v_{\bar{g}}$, a conclusion confirmed by many experiments. SPEIDEL e.g. found that in 53 of 55 stands he inspected, the diameter of the mean volume tree was 2–5 mm higher than $d_{\bar{g}}$.
2. $\bar{v} - v_{\bar{g}}$ depends on b and the coefficient of variation.

An estimate of $\bar{v} - v_{\bar{g}}$ is found by taking $q = s\bar{d}^{-1} = \frac{1}{4}$. For Douglas fir and Larch $c'' = 0.03$ and for Scotch pine 0.014.

In our examples, the Scotch pine stand had $\bar{v} = 0.1008$ m^3 and $v_{\bar{g}} = 0.1000$ m^3, (deviation: –0.8 %) and the Douglas fir $\bar{v} = 0.8669$ m^3 and $v_{\bar{g}} = 0.8478$ m^3 (deviation: –2.2 %).

The formulas also show that when there is a linear relation between v and d^2, $(b = 2)$ $\bar{v} = v_{\bar{g}}$.

The formulas allow us to explain the views of some authors.

We cite:

TISCHENDORF: "... allerdings bezieht sich GEHRHARDT nur "auf gleichmäszige, reine und gleichaltrige Fichtenbestände." and "In Beständen wie sie GEHRHARDTS Beispiel zugrunde liegen, wird der Unterschied, practisch überhaupt nicht in Erwägung gezogen werden".

GEHRHARDT: "Die theoretische Verschiedenheit des arithmetrische Massen- und Grundflachen-Mittelstammes gleichmäszige gut durchforsteter Bestände ist in Praxis meist so geringfügig und..."

We see that experience shows that $\bar{v} - v_{\bar{g}}$ is smaller if the stands are regular (small s). Our formula indicates the same. It is also clear why investigators in elder Scotch pine stands find a smaller difference. In these cases $(b-2)$ as well as $s\bar{d}^{-1}$ are small.

But as we cannot expect a small value of c'' in every case, so $v_{\bar{g}}$ gives a considerably biased estimate of $\bar{v}$.

8. HOHENADL'S METHOD

HOHENADL states that the average of the trees with diameters $\bar{d} + s$ and $\bar{d} - s$ equals $\bar{v}$.

If we use BERKHOUT's formula we get:

$$v_{(\bar{d}+s)} = a(\bar{d} + s)^b = v_+$$

$$v_{(\bar{d}-s)} = a(\bar{d} - s)^b = v_-$$

$$v_+ = a(\bar{d}^b + bs\bar{d}^{b-1} + \tfrac{1}{2} b(b-1) s^2 \bar{d}^{b-2} + \ldots$$

$$v_- = a(\bar{d}^b - bs\bar{d}^{b-1} + \tfrac{1}{2} b(b-1) s^2 \bar{d}^{b-2} + \ldots$$

$$\tfrac{1}{2}(v_+ + v_-) = a(\bar{d}^b + \tfrac{1}{2} b(b-1)s^2 \cdot \bar{d}^{b-2}) = a\bar{d}^b (1 + \tfrac{1}{2} b(b-1)s^2 \bar{d}^{-2})$$

The second part of this equation is the same as in (4, 3, 2) and we find $\bar{v} = \tfrac{1}{2}(v_+ + v_-)$.

9. ACCURACY OF THE METHOD

To compute the accuracy of the method we make use of (4, 3, 2):

$$\bar{v} = a\bar{d}^b (1 + \tfrac{1}{2} b(b-1) s^2 \bar{d}^{-2}).$$

In calculating b for Douglas-fir and Larch we found an estimate for $\sigma_{\bar{b}}$.

Estimates of σ_b are found from the formula $\sigma_b = \sigma_{\bar{b}} \sqrt{n}$ where n is the number of observations used to compute $\bar{b}$. We find

$S(\sigma_b)$ for Douglas fir: $0.014 \sqrt{33} = 0.08$

$S(\sigma_b)$ for Larch: $0.012 \sqrt{23} = 0.06$

Assume $\sigma_b = 0.07$.

The error that occurs from b is caused by the term $r = \tfrac{1}{2} b(b-1)$. An estimate of σ_r is found as $S(\sigma_r) = (b^2 \sigma_b^2 + \tfrac{1}{4} \sigma_b^2)^{\frac{1}{2}} = \sigma_b(b^2 + \tfrac{1}{4})^{\frac{1}{2}}$ $S(\sigma_r) = \sigma_b \sqrt{6.01} = 2.45\sigma_b = 0.17$.

If we assume $s^2 \bar{d}^{-2} = \frac{1}{16}$ we find for the standard error in the volume $0.01\, v_{\bar{d}}$. So, if we compare the results of our method with the computation of the total volume by summation of the volumes of each tree (volume table), we must take a standard error of 1.0 % into account.

In practice this error has little influence on the total accuracy. If the standard error estimating $v_{\bar{d}} = 0.05\, v_{\bar{d}}$, we find for the total standard error $(0.05^2 + 0.01^2)^{\frac{1}{2}} v_{\bar{d}} = 0.051\, v_{\bar{d}}$. In such a case the influence is only 0.1 %.

CHAPTER V

CALLIPER AND TREE FORK

1. THE CALLIPER

The instrument commonly used in forestry practice for diameter measurements is the calliper. All investigations in the Netherlands concerning diameters are done with the calliper.

It is a common practice to speak about "the" diameter of a tree although, it is well known that the crosssection of the tree can be sufficiently approximated by an ellips.

Speaking about "the" diameter of a tree we indicate with this expression the expected value (d) we will get by repeating the calliper measurement on the crosssection of the tree many times, assuming that every point of the crosssection has the same probability to be touched by the calliper. This expectation can be computed as follows. Consider the crosssection to be an ellips in the equation:

$$x^2 a^{-2} + y^2 b^{-2} = 1$$

Let: $k = (a^2 - b^2)(a^2 + b^2)^{-1}$ and $c^2 = a^2 + b^2$

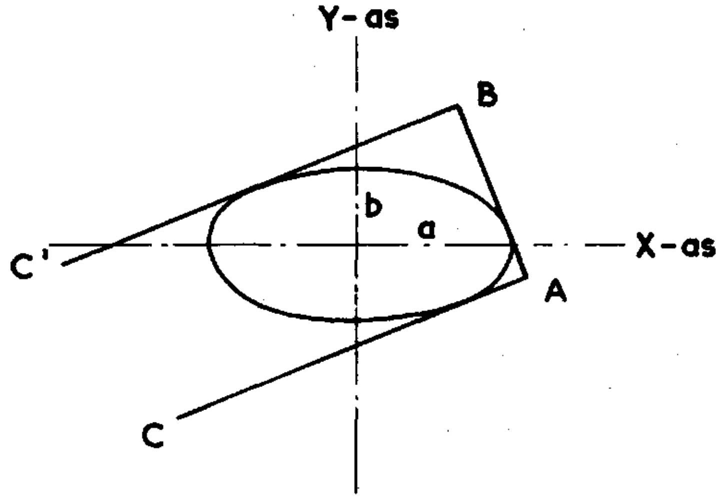


FIG. 6

Consider the calliper AB in the position given in fig. (6).

The tangents AC and BC′ have the equations: $y = mx \pm (a^2 m^2 + b^2)^{\frac{1}{2}}$; $m = \text{arc tg.}\ \varphi$. We find:

$$AB = 2(a^2 m^2 + b^2)^{\frac{1}{2}} (1 + m^2)^{-\frac{1}{2}} = 2(a^2 \sin^2 \varphi + b^2 \cos^2 \varphi)^{\frac{1}{2}} = (1 - k \cos 2\varphi)^{\frac{1}{2}} c\sqrt{2}.$$

As every angle φ has the same probability density of occurence we find, taking the symmetry of the ellips into account

$$(4,1,1) \qquad d = \pi^{-1} c \sqrt{2} \int_0^{\pi} (1 - k \cos 2\varphi) d\varphi = \left(1 - \tfrac{1}{16} k^2 - \tfrac{15}{1024} k^4 \ldots\right) c \sqrt{2}.$$

2. THE TREE FORK

Some time ago Prof. Dr. H. J. BECKING constructed a tree fork which proved to be very useful in practice.

This treefork is composed of two metal rulers with a fixed angle (α). The rulers are provided with a scale corresponding with the diameter classes of a cylinder, caught between the two rulers (fig. 7a).

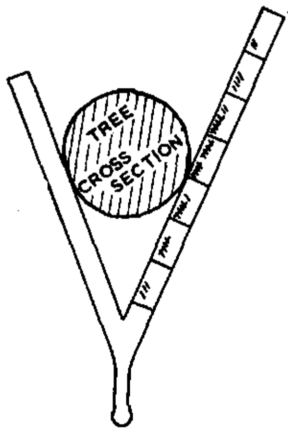


FIG. 7a

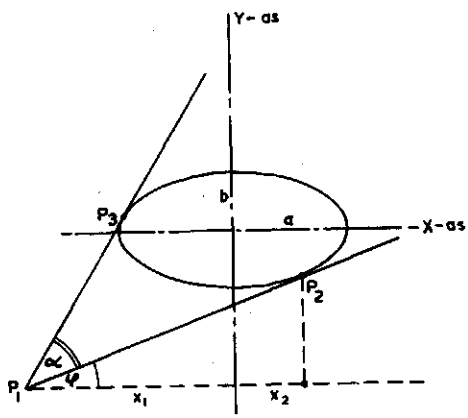


FIG. 7b

The treefork is placed perpendicularly to the tree axis at breast height. The diameter class of the tree can be read on the contact point of the tree circumference and the ruler.

A sheet of paper with the same scale is attached to the ruler to register the diameter classes of the measured trees. In this way the frequency distribution of the diameters of a stand is obtained, precluding reading errors and errors caused by misunderstanding which occur frequently when the calliper is used.

The treefork can be easily handled by one man, which allows a more efficient organization of the work. A time study has shown that the average time needed for a treefork measurement is 0.109 min (2254 observations) and 0.143 min (1318 observations) for a calliper maesurement.

For the computation of the expected value for the fork measurement we use the same notations as in the previous case. Consider the treefork $P_3P_1P_2$ in fig. 7b. The coordinates of P_1 and P_2 are $(-x_1; y_1)$ and $(x_2; y_2)$.

If φ (resp. ψ) is the angle between P_2P_1 (resp. P_3P_1) and the x-axis, we put $\mathrm{tg}\,\varphi = m_1$; $\mathrm{tg}\,\psi = m_2$; $\mathrm{tg}\alpha = q$; $\psi - \varphi = \alpha$;

The value which we are interested in, is the expectation of $2(P_1P_2)\,\mathrm{tg}\,\frac{1}{2}\alpha$. $2(P_1P_2)\,\mathrm{tg}\,\frac{1}{2}\alpha$ can be considered a function of φ. If we assume that every angle φ has the same probability density of occurence and we take the symmetry of the ellips into account, we find

$$2\,E\{(P_1P_2)\,\mathrm{tg}\,\tfrac{1}{2}\,\alpha\} = E\{f(\varphi)\} = \pi^{-1} \int_0^{\pi} f(\varphi)\, d\varphi$$

The lines P_2P_1 and P_3P_1 have the equations:

(5, 2,1) $$y = m_1 x - (a^2 m_1^2 + b^2)^{\frac{1}{2}} \text{ and}$$

(5, 2,2) $$y = m_2 x + (a^2 m_2^2 + b^2)^{\frac{1}{2}}$$

From (5, 2, 1) and (5, 2,2) we find the solution for x_1. For x_2 we find:

(5, 2,3) $$x_2 = a^2 m_1 (a^2 m_1^2 + b^2)^{-\frac{1}{2}}.$$

As $q = \text{tg}\alpha$ we have

(5, 2,4) $$q = (m_2 - m_1)(1 + m_1 m_2)^{-1}.$$

Using (5, 2, 1), (5, 2, 2), (5, 2, 3) and (5, 2, 4) and taking $P_1P_2 = (x_2 - x_1) \cos^{-1}\varphi$, we find after some computations:

(5, 2,5) $$d' = E\{f(\varphi)\} = \left(1 - \tfrac{1}{16} k^2 - \tfrac{15}{1024} k^4 \ldots\right) c\sqrt{2}.$$

Hence we find that the expected value d' is independant of α and equals d. In other words: Any treefork gives the same expected value as the calliper.

In case of stands with large diameters we can use forks with a greater angle α. This is of considerable practical importance, since the size (weight) of the instrument can remain small. It is of some interest to show that the accuracy of the instrument need not decrease if a greater angle is used in case of a greater average diameter.

For practical reasons we can better take the class intervals on the rulers not less than 2 cm. Take 2 cm as a fixed value. This means that the corresponding diameter class increases with the angle α.

It is well known that the coefficient of variation (V_c) of the mean, caused by the use of diameter classes can be computed as $V_c = a.(12n)^{-\frac{1}{2}} \bar{d}^{-1}$, where n is the number of observations from which $\bar{d}$ is computed and a the length of the diameter class. If the scale on the ruler is graduated in classes of length a', we have for the corresponding diameter class $a = 2a'\text{tg}\frac{1}{2}\alpha$ (α is the angle of the treefork); V_c can be kept constant by taking such an α that $\bar{d}^{-1} \text{tg}\frac{1}{2}\alpha$ remains constant.

Moreover, practical use shows that it is feasible to choose such an α and $\bar{d}$ that V_c tends to decrease when α and $\bar{d}$ increase. This is illustrated by an example in table 19 in which we suppose the length of the ruler to be 50 cm., $a' = 2$ cm and $n = 100$.

TABLE 19.

Range of $\bar{d}$, average diameter, in cm	tg $\frac{1}{2}\alpha$	Maximum measurable diameter in cm	Diameter class in cm	max Vc
10– 25	0.50	50	2	0.006
25– 50	0.75	75	3	0.003
50– 70	1.00	100	4	0.002
70– 90	1.25	125	5	0.002
90–150	2.00	200	8	0.003
> 150	2.50	250	10	0.002

CHAPTER VI

THE CONE METHOD FOR THE ESTIMATION OF THE MEAN HEIGHT OF THE TREES IN A STAND

In 1955 I introduced[1]) a method to estimate the mean height of the trees in a stand. This so called Cone Method is related to the method of BITTERLICH for the estimation of the total basal area pro square unit.

1. HORIZONTAL TERRAIN

The method is as follows:

Let the ground in the forest be flat and horizontal. Consider a circular cone C with its top A on the ground and with vertical axis (fig. 8).

The top angle of the cone is $(\pi - 2\alpha)$. We count the number n of the tree tops in this cone[2]) and we will show that this number yields an estimate of Σh^2, the sum of squares of heights pro unit area.

Consider the tree tops in a forest. The number of tops in a small domain with volume W is a random variable with expectation $W \cdot \varphi(h)$.

$\varphi(h)$ is a function, only depending on the height of h, which may be called the expected density of the tree tops at height h. The expected number of tree tops is now

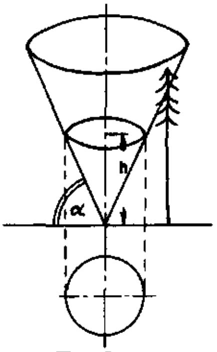


FIG. 8

(6, 1,1)
$$E(\underline{n}) = \int_C \varphi(h)\, dx\, dy\, dh;\; C = x^2 + y^2 \leqslant h^2 \operatorname{cotg}^2 \alpha;\; h > 0$$
$$= \int_0^\infty \varphi(h) \cdot \pi h^2 \operatorname{cotg}^2 \alpha dh$$
$$= \pi \operatorname{cotg}^2 \alpha \int_0^\infty \varphi(h) \cdot h^2\, dh$$

On the other hand, in case D is a vertical cylinder with unit cross section, we have for the expectation of $\Sigma_D \underline{h}^2$

$$\Sigma h^2 = E(\Sigma_D \underline{h}^2) = \int_0^\infty \varphi(h)\, h^2\, dh = E(\underline{n})\, \pi^{-1} \operatorname{tg}^2\alpha$$

An unbiased estimate of Σh^2 is therefore given by

(6, 1,2)
$$\underset{\cdot}{S}(\Sigma h^2) = \pi^{-1} n \operatorname{tg}^2\alpha$$

[1]) At the same time I introduced the cone method (*Nederlandsch Bosbouw Tijdschrift* **27** (1955) 11 (Nov.) 285–287. Mr Taneo Hirata published independently of my studies a similar method called „*Vertical angle count sampling*" (*Journal of the Japanese Foresty Society*) **37** (1955) 11 (Nov.) 479–480.

[2]) In practice we will not take the top of the cone on the ground, but at eye height. We will show later which are the consequences of this.

If N is the number of trees per unit area, then the quadrate mean h_0^2 of the heights in the stand is:

$$h_0^2 = N^{-1} \Sigma h^2$$

h_0 can be estimated by

(6, 1,3) $$\underset{\cdot}{S}(h_0) = (\pi N)^{-\frac{1}{2}} n^{\frac{1}{2}} \operatorname{tg}\alpha$$

h_0 can be regarded as a representative for the heights of the trees in the stand (compare the regression heights: $h_{\bar{d}}$, $h_{\bar{g}}$, $h_{\bar{v}}$). We want to compare this estimate with $h_{\bar{d}}$.

Consider the formula of STOFFELS (STOFFELS and VAN SOEST 1953):

$$E(\underline{h}) = md^n$$

where the expectation is considered under the condition that the diameter of the tree is d. m and n are constants.

The diameters d in the stand have a frequency distribution. Let $\underline{d}$ be the relevant random variable with expectation $\bar{d}$. We call $\underline{d} - \bar{d} = \underline{u}$. We now compute, unconditionally,

$$h_0^2 = E(\underline{h}^2) = E\, m^2(\bar{d} + \underline{u})^{2n}$$

This is approximately

$$E\{ m^2(\bar{d}^{2n} + 2n\bar{d}^{2n-1}\,\underline{u} + \tfrac{1}{2}\cdot 2n(2n-1)\,\bar{d}^{2n-2}\,\underline{u}^2)\} =$$
$$= m^2\,\bar{d}^{2n}\,[1 + 0 + n(2n-1)\,\bar{d}^{-2}\,E(\underline{u}^2)]$$

As $m^2\bar{d}^{2n} = h_{\bar{d}}^2$, and $E(\underline{u}^2) = s^2$ we have:

$$h_0 = h_{\bar{d}}\,\{1 + \tfrac{1}{2}\,n\,(2n-1)\,\bar{d}^{-2}\,s^2\}$$

Estimates of n (in STOFFELS' formula) for Douglas fir and Scotch pine are $n = 0.545$ and $n = 0.305$ respectively. (Estimated from 8 and 10 stands resp.)

Taking the coefficient of variation $s\bar{d}^{-1} = \frac{1}{4}$ we find for Scotch pine $h_0 \approx 0.996\,h_{\bar{d}}$ and for Douglas fir $h_0 = 1.0015\,h_{\bar{d}}$. So h_0 can be considered a good approximation of $h_{\bar{d}}$ for practical purposes.

Remark: In case HENRIKSEN's formula is used to compare h_0 and $h_{\bar{d}}$ we find some more complicated equations, but the conclusion that the difference $h_0 - h_{\bar{d}}$ is small enough to be neglected in practice, remains the same.

However, it is not convenient to choose the top of the cone on the ground. It is chosen on eye height: $c = 1.70$ m above the ground. Using the same formula we do not get an estimate of Σh^2 but of $\Sigma(h-c)^2$, which we denote by $N\,h_c^2$.

$$h_c^2 = \Sigma(h^2 - 2ch + c^2)\,.\,N^{-1}$$
$$= h_0^2 - 2c\bar{h} + c^2;\ \text{substituting}\ \bar{h} = h_0[1 - \tfrac{1}{2}\{\sigma_{(h)}\}^2 h_0^{-2}]$$

we get $$= (h_0 - c)^2 + c\{\sigma_{(h)}\}^2 h_0^{-1}$$

$$h_c^2 = (h_0 - c)^2\,[1 + c\{\sigma_{(h)}\}^2 h_0^{-1}\,(h_0 - c)^{-2}]$$

$$h_c = (h_0 - c)\,(1 + \tau)$$

with $$\tau = \tfrac{1}{2}c\,\{\sigma_{(h)}\}^2 h_0^{-1}\,(h_0 - c)^{-2}$$

In order to estimate τ we computed $\sigma_h^2 h_0^{-2}$ in 30 different stands. For $h_0 > 9$ all values $\sigma_h^2 h_0^{-2}$ were $\ll 0.04$. If $c = 1.70$ m, $h_0 = 10.2$ m and $\sigma_h^2 = 0.04\, h_0^2$, then $\tau = 0.005$. For $h_0 = 17$ m we find $\tau < 0.001$. We may conclude that for practical purposes τ can be neglected. Hence:

(6, 1,4) $$h_0 = h_c + c$$

2. Slopes

In case the top of the cone is on sloping ground (fig. 9) the discussion remains the same as for horizontal ground.

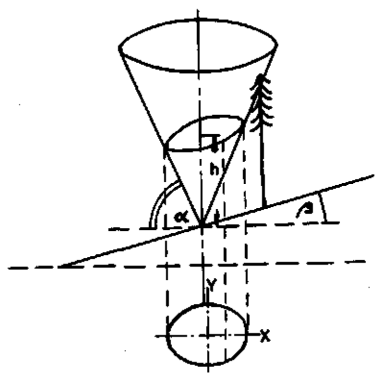


Fig. 9

Analogous to (6, 1,1) we find:

$$E(\underline{n}) = \int_{C'} \varphi(h)\, dx\, dy\, dh; \quad \text{with}$$

$$C' = x^2 + y^2 \leqslant (h + x\,\mathrm{tg}\,\beta)^2 \,\mathrm{cotg}^2\,\alpha$$

$$= \int_0^\infty \varphi(h)\, \pi h^2 \,\mathrm{cotg}^2\,\alpha\, (1 - \mathrm{cotg}^2\,\alpha\,\mathrm{tg}^2\,\beta)^{-1\frac{1}{2}}\, dh$$

(6, 2,1) $$E(\underline{n}) = \pi\,\mathrm{cotg}^2\,\alpha\,\Sigma h^2 (1 - \mathrm{cotg}^2\,\alpha\,\mathrm{tg}^2\,\beta)^{-1\frac{1}{2}}$$

$$\underset{\cdot}{S}\,\Sigma h^2 = n\pi^{-1}\,\mathrm{tg}^2\,\alpha\,(1 - \mathrm{cotg}^2\,\alpha\,\mathrm{tg}^2\,\beta)^{1\frac{1}{2}}$$

(6, 2,2) $$h_0 = N^{-\frac{1}{2}} (\Sigma h^2)^{\frac{1}{2}} = (\pi N)^{-\frac{1}{2}} n^{\frac{1}{2}}\,\mathrm{tg}\alpha\,(1 - \mathrm{cotg}^2\,\alpha\,\mathrm{tg}^2\,\beta)^{\frac{3}{4}}$$

The factor $(1 - \mathrm{cotg}^2\,\alpha\,\mathrm{tg}^2\,\beta)^{\frac{3}{4}}$ may be regarded as a correction factor for sloping ground with a slope β.
Table 20 (page 44) column 1, 2 and 3 shows the factor $(1 - \mathrm{cotg}^2\,\alpha\,\mathrm{tg}^2\,\beta)^{\frac{3}{4}}$ for some values of α and β.

Formula (6, 2,2) is the general formula for the expectation of h_0 when the cone is used with vertical axis.
The cone can also be used with an axis perpendicular to the ground (fig. 10).

TABLE 20.

β (slope) \ tgα (cone)	$(1-tg^2\alpha\ cotg^2\beta)^{\frac{1}{2}}$: 1 ($\alpha = 45°$)	$\sqrt{\frac{\pi}{2}}$ ($\alpha = 51° 25'$)	$\sqrt{\pi}$ ($\alpha = 60° 35'$)	$\cos^{-1\frac{1}{2}} \beta$
2° 30′	0.9986	0.9991	0.9995	1.0014
5°	0.9943	0.9963	0.9982	1.0057
7° 30′	0.9870	0.9917	0.9959	1.0130
10°	0.9766	0.9851	0.9926	1.0232
12° 30′	0.9629	0.9764	0.9882	1.0366
15°	0.9457	0.9655	0.9828	1.0533
17° 30′	0.9245	0.9521	0.9762	1.0737
20°	0.8989	0.9361	0.9682	1.0978
22° 30′	0.8683	0.9169	0.9588	1.1261

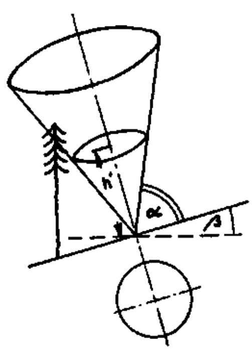


FIG. 10

In that case we find:

(6, 2,3) $$h_0 = n^{\frac{1}{2}} (\pi N)^{-\frac{1}{2}} \operatorname{tg} \alpha \cos^{-1\frac{1}{2}} \beta$$

If the cone is used in this way, the correction is independent of α.

Table 20 column 4 gives the correction factor for some values of β.

We notice in formula (6, 2,2) for vertical axis that the correction is smaller if a greater angle α is used.

As we are usually faced with more or less broken ground, we have to assume an average ground surface and to consider the errors due to unevenness of the ground as accidental errors, which do not affect the expected value but only enlarge the variance. Experimental results proved this statement to be acceptable.

3. THE CONOMETER

We introduced two types of instruments (conometers) to estimate the height (h_c) with the cone method.

a. *The mirror type* (fig. 11). The mirror type consists of a holder c that can be mounted on a stick and of three bars a, b and e. a is a freely suspended bar attached to c and at the lower end made heavier with a weight g. Bar a carries bar b, wich is revolving around an axis b′ in such a way that b can form an arbitrary angle β with a. The third bar e carries a mirror s and revolves on the axis of b (the mirror s revolves about an axis e′). Sights VV′ are attached to e. The use of the instrument is as follows: The line of sight VV′ is directed parallel to the ground surface by turning b and is fixed in this position. The angle β of a and b is the angle of the slope and can be read on a scale. If we need a cone angle $(\pi - 2\alpha)$ the mirror is fixed in such a way that the angle between s and e is $\frac{1}{2}\alpha$. We now count the number of trees that occur above the line of sight VV′ while e is turning around the axis b.

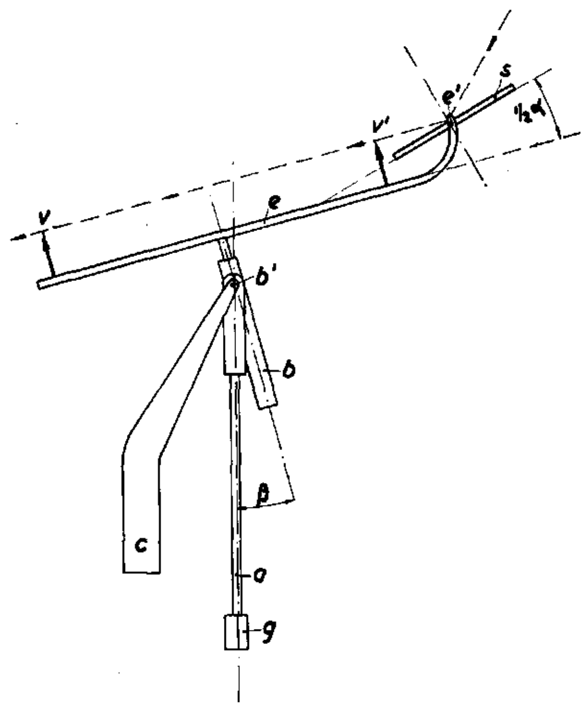


FIG. 11

Remark: It proved however, to be difficult to find (count) trees that are partially obscured by others, due to the limited range of vision of the mirror.

b. *The "free sight" type* (fig. 12 and foto). The "free sight" type also consists of a holder c and of three bars a, b and e, having the same function. e, however, is also revolving about an acis e″, carries no mirror and can be fixed, in such a position that the angle of e and b is $(\frac{1}{2}\pi - \alpha)$ (α resp. arc tg 0, arc tg 1, arc tg $\{2^{-\frac{1}{2}}\pi^{\frac{1}{2}}\}$). The use is as follows: We first put e perpendicular to b. ($\alpha =$ $=$ arc tg 0). Finally e is fixed in such a position that α is e.g. arc tg $\pi^{\frac{1}{2}}$. The trees are counted in the same way as discussed above.

This last type of instrument does not present the difficulties mentioned for the mirror type and therefore can be recommended as most suited for practical use.

4. Some experimental results

We applied the cone method in some Scotch pine stands in the forest ranges Ommen and Nijverdal and also in some chir pine (*Pinus longifolia*) stands in Dehra Dun (India).

Table 21 gives the results:

The "true heights" are found from height curves constructed with 30–40 measurements of diameter and height.

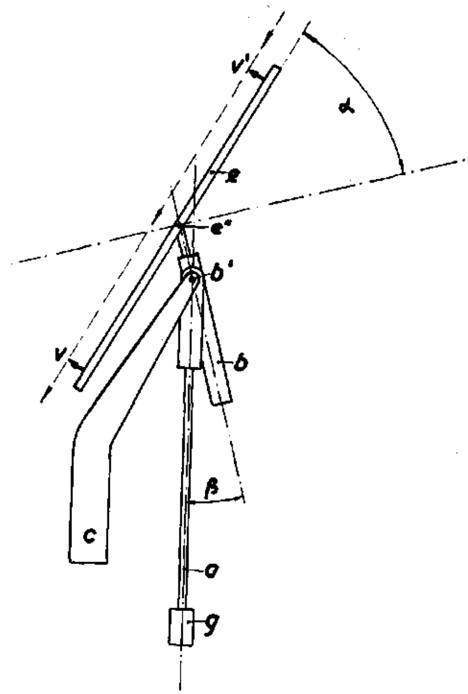


FIG. 12

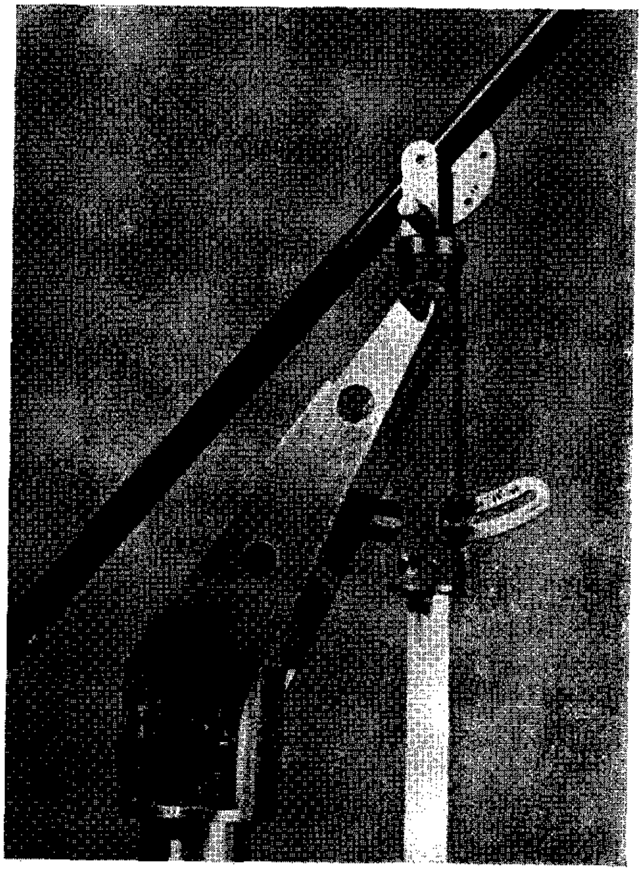
TABLE 21.

Scotch Pine Ommen Number of stand	number of cones	deviation from true height in %	Scotch Pine Nijverdal Number of stand	number of cones	deviation from true height in %	Chir Pine Dehra Dun Number of stand	number of cones	deviation from true height in %
9^a	9	1.9	77^a	2	0	75	4	–2.5
11^c O	4	0.8	135	4	+3.4	76	4	+1.1
11^c W	4	2.4	97^a	3	–3.3	77	4	–0.9
12^a O	4	–1.7	99^b	6	–4.5	78	4	–1.2
12^a W	4	+1.7	72^b	2	16.9	79	4	+1.5
13^b	4	+2.1	92^a	8	0	80	4	+1.7
31^b	4	–1.4	78^a	2	3.7	87	4	–1.5
40^a	4	+1.8				88	4	+4.7
60^a	5	–2.5				89	4	–1.8
						90	4	–10.2

5. ASPECTS OF THE CONE METHOD

a. A time study made in Nijverdal to compare the efficiency of the cone method with the measurement with the BLUME-LEISS hypsometer, shows the following results:

Cone method	(1 man labor)	0.854 min. pro cone
BLUME-LEISS	(2 man labor)	1.01 min. pro tree pro man
BLUME-LEISS	(1 man labor)	1.55 min. pro tree pro man



The conometer

b. The standard error in the estimate of $n^{\frac{1}{2}}$ was calculated and turns out to be smaller than the standard error in the estimate of $h_{\bar{d}}$ with the BLUME-LEISS, but the sample was too small to enable us to draw definite conclusions.
c. The estimation of the volume $Nv_{\bar{d}}$ is more efficient when the cone is applied even in the cases that the standard error of $n^{\frac{1}{2}}$ equals the standard error of $h_{\bar{d}}$ (BLUME-LEISS).

To illustrate this statement we use formula (4, 2, 1) in connection with BERKHOUT's formula.

$$N\, v_{\bar{d}} = N \gamma\, \bar{d}^{\beta}; \qquad \gamma = p\, \bar{d}^{q}\, h_{\bar{d}}^{r} \qquad \text{hence}$$
$$N\, v_{\bar{d}} = N\, p\, \bar{d}^{\beta+q}\, h_{\bar{d}}^{r}.$$

Suppose the standard error of N, $h_{\bar{d}}$ and $n^{\frac{1}{2}}$ is 5 % of the relevant expected values and that of $\bar{d} = 2.5\,\%$. In case the old hypsometer is used we find for the standard error in the estimate of $Nv_{\bar{d}}$

$$Nv_{\bar{d}}\{0.05^2 + (\beta+q)^2 \times 0.025^2 + r^2 \times 0.05^2\}^{\frac{1}{2}} = 0.05\left\{1 + r^2 + \left(\frac{\beta+q}{2}\right)^2\right\}^{\frac{1}{2}} Nv_{\bar{d}}$$

In case the cone is used the formula is

$$N\,v_{\bar{d}} = N\,p\,d^{\beta+q}\,\bar{d}^{\frac{1}{2}r}\,N^{-\frac{1}{2}r}$$
$$N\,v_{\bar{d}} = N^{(1-\frac{1}{2}r)}\,\bar{d}^{\beta+q}\,n^{\frac{1}{2}r}$$

and the standard error is $0.05\left\{(1-\frac{1}{2}r)^2 + \frac{1}{4}r^2 + \left(\frac{\beta+q}{2}\right)^2\right\}^{\frac{1}{2}} Nv_{\bar{d}}$

We notice that the influence of the standard error of N on the total estimate is much lower as N occurs in the denominator of the cone estimate.

The computations can be facilated by the use of tables (Table 22).

CHAPTER VII

ANOTHER METHOD FOR THE ESTIMATION OF VOLUME PRO H.A.

1. THE METHOD

As we saw the cone measurements give us an unbiased estimate of $\Sigma(h-c)^2$ pro ha $\approx N(h_{\bar{d}}-c)^2$ in a very easy way.

In the preceding chapter we used this estimate for the computation of $h_{\bar{d}}$. In this chapter we want to use the cone measurement to estimate directly the volume pro ha.

We assume that a relation exists between the expected value of $\underline{V}$ pro ha and the values $\Sigma(h-c)^2$, $\bar{d}$ and s and that the variance of the random variable $\underline{V}$ is small enough to permit us to estimate not only $E(\underline{V})$, but even V with the help of the relation; to simplify computations we used $b_{15} = d_{85\%} - d_{15\%} = 2.11$ s instead of s.

In order to find the relevant function for Douglas fir, we used 65 Douglas fir surveys in stands of different ages. The data were kindly provided by the Institute of Forestry Research of the University of Wageningen.

In all these stand V, $\Sigma(h-1.70)^2$, $\bar{d}$ and b_{15} were computed. (The number of trees counted in a cone with angle $\alpha = \text{arc tg}\pi^{\frac{1}{2}}$ and with top 1.70 m above the ground, is an estimate of $\Sigma(h-1.70)^2\,10^{-4}$ pro ha; this number will be designated by n.) The subdivision Wageningen of the statistic department TNO was kind enough to analyse the data. First the relation between V, n, $\bar{d}$ and b_{15} was studied roughly with a graphical method. (Compare: EZEKIEL). We found that a linear relation would suffice. There was no reason to assume that the variance of V was correlated to one of the variables, except in the case of V and n, where it seemed that the variance of V was positively correlated with n. Secondly a short computation was made to see if the variables n, $\bar{d}$ and b_{15} were too highly correlated, in which case they can not be considered to represent independent influences. In spite of the fact that $\bar{d}$ and b_{15} had a fairly high correlation the "bunch analysis" gave no reasons to exclude one of these variables if the whole set is considered. We came to the conclusion that

TABLE 22. For $(h_{\bar{d}} - c)$ in meters. N is the stem number pro ha. n is the number of tree tops counted in a vertical cone with top angle $\frac{1}{2}\pi$ ($\alpha = \frac{1}{4}\pi$). c is the height of the cone top above the ground

N \ n	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40
100	11.28	13.82	15.96	17.84	19.54	21.11	22.57	23.94	25.23	26.46	27.64	28.77	29.85	30.90	31.91				
125	10.09	12.36	14.27	15.96	17.48	18.88	20.18	21.41	22.57	23.67	24.72	25.73	26.70	27.64	28.55	29.42	30.28	31.11	31.91
150	9.21	11.28	13.03	14.57	15.96	17.23	18.42	19.54	20.60	21.61	22.57	23.49	24.37	25.23	26.06	26.86	27.64	28.40	29.13
175	8.53	10.45	12.06	13.49	14.77	15.96	17.06	18.09	19.07	20.00	20.89	21.75	22.57	23.36	24.12	24.87	25.59	26.29	26.96
200		9.77	11.28	12.61	13.82	14.93	15.96	16.92	17.84	18.71	19.54	20.35	21.11	21.85	22.57	23.26	23.94	24.59	25.23
225		9.21	10.64	11.89	13.03	14.07	15.05	15.96	16.82	17.64	18.42	19.17	19.90	20.60	21.28	21.93	22.57	23.18	23.79
250		8.74	10.09	11.28	12.36	13.35	14.27	15.14	15.96	16.73	17.48	18.19	18.88	19.54	20.18	20.81	21.41	22.00	22.57
275		8.33	9.62	10.76	11.79	12.73	13.61	14.43	15.21	15.96	16.67	17.35	18.00	18.63	19.25	19.84	20.41	20.98	21.52
300			9.21	10.30	11.28	12.19	13.03	13.82	14.57	15.27	15.96	16.60	17.23	17.84	18.42	18.97	19.54	20.08	20.60
325			8.85	9.90	10.84	11.71	12.52	13.27	13.99	14.68	15.33	15.96	16.55	17.13	17.71	18.25	18.78	19.29	19.79
350			8.53	9.54	10.45	11.28	12.06	12.80	13.49	14.14	14.77	15,37	15.96	16.51	17.06	17.58	18.09	18.59	19.07
375			8.24	9.21	10.09	10.90	11.65	12.36	13.03	13.66	14.27	14.86	15.42	15.96	16.48	16.99	17.48	17.95	18.42
400				8.92	9.77	10.56	11.28	11.97	12.61	13.23	13.82	14.39	14.93	15.45	15.95	16.45	16.92	17.39	17.84
425				8.65	9.48	10.24	10.95	11.61	12.24	12.84	13.41	13.95	14.48	14.99	15.48	15.96	16.42	16.87	17.31
450				8.41	9.21	9.95	10.64	11.28	11.89	12.47	13.03	13.56	14.07	14.57	15.05	15.51	15.96	16.39	16.82
475				8.19	8.97	9.69	10.35	10.98	11.58	12.14	12.68	13.20	13.70	14.18	14.64	15,10	15.53	15.96	16.37
500					8.74	9.44	10.09	10.70	11.28	11.83	12.36	12.86	13.35	13.82	14.27	14.71	15.14	15.55	15,96
525					8.53	9.21	9.85	10.45	11.01	11.55	12.06	12.56	13.03	13.49	13.93	14.35	14.77	15.17	15,57
550					8.33	9.00	9.62	10.21	10.76	11.28	11.79	12.27	12.73	13.18	13.61	14.03	14.43	14.83	15,21
575					8.15	8.80	9.41	9.98	10.52	11.04	11.52	12.00	12.45	12.88	13.31	13.72	14.11	14.50	14,87
600						8.61	9.21	9.77	10.30	10.80	11.28	11.75	12.19	12.61	13.03	13.43	13.82	14.20	14.57
625						8.44	9.03	9.57	10.09	10.59	11.06	11.51	11.95	12.36	12.77	13.16	13.53	13.91	14.27
650						8.28	8.85	9.39	9.90	10.38	10.84	11.28	11.71	12.12	12.52	12.90	13.27	13.64	13.99
675						8.13	8.69	9.21	9.71	10.19	10.64	11.07	11.49	11.89	12.29	12.66	13.03	13.38	13.74
700							8.53	9.05	9.54	10.00	10.45	10.87	11.28	11.68	12.06	12.43	12.80	13.14	13.49
750							8,24	8.74	9.21	9.66	10.09	10.50	10.90	11.28	11.65	12.01	12.36	12.70	13.03
800								8.46	8.92	9.36	9.77	10.17	10.56	10.93	11.28	11.63	11.97	12.30	12.61
850								8.21	8.65	9.07	9.48	9.87	10.24	10.60	10.95	11.28	11.61	11.93	12.24
900									8.41	8.82	9.21	9.59	9.95	10.30	10.64	10.97	11.28	11.60	11.89
950									8.19	8.59	8.97	9.33	9.69	10.03	10.35	10.67	10.98	11.28	11.58
1000										8.37	8.74	9.10	9.44	9.77	10.09	10.40	10.70	11.00	11.28
1050										8.17	8.53	8.88	9.21	9.54	9.85	10.15	10.45	10.73	11.01
1100											8.33	8.67	9.00	9.32	9.62	9.92	10.21	10.49	10.76
1150											8.15	8.48	8.80	9.11	9.41	9.70	9.98	10.26	10.52

for practical purposes a sufficient estimate of V can be given by a simple linear regression formula as:

(7, 1,1) $$V = p + qn + r\bar{d} + t\, b_{15}$$

The calculation gave:
$q = 6.76$, $r = 6.89$, $t = 4.65$ and $p = -78.58$ for V in m^3, n as stated, $\bar{d}$ in cm and b in cm. Tables 23 gives the correlation matrix.

TABLE 23

	V	100 n	$\bar{d}$	b_{15}
V	1	0.72621	0.93710	0.71425
100 n	0.72621	1	0.50524	0.27658
$\bar{d}$	0.93710	0.50524	1	0.72254
b_{15}	0.71425	0.27658	0.72254	1

The multiple correlation coefficient is $R_{V \rightarrow n, \bar{d}, b_{15}} = 0.9867$. The standard error around the regression plane is 6.8 % of the mean.

For the construction of a table with (7, 1, 1) we use the following expression:

(7, 1,2) $$V = N + B + D$$

We obtain $V = N + B + D$ with $N = qn$, $B = t\, b_{15}$ and $D = r\,\bar{d} + p$ with the help of table 24 for N, B and D.

TABLE 24

V = N + B + D					
I	N	B	D	I	D
1	6.76	4.65	Δ D = 6.89	26	100.5530
2	15.52	9.302	–	27	107.4426
3	20.28	13.953	–	28	114.3322
4	27.04	18.604	–	29	121.2218
5	33.80	23.255	–	30	128.1114
6	40.56	27.906	–	31	135.0010
7	47.32	32.557	–	32	141.8906
8	54.08	37.208	–	33	148.7802
9	60.84	41.859	–	34	144.6698
10	67.60	46.510	–9.6806	35	162.5594
11	74.36	51.161	–2.7910	36	169.4490
12	81.12	55.812	4.0986	37	176.3386
13	87.88	60.463	10.9882	38	183.2282
14	94.64	65.114	17.8778	39	190.1178
15	101.40	69.765	24.7674	40	197.0074
16	108.16	74.416	31.6570		
17	114.92	79.067	38.5466		
18	121.68	83.718	45.4362		
19	128.44	88.369	52.3258		
20	135.20	93.020	59.2154		
21	141.96	–	66.1050		
22	148.72	–	72.9946		
23	155.48	–	79.8842		
24	162.24	–	86.7738		
25	169.00	–	93.6634		

Use of table 24:

Suppose our measurement is done with a cone angle $\alpha = $ arc tg$\pi^{\frac{1}{2}}$ ($\alpha = 60°35'$). We compute $\bar{n}$, $\bar{d}$ and b_{15}. In column N we find the number N_1 that corresponds with n in column I. In column B we find the number B_1 that corresponds with b_{15} in column I, and in column D we find the number D_1 that corresponds with $\bar{d}$ in I. We take for V in m^3:

$$V = N_1 + B_1 + D_1$$

2. APPLICATION IN PRACTICE

In practice the survey can be done by two persons A and B as follows: A counts the trees in about 10 cones, chosen systematically in the stand and B measures about 100 diameters, using BECKING's tree fork. The average $\bar{n} = 10^{-1}\Sigma n$ is calculated.

a. In the forest stand O.N.O. (1) in Wageningen the cone method was applied ($\alpha = 45°$) to estimate V. From 4 cone measurements we found $\bar{n}' = 40.75$. This corresponds with $\bar{n} = 40.75\,\pi^{-1} = 12.97$ for $\alpha = 60°35'$. $\bar{d}$ was estimated from 100 diameter measurements as 12.35 cm, b_{15} is computed as (14.89 – 10.02) cm = 4.87 cm.

We find in column N for 12.97 the value 87.68, in column B for 4.87 the value 22.65 and in column D for 12.37 the value 6.51.

V pro ha. = $(87.68 + 6.51 + 22.65)m^3 = 116.84\ m^3$.

From a total measurement[1]) (all diameters and height curve) we knew: V pro ha = 110.31; deviation is 6.53 m^3 or + 5.9%.

b. The method can be applied, using the data in table 18. From this table we compute $n = 10^{-4}\,\Sigma(h - 1.70)^2 = 20.87$, $\bar{d}$ is 30.45 cm and $b_{15} = (37.78 - 24.10)$ cm = 13.68 cm. In the same way as discussed in the first example we find from table 24:

V pro ha = $(141.08 + 131.21 + 63.60)m^3 = 336\ m^3$.

From the total measurement we find $V_{\text{pro ha.}} = 90.162 : 0.2763 = 330\ m^3$; deviation 1.8%.

3. A STANDARD VOLUME TABLE

The formula $v = ad^{2.4}$ with $a = 0.0597\,\bar{d}^{-0.5403}\,h_{\bar{d}}^{0.978}$, gives

$$v = 0.0597\, d^{2.4}\, \bar{d}^{-0.5403}\, h_{\bar{d}}^{0.978}, \quad (7, 3,1)$$

For Douglas fir, according to (7, 3,1), we can find the volume of a tree if its diameter d, the average diameter in the stand $\bar{d}$ and the regression height $h_{\bar{d}}$ are known, so (7, 3,1) defines a standard volume table for Douglas fir (compare STOFFELS 1953).

[1]) Measurement after thinning 1957.

CHAPTER VIII

ON THE ESTIMATION OF THE INCREMENT BY THE AID OF THE INCREMENT BORER

1. INTRODUCTION

For an efficient forest management the increment of the volume is as important as the volume itself.

Several methods have been introduced to estimate the increment. One of the most simple methods is the estimation of the increment of a single tree whose increment is considered to be the average increment of the whole stand. Usually the mean basal area tree (diameter $d_{\bar{g}}$) is chosen for this purpose. By the aid of the increment borer the average number of annual rings on the last cm is determined and SCHNEIDER's formula $i_p = k(n\, d_{\bar{g}})^{-1}$ is used, in which k is a constant, n the number of annual rings on the last cm and i_p the increment percentage. In this chapter the increment is also estimated by the aid of an increment borer, but in connection with our discussions in Chapter IV the problem is solved along new lines.

2. INCREMENT OF THE MEAN VOLUME TREE

Consider a forest stand at present. The volume pro ha is V, the number of trees N, the mean volume $\bar{v}$, the average diameter $\bar{d}$, etc. Consider on the other hand the same stand a short period ago (3–5 years). The properties of the old stand will be denoted as above, but with $'$. (The mean volume was $\bar{v}'$ etc.).

The increment is $V - V' = N\bar{v} - N\bar{v}' = N(\bar{v} - \bar{v}')$. (provided there have been no thinnings). We wish to express the increment $\bar{v} - \bar{v}'$ as a function of $d_{\bar{v}} - d'_{\bar{v}}$. If we use the increment borer, we find the difference $d_{\bar{v}} - d'_{\bar{v}}$. If $d'_{\bar{v}} = d_{\bar{v}'}$ is an acceptable approximation, the increment given by the borer can be used as an estimate of $d_{\bar{v}} - d_{\bar{v}'}$.

It is commonly assumed in forestry that in even aged stands a linear relation exists between the expectation of the diameter increment and the diameter. As experiments showed this statement to be acceptable (PRODAN (1951) and others) we will use it for our computations. So we take

(8, 2,1) $$d' = ud + w$$

From (4, 3,4) and (8, 2,1) we find:

$$d'_{\bar{v}} = ud_{\bar{v}} + w = u\bar{d}c + w \quad \text{with} \quad c = 1 + \tfrac{1}{2}(b - 1)\, s^2\bar{d}^{-2}$$

and

$$d_{\bar{v}'} = \bar{d}'\, c' = (u\bar{d} + w)c' = u\bar{d}c' + c'w \quad \text{with} \quad c' = 1 + \tfrac{1}{2}(b - 1)\, s'^2\, \bar{d}'^{-2}$$

The difference $d'_{\bar{v}} - d_{\bar{v}'} = u\bar{d}(c - c') + w(1 - c')$ is usually small enough to be neglected in older stands (in such cases c can be taken equal to c′ for a short period. The regression increment w of the tree with diameter 0 multiplied by $0.7\ s^2\,\bar{d}^{-2}$ is also small enough). For practical purposes we therefore take $d'_{\bar{v}} = d_{\bar{v}'}$ and in consequence we state: The increment of the mean volume tree equals the mean increment of the stand.

3. COMPUTATION OF THE INCREMENT

The mean volume is $\bar{v} = a d_{\bar{v}}^{b}$ (BERKHOUT) with $a = p\bar{d}^{q} h_{\bar{d}}^{r}$ (compare 4, 2,1). From these two equations we find:

(8, 3,1) $$\bar{v} = p\bar{d}^{q} h_{\bar{d}}^{r} d_{\bar{v}}^{b} \quad \text{or}$$

$$\ln \bar{v} = \ln p + q \ln \bar{d} + r \ln h_{\bar{d}} + b \ln d_{\bar{v}}$$

Hence

(8, 3,2) $$d \ln \bar{v} = qd \ln \bar{d} + rd \ln h_{\bar{d}} + bd \ln d_{\bar{v}}$$

we now take

(8, 3,3) $$d \ln \bar{d} = d \ln d_{\bar{v}} = k^{-1} d \ln h_{\bar{d}}$$

From (8, 3,2) and (8, 3,3) we find

$$d \ln \bar{v} = qd \ln d_{\bar{v}} + krd \ln d_{\bar{v}} + bd \ln d_{\bar{v}} = (q + kr + b) d \ln d_{\bar{v}}$$

(8, 3,4) $$d \ln \bar{v} = Cd \ln d_{\bar{v}} \quad \text{with} \quad C = q + kr + b$$

(8, 3,4) gives: $$\ln \bar{v}' - \ln \bar{v} = C(\ln d'_{\bar{v}} - \ln d_{\bar{v}}) \quad \text{or,}$$

$$\ln \frac{\bar{v}'}{\bar{v}} = C \ln \frac{d'_{\bar{v}}}{d_{\bar{v}}}$$

Take $$\bar{v} - \bar{v}' = i \quad \text{and} \quad d_{\bar{v}} - d_{\bar{v}'} = \Delta \quad \text{then,}$$

$$\ln \frac{\bar{v} - i}{\bar{v}} = \ln \left\{\frac{d_{\bar{v}} - \Delta}{d_{\bar{v}}}\right\}^{C} = \ln (1 - \Delta\, d_{\bar{v}}^{-1})^{C}$$

$$\bar{v} - i = (1 - \Delta\, d_{\bar{v}}^{-1})^{C}\, \bar{v} = (1 - C\Delta\, d_{\bar{v}}^{-1} + \tfrac{1}{2} C (C-1) \Delta^{2} d_{\bar{v}}^{-2} + \ldots)\bar{v}$$

$$i = \{C\Delta\, d_{\bar{v}}^{-1} - \tfrac{1}{2} C(C-1) \Delta^{2} d_{\bar{v}}^{-2}\}\, \bar{v} = K\bar{v}$$

As i is the mean increment we find for the total increment $I = Ni$

(8, 3,5) $$I = K\,V \quad \text{with} \quad K = C\Delta\, d_{\bar{v}'}^{-1} + \tfrac{1}{2} C (C-1) \Delta^{2} d_{\bar{v}}^{-2}$$

The increment percentage is:

$$I_{p} = 10^{2}\, I \{\tfrac{1}{2} (V + V')\}^{-1} = 10^{2}\, I\, (V - \tfrac{1}{2} I)^{-1} = 10^{2}\, KV\, (V - \tfrac{1}{2} KV)^{-1} =$$
$$= 10^{2}\, K\, (1 - \tfrac{1}{2} K)^{-1} \approx (K + \tfrac{1}{2} K^{2})\, 10^{2}$$

4. SOME DATA FOR PRACTICAL USE

The factor C equals b only if $q + kr = 0$. This is usually not the case. Taking $q + kr = 0$ means that we assume that the height curve does not change (a remains constant). We know however from experience that the height curve changes with the age. Only in selection forests the curve remains constant. If we still use $C = b$, a considerable bias can be expected.

From (4, 2,1) we know that $q = -0.54$ and $r = 0.98$ for Douglas fir and -0.268 resp. 0.865 for Scotch pine (section 4, 2). We also saw in section (6, 1) that the height curve $h = md^{n}$ has an average $n = 0.545$ for Douglas fir and 0.305 for Scotch pine. For Douglas fir we find $C = 1.86 + 0.98\, k$ and for Scotch pine $C = 1.942 + 0.865\, k$.

Using these data we notice that if $k = n$, $q + kr \approx 0$ for both species. (Height curve remains on the same level).

The value k can be estimated from yield tables. From data published by VAN LAAR (1954) and GRANDJEAN-VAN SOEST (1953) we see that k varies with diameter and site class. From these data we took some rough estimates of k (an average site class) and computed C for Douglas fir. Table 25.

TABLE 25

diameter	k	$C = 1.86+0.98\,k$
10–20	0.95	2.80
20–30	0.80	2.65
>30	0.65	2.50

The value $C = 1.942 + 0.865\,k$ is also estimated for Scotch pine taking the values of k from the yield table constructed by GRANDJEAN and STOFFELS (1955), and given below in table 26.

TABLE 26

$\bar{d}$ \ Site Class	II	III	IV
10	2.91	2.84	2.84
14	2.77	2.71	2.63
18	2.67	2.59	2.50
22	2.55	2.44	2.29
26	2.43	2.25	–
28	2.38	2.17	–

5. APPLICATION IN PRACTICE

In the 24 years old Douglas fir stand O.N.O. (1) the volume pro ha was calculated in 1953 and in 1957 from a complete measurement[1]) (all diameters and height curve). Table 27 gives the calculation of I.

The diameter increment was also measured by the aid of the increment borer. We took 28 borings in all diameter classes and a straight line was fitted, giving the relation between the diameter and its increment. The equation was: $d' = 0.802\,d + 0.988$.

We found: $\bar{d} = 11.83$; $s^2\bar{d}^{-2} = 0.0416$; $d_{\bar{v}} = \bar{d}(1 + 0.7\,s^2\,\bar{d}^{-2}) = 12.17$ $c = 1.0291$; $c' = 1,0195$.

From the regression line we find $\Delta = 1.42$ cm. As $C = 2.80$ we calculated the increment as follows:

$$C\,\Delta\,d_{\bar{v}}^{-1} = 0.327; \quad \tfrac{1}{2}\,C\,(C-1)\,\Delta^2\,d_{\bar{v}}^{-2} = 0.034;$$

$$K = C\,\Delta\,d_{\bar{v}}^{-1} - \tfrac{1}{2}\,C\,(C-1)\,\Delta^2\,d_{\bar{v}}^{-2} = 0.293$$

$I = KV = (0.293 \cdot 133.8)\ m^3 = 39,3\ m^3$. The deviation is $(39.3 - 41.4\ m^3 = = -2.1\ m^3 = -5.1\%$.

[1]) 1953 after thinning; 1957 before thinning.

TABLE 27

	1953			1957		
d in cm	n	h in m	v in dm³	n	h in m	v in dm³
7	5	7.5	16	3	8.4	18
8	14	8.3	23	10	9.2	25
9	19	9.0	31	11	9.9	34
10	32	9.5	39	23	10.5	44
11	28	10.0	50	18	11.1	56
12	21	10.4	60	28	11.6	68
13	18	10.8	74	17	12.0	81
14	8	11.1	87	13	12.4	96
15				11	12.8	113
16				5	13.1	131
17				5	13.3	148
18				1	13.5	167

area = 0.075 ha $V_{\text{pro ha}}$ 1957 = 133.8 m³
$V_{\text{pro ha}}$ 1953 = 92.4 m³

$I_{\text{pro ha}}$ = 41.4 m³

For this young stand we also calculated $d'_{\bar{v}} - d_{\bar{v}'} = u\bar{d}(c - c') + w(1 - c')$; $d'_{\bar{v}} - d_{\bar{v}'} = 0.072$. If we take for Δ the correct value $1.42 + 0.07 = 1.49$ we find $K = 0.305$ and $I = KV = 40.7$ m³. Deviation: – 1.9 %.

SUMMARY

In this thesis we develop some theory about new methods to estimate the stem number N, the regression height $h_{\bar{d}}$, the diameter of the mean volume tree $d_{\bar{v}}$, and the increment I.

In Chapter I and II we discuss some well known methods for the determination of the stem number.

Chapter III gives the theory about the distance method. We propose to use the median m_4 of the distances measured from points chosen systematically (or at random) in the stand, to the fourth tree. Moreover, it is shown that in any case the estimate with m_4 is more efficient than estimates with the other medians (m_1 etc.) or with other parameters.

In Chapter IV we show that the mean volume of the trees in even aged stands can be calculated from the volume of the tree with the arithmetic mean diameter, the constant introduced by BERKHOUT and the coefficient of variation. *A formula for the diameter of the mean volume tree is introduced.*

In Chapter V we discuss the tree fork introduced by BECKING. The fork proved to be very useful in practice. It is shown that the expected value of „the" diameter of a tree is the same when the tree fork is used as when the calliper is used.

In Chapter VI the Cone Method, introduced in 1955 by MR. HIRATA and by the author independently of each other, is discussed for horizontal terrain and

for sloping ground. The conometer, an instrument for the application of the Cone Method, is introduced. It is shown that the height, estimated with the Cone Method can be regarded as an estimate of $h_{\bar{d}}$.

In Chapter VII we introduce a method to estimate the volume pro ha directly. Using this method we take full advantage of the Cone Method, discussed in Chapter VI.

In Chapter VIII we derive a formula to estimate the increment by the aid of the increment borer.

In all Chapters the relevant tables are given. The methods are tested in practice and examples are given concerning the application in forestry.

SAMENVATTING

In dit proefschrift worden na enige korte statistische beschouwingen, gewijd aan bekende schattingsmethoden in de bosbouw, nieuwe schattingsmethoden voorgesteld. Deze worden aan de hand van de statistische theorie afgeleid en geanalyseerd en experimenteel getoetst. Met tijdstudies wordt de efficiency der methoden nagegaan. Enige voor de praktijk belangrijke tabellen zijn geconstrueerd en voorts is bij een der methoden een voor de praktijk geschikt instrument voorgesteld. Aangezien de hoofdstukken verschillende uiteenlopende onderzoekingen en beschouwingen bevatten, volgt thans een samenvatting van elk der hoofdstukken afzonderlijk.

Na in de inleiding enige redenen te hebben opgegeven waarom de schatting van de staande houtvoorraad voor de bosbouw van belang is, is uiteengezet op welke wijze men streeft naar een schatting van de houtvoorraad die nauwkeurig genoeg is voor het gestelde doel en zo min mogelijk kost. In het bijzonder wordt aandacht besteed aan een schatting van de houtvoorraad volgens de formule:

$$\mathcal{S}(V) = \mathcal{S}(N) \cdot f\{\mathcal{S}(\bar{d}),\ \mathcal{S}(h_{\bar{d}})\}$$

Hierin is $\mathcal{S}(V)$ een schatting van de houtvoorraad V per oppervlakteeenheid, N het aantal bomen per oppervlakte eenheid (verder kortweg stamtal genoemd), $\bar{d}$ de gemiddelde diameter, $h_{\bar{d}}$ de regressiehoogte bij $\bar{d}$, en f een functie van twee variabelen. Al deze schattingen worden in de volgende hoofdstukken afzonderlijk besproken.

Hoofdstuk I bevat een korte uiteenzetting van de twee belangrijkste methoden ter bepaling van het stamtal nl. de z.g. monstervlakte-methode waarbij het aantal bomen op bepaalde steekproefoppervlakken (monstervlakten) wordt beschouwd, en de afstandmethode, waarbij afstanden tussen bomen worden gemeten (BAUERSACHS, KÖHLER). De door KÖHLER voor zijn formules gegeven afleiding blijkt onvoldoende gesteund te worden door de statistische theorie.

In hoofdstuk II wordt de door mij in 1954 gegeven benaderings formule voor de afstandmethode van BAUERSACHS en KÖHLER nader verklaard. De benaderingsformule toont goede overeenkomst met de door vorige onderzoekers gepubliceerde uitkomsten van experimenten; ook ons eigen experimenteel werk gaf aanleiding enig vertrouwen te schenken aan de benadering. Aan deze werkwijze kleven echter bezwaren.

Hoofdstuk III introduceert een nieuwe afstandmethode. Beschouw een willekeurig punt in een bos. De dichtstbijzijnde boom wordt 1e boom genoemd, de op één na dichtstbijzijnde 2e boom, enz.

De afstand $\underline{a}_n$ vanuit een willekeurig punt tot de n^{de} boom is een stochastiek (stochastische variabele). Is de verwachting $E(\underline{a}_n)$ klein (groot), dan bevat het bos veel (weinig) bomen per oppervlakte eenheid. Deze verwachting is dus afhankelijk van het aantal bomen per oppervlakte eenheid. Zij blijkt eveneens (zij het in mindere mate) afhankelijk van de onderlinge ligging der bomen. In verband hiermede wordt de kansverdeling van $\underline{a}_n$ (eventueel slechts enige parameters) afgeleid bij verschillende veronderstellingen omtrent de onderlinge ligging der bomen, te weten:

1. De bomen staan volgens toeval op het oppervlak verspreid (Poisson-bos, „Random Forest"). Aangezien de kansverdeling van Poisson de grondslag hierbij vormt, wordt een ruime plaats aan de afleiding van deze kansverdeling geschonken.
2. De bomen staan in de hoekpunten van een vierkantsnet (vierkantsrooster).
3. De bomen staan in de hoekpunten van een driehoekig rooster.

De gevallen 2 en 3 worden Systematische bossen („Systematic Forest") genoemd. De kansverdeling $\underline{a}_n$ wordt voor het Poisson-bos volledig afgeleid, terwijl voor de Systematische bossen de eerste twee momenten en de mediaan worden berekend.

De verkregen schattingen in een Poisson-bos zijn asymptotisch zuiver.

De drie genoemde parameters worden voor de drie gevallen vergeleken, waarbij blijkt, dat de medianen de minste onderlinge verschillen vertonen. De onderlinge verschillen nemen af met toenemende n. Het verschil tussen de mediaan van het Poisson-bos en het Systematisch bos is bij de eerste boom 15%, bij de vierde boom slechts 2,5% van de mediaan van het Poisson-bos.

Op grond van een voor de hand liggend model wordt aangenomen, dat bij de ligging der bomen in een bos zoals dat in de natuur wordt aangetroffen, de mediaan van de kansverdeling van $\underline{a}_n$ tussen die van het Systematisch en het Poisson-bos inligt, indien het stamtal in al deze gevallen gelijk is. Het verschil met de mediaan van het Poisson-bos zal kleiner zijn naarmate het verwachte aantal bomen per oppervlakte eenheid kleiner is.

Het stamtal kan het beste worden geschat met behulp van de mediaan van $\underline{a}_4$. Men berekent daartoe het stamtal alsof men met een Poisson-bos te doen heeft, en past een correctie toe die afhankelijk is van het gevonden stamtal. De toepassing wordt vergemakkelijkt door enige tabellen, waarin voor elke waarde van m_4 de verwachtingswaarde van het stamtal is gegeven.

De gestelde hypothesen betreffende de ligging der bomen in de natuur is getoetst door in zeven opstanden de afstanden $\underline{a}_1$, $\underline{a}_2$ $\underline{a}_3$ en $\underline{a}_4$ te meten. Gemiddeld zijn er per steekproef 300–400 metingen gedaan.

Het stamtal N was in elk der gevallen bekend. Bovendien wordt het stamtal geschat onder de veronderstelling dat men met een systematische en een toevallige ligging te doen heeft. De zo verkregen schattingen N_s en N_p liggen inderdaad aan weerszijden van N.

De gestelde hypothese is ook bevestigd door vergelijking van de in deze opstanden geschatte variatiecoëfficient met die van a_n in Systematische en Poisson-bossen.

De resultaten van een tijdstudie gecombineerd met de bekende variantie van $\underline{a}_n$ deden zien dat de afstand tot de vierde boom de meest efficiente is.

In hoofdstuk IV wordt een kort overzicht gegeven van de methoden die bestaan om de gemiddelde inhoud in bosopstanden te schatten. De door BERKHOUT gevonden betrekking tussen de verwachtingswaarde van de inhoud van een boom en de diameter daarvan ($E(\underline{v}) = ad^b$) wordt als uitgangspunt gekozen voor de berekening van de gemiddelde inhoud. De volgende formules worden afgeleid:

$$\bar{v} = v_{\bar{d}} \{ 1 + \tfrac{1}{2} b (b-1) s^2 \bar{d}^{-2} \}$$

$$\bar{v} - v_{\bar{g}} = \tfrac{1}{2} v_{\bar{d}}\, b (b-2) s^2 \bar{d}^{-2},$$

waarmee een zuivere schatting van $\bar{v}$ gevonden wordt en de resultaten van vorige onderzoekers kunnen worden verklaard. s kan worden geschat met de empirische betrekking $b_{15} = d_{0.85} - d_{0.15} = 2.1$ s. De constanten a en b in de formule van BERKHOUT werden berekend voor Douglas, terwijl voor Larix alleen b berekend is. De methode HOHENADL wordt eveneens met de gevonden betrekkingen verklaard.

Tenslotte wordt de schatting van $\bar{v}$ met enige voorbeelden uit de praktijk toegelicht.

Hoofdstuk V geeft een beschrijving van de door BECKING geconstrueerde boomvork.

De verwachtingswaarde van de meting met de boomvork blijkt onafhankelijk van de hoek van de vork te zijn, en gelijk aan de verwachtingswaarde van de meting met de boomklem.

De boomvork heeft vele voordelen in de praktijk.

Enige resultaten van tijdstudies worden vermeld.

De door mij in 1955 geïntroduceerde kegelmethode wordt besproken in hoofdstuk VI.

Deze methode komt op het volgende neer: Telt men in een bos het aantal boomtoppen n, dat zich binnen een kegel bevindt met verticale as en top op hoogte c boven de grond, dan is n een zuivere schatting voor $\Sigma(h-c)^2$ per eenheid van oppervlak, waarbij gesommerd is over alle bomen op de eenheid van oppervlakte, terwijl h de hoogte van de bomen voorstelt. Deze schatting kan worden gebruikt om met behulp van het stamtal N de hoogte $h_{\bar{d}}$ (regressiehoogte bij $\bar{d}$) op een voor de praktijk voldoende nauwkeurige wijze te schatten. De afwijkingen, die optreden bij hellend terrein, worden besproken en een correctie wordt vermeld. Een door mij geconstrueerd instrument (konometer), dat het mogelijk maakt ook bij hellend terrein een zuivere schatting te krijgen, wordt besproken. Tijdstudies waarbij de kegelmethode met de klassieke methode voor hoogtemeting wordt vergeleken, doen zien dat het bepalen van n korter duurt dan een hoogtemeting. Enkele resultaten van experimenten in Nederland en India worden gegeven.

In hoofdstuk VII wordt een werkwijze behandeld, waarbij de inhoud per oppervlakte eenheid wordt verkregen als een functie van n, het aantal getelde boomtoppen binnen de kegel, $\bar{d}$ en b_{15}. Voor de Douglas is een tabel geconstrueerd met behulp waarvan de massa bij toepassing van deze werkwijze gemakkelijk kan worden geschat.

Tenslotte worden de formules voor de samenstelling van standaard massatafels voor de Douglas gegeven.

Hoofdstuk VIII behandelt de schatting van de aanwas met behulp van boorspanen. Een korte afleiding doet zien, dat in de meeste gevallen de aanwas van de massa-middenstam kan worden beschouwd als de gemiddelde aanwas. Met behulp van enige, in vorige hoofdstukken gevonden vergelijkingen, wordt een formule voor de aanwasberekening afgeleid. Het in deze formule voorkomend onbenoemd getal blijkt van de gemiddelde diameter en de boniteit af te hangen. Een opgave van enige gegevens voor de practijk wordt gevolgd door een voorbeeld van aanwasbepaling in een te Wageningen gelegen Douglas opstand.

REFERENCES

1. Bauersachs, E. Bestandesmassenaufnahme nach dem Mittelstammverfahren des zweitkleinsten Stammabstandes. Forstw. Centrallblatt **64**, 182–186 (1942).
2. Becking, J. H. Massatafels voor de bepaling van de houtmassa van de douglasden (*Pseudotsuga taxifolia* Britt.) in Nederland (summary: Volume tables for the douglas fir in the Netherlands). Mededelingen van de Landbouwhogeschool Wageningen, Netherlands **50**, 1–16 (1950).
3. Berkhout, A. H. Het meten der boomen in verband met hun aanwas (mit deutscher Zusammenfassung). Mededelingen van de Landbouwhogeschool Wageningen, Netherlands **17**, 109-225 (1920).
4. Bitterlich, W. Die Winkelzählmessungen. Algemeine Forst und Holzwirtschaftliche Zeitung **58**, 94–96 (1947).
5. Eeden, Constance van. Maximum likelihood estimation of ordered probabilities. Indagationes Mathematicae **18**, 444–455 (1956).
6. Eeden, Constance van. Maximum likelihood estimation of partially or completely ordered parameters I. Indagationes Mathematicae **19**, 128–136 (1957).
7. Essed, F. E. De kegelmethode voor de bepaling van de gemiddelde hoogte in eensoortige bosopstanden. Nederlandsch Boschbouw-Tijdschrift **27**, 285–287 (1955).
8. Essed, F. E. A quick, simple and at the same time accurate method for estimating the total volume and the increment per cent of even-aged stands. Paper for the IV World Forestry Congress 1954, and in: Indian Forester **82**, 260–263 (1956).
9. Ezekiel, Mordecai. Methods of correlation analysis 2nd ed. New York en London, 1953.
10. Gehrhardt, E. Die theoretische und praktische Bedeutung des arithmetischen Mittelstammes. Meiningen, 1901.
11. Gehrhardt, E. Ein Vergleich des Kopezky-Gehrhardtschen Verfahren der Bestandsmassen ermittlung mit demjenigen von Neubauer und Tischendorf in Bezug auf Einfachheit und Genauigkeit. Tharandter forstl. Jahrbuch **84**, 328–344 (1933).
12. Grandjean, A. J. en J. van Soest. Opbrengstgegevens van de Douglas in Nederland. With a summary. Nederlandsch Boschbouw-Tijdschrift **25**, 90–112 (1953).
13. Grandjean, A. J. en A. Stoffels. Opbrengsttabellen van de groveden in Nederland. Nederlandsch Boschbouw-Tijdschrift **27**, 215–231 (1955).

14. HIRATA, TANEO. Vertical angle count sampling. Journal of the Japanese Forestry Society **37**, 479–480 (1955).
15. HOHENADL. Einführung in die Bestandsberechnung mit Hilfe von zwei Mittelstämmen. Forstwissenschaftliches Centralblatt **61**, 261–280 (1939).
16. KENNEY, J. F. & E. S. KEEPING. Mathematics of statistics. Part II, 2nd ed. New York, 1951.
17. KÖHLER, H. Vorratsermittlung in Buchenbeständen nach Stammdurchmesser und Stammabstand. Allgem. Forst- und Jagdztg. **123**, 69–74. (1952).
18. KUIPER, N. H. Differentiaal en integraalrekening. Wageningen, 1956.
19. KUIPER, N. H. Een opmerking over het aanpassen van functies aan een groot aantal waarnemingsuitkomsten. Statistica **8**, 1–6 (1954).
20. LAAR, A. VAN. Stand development and stem form of the Douglas fir (*Pseudotsuga taxifolia* BRITT.) in the Netherlands. Netherlands Journal of Agricultural Science **2**, 265–272 (1954).
21. LAER, W. VON & M. SPIECKER. Massenberechnungstafeln zur Ermittlung von Vorrat und Zuwachs von Waldbeständen. Frankfurt, 1951.
22. NEUBAUER, W. Zur Frage der kombinierten Verwendung von Massentafeln und Probestämmen bei Bestandesaufnahmen. Centrallblatt für das gesamte Forstwesen **58**, 305–320 (1932).
23. PRODAN, M. Messung der Waldbestände. Frankfurt a/M, 1951.
24. SPEIDEL, EMIL. Beiträge zu den Wuchsgesetzen des Hochwaldes. Tübingen, 1893.
25. STOFFELS, A. Die Genauigkeit der Bestimmung der Stammzahl pro Hektar durch Messung von Stammabständen Forstwissenschaft. Centr.bl. **74**, 211–218 (1955).
26. STOFFELS, A. De inhoudsbepaling van grovedennen-opstanden met behulp van standaardkrommen. Avec un résumé. Nederlandsch Boschbouw-Tijdschrift **25**, 29–42 (1953).
27. STOFFELS & J. VAN SOEST. Principiële vraagstukken bij proefperken, 3. Hoogteregressie. With a summary. Nederlandsch Boschbouw-Tijdschrift **25**, 190–199 (1953).
28. SPURR, STEPHEN H. Forest Inventory, New York, 1952.
29. TISCHENDORF, W. Grundflächen oder Massenmittelstamm? Allgemeine Forst- und Jagdzeitung **103**, 346–348 (1927).
30. TISCHENDORF, W. Der arithmethische und der zentrale Mittelstamm und andere Mittelstämme. Österreichisches Vierteljahresschrift für Forstwesen, **90**, 190–208 (1949).
31. TISCHENDORF, W. Mittelstammdimensionen, Forstwissenschaftliches Centrallblatt **47**, 787–798 (1925).
32. Verslag over de inhoudstabellen over de groveden. Nederlandsch Boschbouw Tijdschrift **27**, 290–293 (1955).
33. WECK, J. Untersuchung über Brauchbarkeit und Genauigkeit eines Verfahrens der Bestandesmessung unter Verwendung von Stammabständen. Forstarchiv **24**, 257–260 (1953).
34. WECK, J. Schnellverfahren zur Ermittlung von Holzvorrat und Zuwachs in Einzelbeständen des Gemässigten Zone. Union Forest Res. Organ. Cong. (Proc.) **11**, 865–872 (1953, publ. 1954).
35. WECK, J. Vorratsermittlung durch Abstandmessung. Zeitschrift für Weltforstwirtschaft **18**, 24–27 (1955).