MULTIPLE REGRESSION OF INCREMENT PERCENTAGE ON OTHER CHARACTERISTICS IN SCOTCH-PINE STANDS

KULLERVO KUUSELA AND PEKKA KILKKI

SELOSTE:

KASVUPROSENTIN JA MUIDEN METSIKKÖTUNNUSTEN VÄLINEN YHTEISKORRELAATIO MÄNNIKÖISSÄ

HELSINKI 1963

Suomalaisen Kirjallisuuden Kirjapaino Oy Helsinki 1963

Preface

This paper has been prepared as part of a research project entitled »Increment Forecast Methods for a Large Forest Area» under *United States Public Law N:o 480, 83rd Congress* (The Agricultural Trade Development and Assistance Act of 1954). The research funds have been granted to the *Institute of Forest Mensuration and Management of the University of Helsinki*; of the authors Dr. KULLERVO KUUSELA is the principal investigator of the project.

The investigation has been worked out through a close cooperation of the two authors. KUUSELA has prepared the overall plan and has directed the work. KILKKI has accomplished most of the calculations and has developed independently the mathematical form of many increment percentage functions. He has written the original manuscript in Finnish and presented it for a degree in forest mensuration at the *University of Helsinki*.

Professor AARNE NYYSSÖNEN, Director of the Institute of Forest Mensuration and Management, has given much valuable assistance to the authors in their work. For this the authors wish to express their sincere gratitude to him.

The authors acknowledge with pleasure the support ot the *Finnish Society of Forestry* in publishing this paper.

Helsinki, December 1, 1962.

Kullervo Kuusela *Pekka Kilkki*

Contents

Introduction

The increment of tree stands is usually described in increment and yield tables. In yield tables the increment is frequently given in terms of volume, although in Finland and the Scandinavian countries in particular it also appears in inventory data as a percentage. The loose dependency of increment percentage on site and even on tree species has been considered a marked advantage of this expression of increment (ILVESSALO 1942; 1942 a; HAGBERG 1943; PETRINI 1948; etc.). The increment percentage series obtainable from forest inventories have been given according to main tree species, age, density and site class.

Increment as the dependent variable in regression analysis has been the subject of concentrated study especially in the United States (SPURR 1952); it has been expressed as volume measure, either in cubic feet or in commercial units. The regression method has also been used in Europe. Most commonly volume increment has been used as a dependent variable (WECK 1948; MICHAJLOW 1952; SCHLETTER 1954; ERTELD 1957; VUOKILA 1960; etc.) and also as the increment percentage of mean diameter (PETTERSON 1955;), of basal area EIDE-LANGSAETER 1942), and of diameter and height of single trees (JONSSON 1961) have been used.

Although the use of the increment percentage is said to resemple »putting the cart before the horse» (CHAPMAN and MEYER 1949, p. 464), there is obviously no difference in principle between using the increment in terms of either volume or percentage if the percentage is calculated and used correctly (KUUSELA 1953). From a methodological point of view the percentage has an advantage in that since the relationship between it and age (or height) is more or less hyperbolic within a considerable age (or height) range. The corresponding relationship between the increment in volume measure and other stand characteristics is much more complex. Thus it is easier to begin the analysis of increment functions with the percentage than with the volume-measure increment.

The object of this study has been to discover some of the basic principles on which an increment for a large forest area might be forecast. Because the stands in a large forest area vary considerably in density and are subject to different kinds of treatment, the main interest falls on the stand characteristics which determine the increment percentage in such forest conditions as these.

The material and its preliminary treatment

This study comprised the first phase in a larger research project, and it was therefor considered unnecessary to collect new field material. On the other hand, most of the available sample plots and yield tables are not really suitable for the purposes of the study as defined in the introduction. Almost all yield tables, and the sample plots measured for the yield tables, describe the development of stands which are more or less fully stocked and have been under regular silvicultural treatment (so called normal stands). The sample plots measured in Finnish forest inventories cannot be used because they do not include an estimate for the increment of trees cut, or dying from natural causes, during the measurement period.

The most suitable material for the present study was that published on the sample plots of Scotch-pine stands (NYYSSÖNEN 1954). It has been used for this analysis and its main features are shown in Table 1 and Fig. 1.

The structure of these stands varies considerably. The under-stocked stands treated with silviculturally unsound cuttings comprise 26 % of the total. The great variation in the growing-stock volume is revealed in the correlation coefficient for age and volume, which is + 0.098.

For volume and increment, the total bole is taken from the stump to the top of the tree, in cubic meters, solid measure, excluding bark (ILVESSALO 1947). The mean annual increment in Table 1 describes the average potential capacity of the three forest site types to grow trees in fully stocked stands during a rotation of 100 years.

The increment (i) of each stand is measured for a 5 or 10 year period, and calculated as a mean annual increment over the period. It is the gross increment, including the increment of those trees living at the end of the measurement period and the increment of those trees cut during the period. In the increment percentage (p_i), the mean annual increment during the period is compared with the mean growing stock volume (v) during the same period:

$$p_i = 100\frac{i}{v}$$

Age (t) is the true biological age at the middle of the measurement period, estimated by borings.

Table 1. Sample plot stands used in this study.
Taulukko 1. Tutkimuksessa käytetty koeala-aineisto.

Treatment class *Käsittelyluokka*	Forest site type *Metsätyyppi* MT	VT	CT	Total number of stands *Metsiköiden lukumäärä yhteensä*
	number of stands *metsiköiden lukumäärä*			
A_1	27	50	19	96
A_2	3	6	5	14
B	6	12	12	30
C_1	5	20	13	38
C_2	1	8	3	12
Total *Yhteensä*	42	96	52	190
Range of age, years *Iän vaihtelualue, vuosia*	41—127	39—145	51—135	39—145
Range of volume, m³/ha *Kuutiomäärän vaihtelualue*	65—321	40—240	41—158	40—321
Annual mean increment, m³/ha *Vuotuinen keskikasvu* [1]	6.5	5.1	2.8	4.8

Description of the treatment classes (NYYSSÖNEN 1954, p. 177):

A. Stands treated with silvicultural thinnings.
  1. Repeatedly thinned stands.
  2. Heavily thinned stands.
B. Stands treated with indefinite cuttings.
C. Stands treated with selective cuttings.
  1. Stands subjected to prolonged selective cuttings.
  2. Selectively cut, rested stands.

Käsittelyluokkien kuvaus (NYYSSÖNEN 1954, s. 177):

A. Harventaen käsitellyt metsiköt.
  1. Toistuvasti harvennetut.
  2. Väljentäen harvennetut.
B. Epämääräisesti käsitellyt metsiköt.
C. Harsien käsitellyt metsiköt.
  1. Toistuvasti harsitut.
  2. Harsitut, lepoa saaneet metsiköt.

In those stands where the dominant crown storey existed at the time of measurement, the site index for the purpose of this examination is estimated from the mean increment of the dominant height. The site index (B) scale contains 10 classes and was prepared using data from the yield tables for naturally

[1] Vuotuinen keskikasvu on hyvin hoidetun täysitiheän metsikön keskikasvu 100 vuoden kiertoajalla.

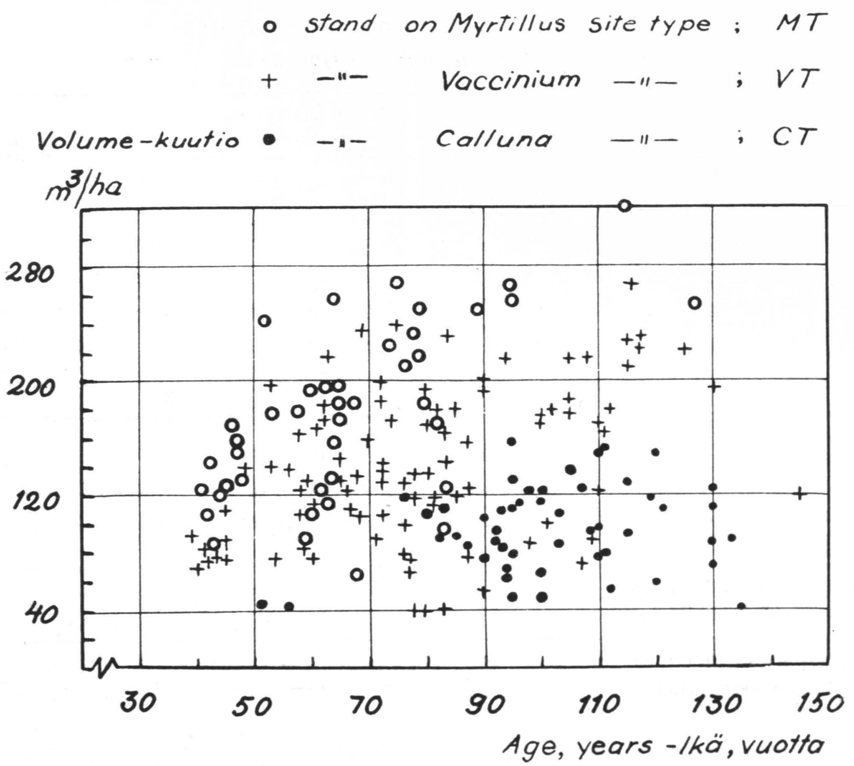

Fig. 1. Volume, excl. bark, of the sample plot stands against age.
Kuva 1. Koealametsiköiden iän mukainen kuoreton kuutiomäärä.

normal pine stands published by ILVESSALO (1920). Using the scale in Fig. 2 the site quality-class was determined for each stand where the dominant height could be estimated. For selectively cut stands, where the dominant crown storey had been destroyed, the site index was estimated on the basis of the forest site type.

The mean diameter at breast height, including bark (d_g), and the mean height (h_g) at the end of the measurement period could be used as independent variables. The former is the diameter of a tree which is the median of the basal area of the stand. It is approximately equal to the mean diameter weighted by the basal area.

Basal area including bark at the end of the measurement period (g_e) and the average basal area during the period (g_m) have also been used as independent variables.

As for mean diameter, mean height and basal area at the end of the measurement period, it should be remembered that they do not exactly meet the require-

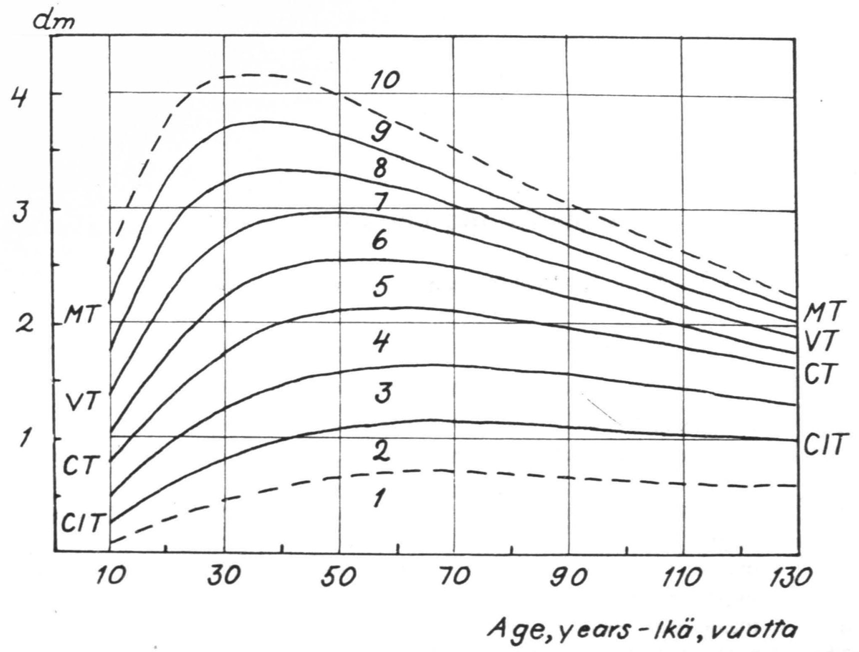

Fig. 2. Site quality-classes based on the mean annual increment of the dominant height.
Kuva 2. Valtapituuden keskikasvuun perustuvat kasvupaikan boniteettiluokat.

ments of an independent variable in regression analysis because they are measured at the end of a period, while the increment and the volume are estimates of the mean annual values during that period. Where a cut has led to a great decrease in the growing stock volume during the period, the stand characteristics at the end of the period are particularly unsuitable as independent variables. As a result, regression analysis with these characteristics can provide a foundation for only tentative conclusions.

Symbols and units of measure

The dependent variable

p_i	= percentage of annual increment

Independent variables

t (year)	= age of stand at time of observed increment percentage.
v (m^3/ha)	= average cubic volume of growing stock, excluding bark, during measurement period.
B	= site quality index, estimated in 10 classes based on mean increment of dominant height.
d_g (cm)	= diameter corresponding to mean basal area of stand at the end of increment-measurement period.
h_g (m)	= mean heigt, weighted by basal area, at end of measurement period.
g_e (m^2/ha)	= basal area of stand, including bark, at end of measurement period.
g_m »	= basal area of stand, including bark, in middle of period.

Some statistical characteristics

y	$= a + b x_1 + c x_2 + \ldots\ldots$ Equation of regression of y (dependent variable) on $x_1, x_2, \ldots\ldots$ (independent variables) where a, b, c . . . . . , constant and regression coefficients, are solved by the method of least squares.
r_{yx1}	= correlation coefficient between y and x_1.
R	= multiple correlation coefficient between dependent and independent variables in regression equation.
s_y	= standard deviation of dependent variable.
$s_{y \cdot 1}$	= standard error of estimate (or standard deviation from regression expressed by equation with x_1 as independent variable).
$s_{y \cdot 12 \ldots}$	= standard error of estimate calculated by regression equation with $x_1, x_2, \ldots\ldots$ as independent variables.

Research procedure

Preliminary examination

To determine permissible forms for the regression equations the material was examined by a combined graphic and analytic method. This aspect of the study consisted of several experiments, one of which is described here as an example.

The increment percentages were plotted on rectangular coordinates against age (Fig. 3). The regression of percentage on age is seen to be curvilinear. There are several curves which follow this kind of cluster of dots, e.g.:

$$y = \frac{a}{x}; \quad y = \frac{a}{x^b}; \quad y = \frac{a}{x + b}.$$

The constants were found by trial and error. The deviation of each percentage from a suitable curve (Fig. 3) was calculated and these remainder deviations were plotted against another independent variable, e.g. against the stand volume (Fig. 4), where there is a curvilinear regression between the remainder deviations and the volume. The course of the cluster of dots could be followed with another fixed curve and the remainder deviation calculated in this second step in the analysis. The procedure ended when the dots of the remained deviations formed a cluster without any regression on the remaining independent variables.

Using the above procedure the regression of the percentages was examined for all independent variables. The multiple regression was also examined, using forms of functions determined empirically. The regression equation:

$$p_i = 3.7853 + 0.8561 \frac{1}{(t\text{-}25)^{1.5}} + 0.9623 \frac{2}{v}$$

for example, was solved by the method of least squares for the stands of VT site type.

Another method of examining the multiple regression of the increment percentage on the stand characteristics is to combine two or more independent variables into one variable. E.g., a combination of age and volume might replace separate variables for age and volume. The best combination was the form $t\,(v + b)$ and the corresponding regression equation is

$$p_i = a \frac{1}{t\,(v + b)} + c$$

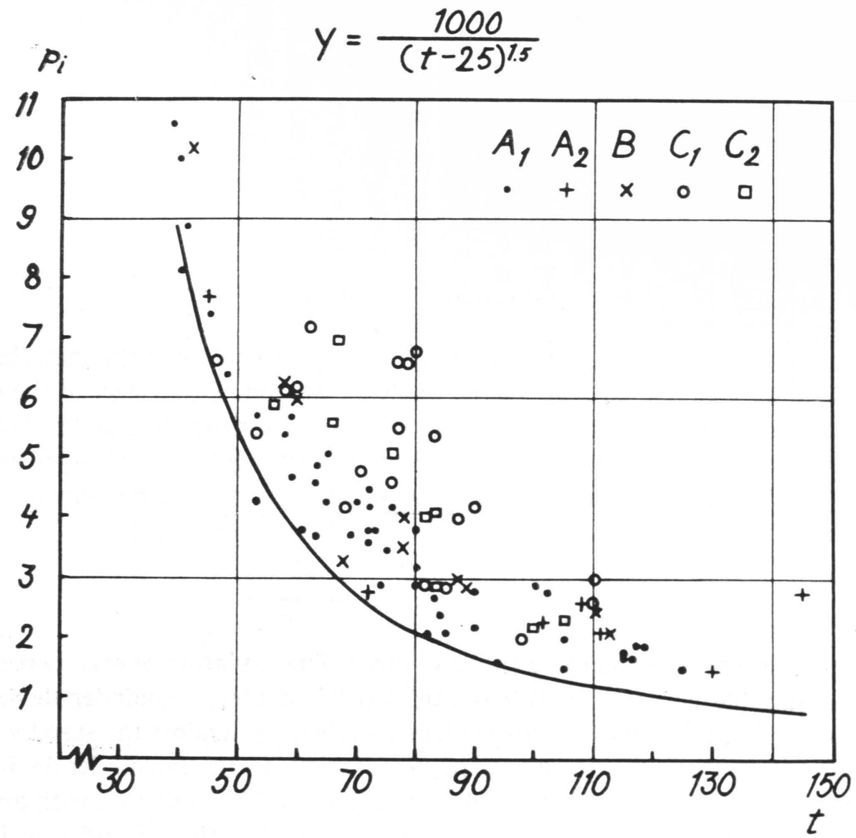

Fig. 3. Increment percentage against age, and a regression curve established by trial. (Explanation of signs A_1, A_2, etc. on p. 7).

Kuva 3. Kasvuprosentti iän funktiona ja kokeillen saatu regressiokäyrä. (Merkkien A_1, A_2 jne. selitykset ovat sivulla 7).

Of the constants, *b* had to be solved by trial and *a* and *c* by the method of least squares. A suitable value for *b* is about 100.

Fig. 5 shows increment percentage estimates as functions of $t\,(v + 100)$, calculated by a regression equation as described above. Compared with Fig. 3 the decrease in deviation is quite marked. In the total sample of 190 stands the standard deviation of the increment percentages is 61 per cent of the mean percentage (i.e. the coefficient of variation is 0.61). The error in an estimate which uses the combination of age and volume as the independent variable is 21 percent of the mean percentage. In other words, about $^2/_3$ of the variation of the observed percentages could be explained.

Although the regression equations based on combinations of independent variables make workable solutions possible, the determination of the constants by trial is, from the methodological point of view, a weakness. Even more serious

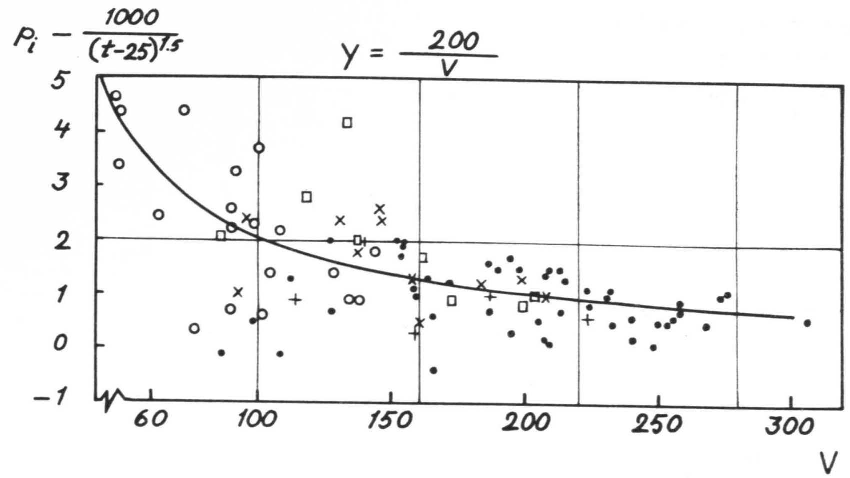

Fig. 4. Remainder deviations from Fig. 3 (observed percentage minus estimate given by the regression curve) against volume, and a regression curve established by trial.

Kuva 4. Jäännöspoikkeamat kuvasta 3 (havaittu prosentti miinus regressiokäyrän arvo) kuutiomäärän funktiona. Regressiokäyrä on saatu kokeillen.

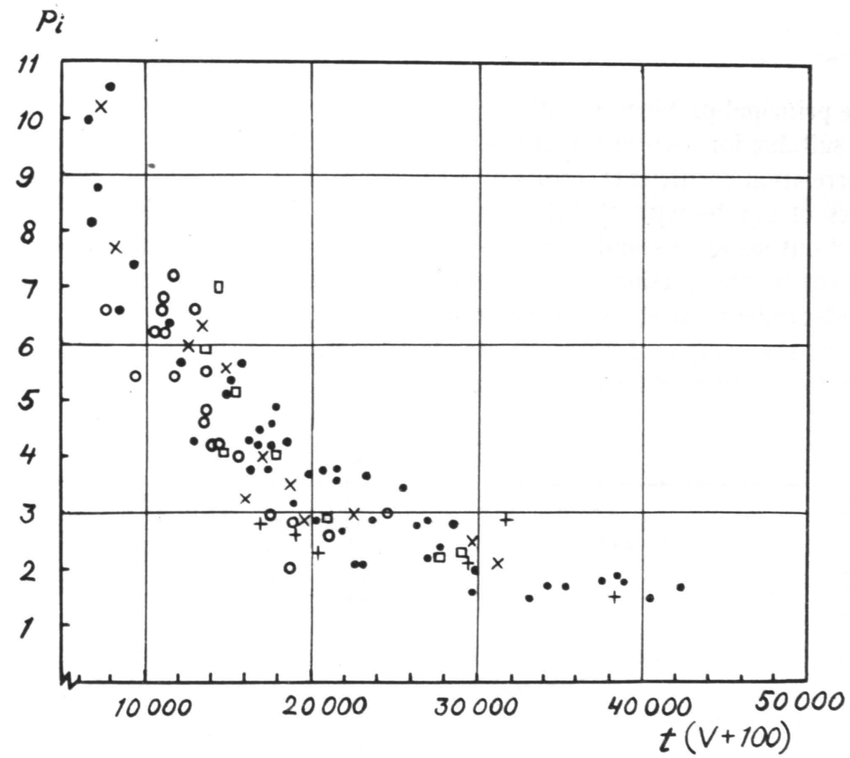

Fig. 5. Increment percentage from Fig. 3 against age and volume as a combined independent variable.

Kuva 5. Kuvassa 3 esitetyt kasvuprosentit ikä- ja kuutiomääräyhdistelmän funktiona.

is the fact that a combination of more than two independent variables is much more difficult to find than a combination of two variables. Because of this, analysis was continued on another basis.

Suitable forms of increment percentage functions

Regression analysis of many variables with curvilinear correlation between them can be based successfully on two auxiliary functions (compare JEFFERS 1960, SIRÉN 1961 and ASSMAN 1961). One is the hyperbolic function

$$y = a\ x^b$$

and another the exponential function

$$y = a\ 10^{bx}$$

In logarithmic form they are

$$\log y = \log a + b \log x$$
$$\log y = \log a + b\ x$$

The principal problem is to find which of these two auxiliary functions is the more suitable for each of the independent variables. Suitability can be assessed by correlation coefficients, errors of estimates and by comparisons between the courses of graphs with the clusters of observed dots. Preliminary tests were carried out using 19 stands chosen systematically from the total sample of 190. To illustrate this section of the analysis, experiments with age and volume as the independent variables are shown in Figures 6, 7, and 8. Although the variation in the sample of these 19 stands is comparatively great the following conclusions may be drawn:

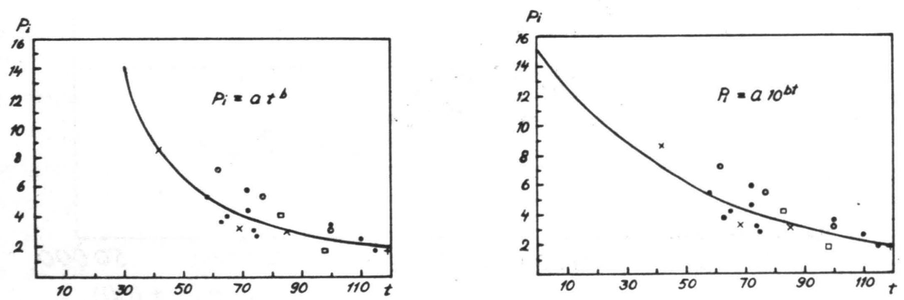

Fig. 6. Increment perscentage in hyperbolic and exponential relationship against age.
Kuva 6. Kasvuprosentin hyperbolinen ja eksponentiaalinen riippuvuus iästä.

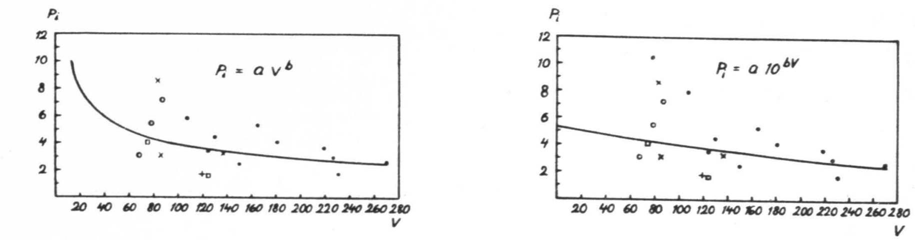

Fig. 7. Increment percentage in hyperbolic and exponential relationship against volume.
Kuva 7. Kasvuprosentin hyperbolinen ja eksponentiaalinen riippuvuus kuutiomäärästä.

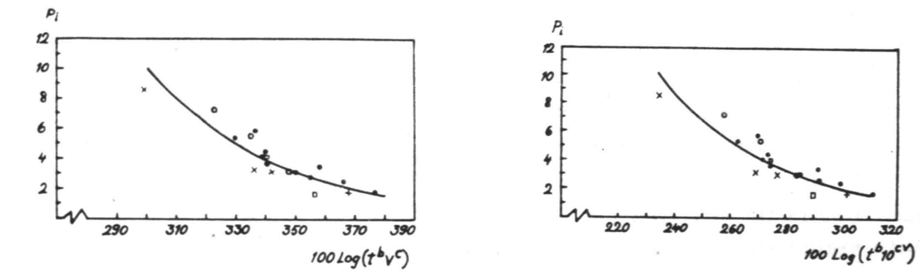

Fig. 8. Increment percentage against the combination of age and volume, where the relationship between percentage and age is hyperbolic and between percentage and volume hyperbolic or exponential.
Kuva 8. Kasvuprosentti iän ja kuutiomäärän yhdistelmän funktiona siten, että riippuvuus iästä on hyperbolinen ja kuutiomäärästä hyperbolinen tai eksponentiaalinen.

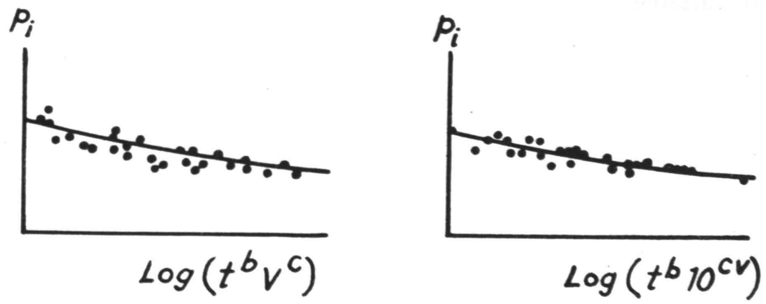

Fig. 9. Course of regression curve which is, in regard to volume, either hyperbolic or exponential, in a group of increment percentages for old and heavily stocked stands.
Kuva 9. Regressiokäyrä, jossa kasvuprosentin ja kuutiomäärän välinen riippuvuus on joko hyperbolinen tai eksponentiaalinen, vanhojen ja suurikuutioisten metsiköiden kasvuprosenttien tasoittajana.

Where age is the independent variable, the hyperbolic function is more flexible than the exponential function and more closely follows the percentage of young stands (Fig. 5). Increment percentage as a function of age alone overestimates the increment of fully stocked stands, especially on good sites. Percentage as a function of volume alone overestimates the increment of old stands and of stands on poor sites, and underestimates the increment of young stands and of stands on good sites. Both age an volume as simultaneous independent variables reduce these systematic errors. In graphic representation the site and treatment do not explain the remainder of the variation.

For age, the hyperbolic function is obviously better than the exponential function. For volume the choice is more difficult and the graphic representations in Figures 6, 7 and 8 do not solve the problem.

Fig. 9 shows the course of the hyperbolic and exponential function among the cluster of increment percentages for old and fully stocked stands, calculated from pine yield-tables (Nyyssönen 1958) which include fully-stocked stands, preparatory and shelterwood stages within a rotation varying from 60 to 130 years. For volume, the hyperbolic function systematically overestimates the increment of fully-stocked old stands while the exponential function more closely follows the percentages.

With young stands, the exponential function is not theoretically the correct one because it does not approach the y-axis asymptotically (Fig. 7). This, however, is not a very serious weakness because the increment of young stands is more of an approximation and less important than the increment of older and fully-stocked stands. In addition, the increment of the young stands is more markedly dependent on age than on volume.

From the above argument, and from other experiments, it may be stated that the hyperbolic function proved to be better for age and the site index (Table 2, p. 17) while the exponential function proved more suitable for the other independent variables.

Table 2. Correlation and multiple correlation coefficients.

Taulukko 2. Korrelaatio- ja yhteiskorrelaatiokertoimet.

	Log p_i	h_g	Log h_g	B	Log B	Log t	v	d_g	g_e	g_m
I										
Log p_i	1.00000	— 0.68707	— 0.68606	+ 0.21222	+ 0.21551	— 0.84213	— 0.44566	— 0.71863	— 0.03929	— 0.21770
h_g	—	1.00000	+ 0.98691	+ 0.45414	+ 0.45347	+ 0.45034	+ 0.78775	+ 0.87803	+ 0.36568	+ 0.52062
Log h_g	—	—	1.00000	+ 0.43588	+ 0.44017	+ 0.46149	+ 0.75206	+ 0.88030	+ 0.33095	+ 0.48979
B	—	—	—	1.00000	+ 0.99258	— 0.48628	+ 0.52535	+ 0.24214	+ 0.50055	+ 0.47632
Log B	—	—	—	—	1.00000	— 0.48120	+ 0.51094	+ 0.24540	+ 0.48225	+ 0.45983
Log t	—	—	—	—	—	1.00000	+ 0.09844	+ 0.62125	— 0.31419	— 0.13536
v	—	—	—	—	—	—	1.00000	+ 0.48913	+ 0.82126	+ 0.92832
d_g	—	—	—	—	—	—	—	1.00000	+ 0.02330	+ 0.17133
g_e	—	—	—	—	—	—	—	—	1.00000	+ 0.94092
g_m	—	—	—	—	—	—	—	—	—	1.00000
II — Log p_i										
h_g	—	—	—	0.90460	0.90653	0.90997	0.70437	0.72813	0.72382	0.70636
B	—	—	—	—	—	0.87187	0.68832	0.82152	0.27072	0.42010
Log B	—	—	—	—	—	0.86949	0.68151	0.82451	0.27050	0.41678
Log t	—	—	—	—	—	—	0.91764	0.87831	0.90092	0.90623
v	—	—	—	—	—	—	—	0.72669	0.72562	0.67440
d_g	—	—	—	—	—	—	—	—	0.71898	0.72500
III$_1$ — Log p_i h_g										
B	—	—	—	—	—	—	0.90497	0.90501	—	0.90489
v	—	—	—	—	—	—	—	0.72826	—	—
d_g	—	—	—	—	—	—	—	—	—	0.72815
III$_2$ — Log p_i Log t										
h_g	—	—	—	—	—	—	0.92051	0.91417	—	—
B	—	—	—	—	—	—	0.91785	0.87946	—	—
v	—	—	—	—	0.91790	—	—	0.91954	—	—

Table 3. Constants and regression coefficients with their standard errors for the regression equations (Log 10 p_i = independent variable), multiple correlation coefficients and the proportional standard errors of the estimate from step by step analyses.

Taulukko 3. Regressiyhtälöiden vakiot ja regressiokertoimet keskivirheineen, yhteiskorrelaatiokertoimet ja arvion suhteellinen keskivirhe regressioon nähden vaiheittaisessa analyysissä (Log 10 p_i = selitettävä muuttaja).

Constant a	Regression coefficients (b, c, d . . . .) and their standard errors					R	$100 \frac{s_{y \cdot 123 \ldots}}{s_y}$
	Log t	v	h_g	d_g	Log B		
3.98181	— 1.28468	—	—	—	—	— 0.841199	54.1
± 0.11400	± 0.06000	—	—	—	—	—	—
4.05600	— 1.22967	— 0.001309	—	—	—	0.916714	40.0
± 0.08446	± 0.04455	± 0.000104	—	—	—	—	—
3.99359	— 1.14432	— 0.000910	— 0.008500	—	—	0.919200	39.4
± 0.08679	± 0.05532	± 0.000188	± 0.003350	—	—	—	—
4.00475	— 1.14929	— 0.000875	— 0.010014	— 0.000936	—	0.918640	39.5
± 0.09697	± 0.05866	± 0.000233	± 0.066692	± 0.003577	—	—	—
3.15030	— 0.75234	— 0.000703	— 0.024433	— 0.000338	0.46851	0.926639	37.6
± 0.21662	± 0.10680	± 0.000225	± 0.007175	± 0.003416	± 0.10746	—	—

Constant	Log t	v	Log B	—	—	R	$100 \frac{s_{y \cdot 123 \ldots}}{s_y}$
3.98181	— 1.28468	—	—	—	—	— 0.841199	54.1
± 0.11400	± 0.06000	—	—	—	—	—	—
4.05600	— 1.22967	— 0.001309	—	—	—	0.916714	40.0
± 0.08446	± 0.04455	± 0.000104	—	—	—	—	—
3.97088	— 1.20304	— 0.001375	0.05553	—	—	0.916564	40.3
± 0.14113	± 0.05691	± 0.000136	± 0.07372	—	—	—	—

Constant	h_g	v	d_g	Log B	—	R	$100 \frac{s_{y \cdot 123 \ldots}}{s_y}$
2.25753	— 0.039295	—	—	—	—	— 0.685017	72.8
± 0.05587	± 0.003031	—	—	—	—	—	—
2.34002	— 0.050644	— 0.0009002	—	—	—	0.700535	71.4
± 0.06129	± 0.004819	± 0.0003011	—	—	—	—	—
2.24736	— 0.011145	— 0.0001143	— 0.021783	—	—	0.723050	69.1
± 0.06445	± 0.011700	± 0.0004012	± 0.005917	—	—	—	—
1.65884	— 0.044486	— 0.0003314	— 0.006161	1.11396	—	0.905937	42.3
± 0.05174	± 0.007417	± 0.0002462	± 0.003734	± 0.06326	—	—	—

Example demonstrating the use of Table 3.

Measured characteristics of a pine stand:

$t = 65$ years $\qquad \log t = 1.8529$
$v = 162$ m³/ha
$h_g = 18.6$ m
$d_g = 19.9$ cm
$B = 7$ $\qquad \log B = 0.8451$

1.
Estimate by regression equation:

$$\log 10\, p_i = 4.05600 - 1.22967 \log t - 0.001309\, v$$

$$\begin{array}{r} -\ 2.22927 \\ -\ 0.21206 \\ \hline 1.61467 \end{array}$$

$10\, p_i = 41.2$
$p_i = 4.12$ $\qquad i = 6.67$ m³/ha

2.
Estimate by regression equation:

$$\log 10\, p_i = 1.65884 - 0.044486\, h_g - 0.0003314\, v - 0.006161\, d_g + 1.11396 \log B$$

$$\begin{array}{r} -\ 0.82744 \\ -\ 0.05369 \\ -\ 0.12260 \\ +\ 0.94141 \\ \hline 1.59652 \end{array}$$

$10\, p_i = 39.5$
$p_i = 3.95$ $\qquad i = 6.40$ m³/ha

The measured annual increment of the stand is 7.0 m³/ha and the estimates are correspondingly 5 % and 9 % below this.

Final analysis

The final analysis, based on the experience gained in the preliminary examination and on the tested forms of the functions, was performed in two stages. In the first the importance of the independent variables and their combinations was examined by the correlation coefficients between the variables, and by the multiple correlation coefficients between the dependent variable and groups of independent variables.

The number and combination of the independent variables and their sequence for promising regression equations were decided on the basis of the correlation tables, and three regression equations were calculated by the step-by-step method of Törnquist (1957). This method provides, after each additional independent variable, the constant and the regression coefficients, the standard errors of the constant and regression coefficients, the sum of the remaining squared deviations from the regression, the multiple regression coefficient, and the deviation of each original observation from the corresponding estimate produced by the regression equation. The importance of each additional independent variable can be determined on the basis of the corresponding multiple correlation coefficient and the standard error of the estimate. The regression equation can be ended after any independent variable in the sequence if the next variables do not reduce the standard error of the estimate.

In the analysis, log p_i was the dependent variable and h_g, log h_g, B, log B, log t, v, d_g, g_e and g_m (p. 10) were used as independent variables. The table (Part I of Table 2) of the correlation coefficients between these variables was calculated by electronic computor. The multiple correlation coefficients between the dependent variable and two of the independent variables (Part II in Table 2), between the dependent variable, mean height (h_g) and two of the most promising other independent variables (Part III_1 in Table 2) and between the dependent variable, age (log t) and two of the most promising other independent variables (Part III_2 in Table 2) were calculated by a hand computor.

The reasoning from the correlation coefficients (from Part I to II, III_1 and III_2) followed two lines. As in the preliminary examination, the regression of the increment percentage on age is very marked ($r = -0.84213$). The first objective was to find the additional independent variables of significant importance for a regression equation with age as the primary independent variable. The multiple correlation coefficient between log p_i, log t and v is 0.91764 (Part II). Because the mean diameter (d_g), and in particular the combination of mean height (h_g) and site (log B), showed a tendency to increase the multiple correlation coefficient, the following equation was chosen for the step by step analysis:

$$\log p_i = a + b \log t + c\, v + d\, h_g + e\, d_g + f \log B$$

Because mean height and mean diameter are not quite faultless as independent variables (p. 9) the following equation was also calculated:

$$\log p_i = a + b \log t + c\, v + d \log B$$

Another objective was a regression equation which did not use age as an independent variable, because determination of age requires borings and age is hard to define in stands of uneven age. On the basis of the correlation coefficient table the following equation was calculated:

$$\log p_i = a + b\, h_g + c\, v + d\, d_g + e \log B$$

A sequence of d_g, v, h_g and log B might have been more logical on the basis of the correlation coefficients (Table 2), but the mean height was wanted as the primary independent variable. Because the site index is difficult to determine objectively it was placed at the end of the sequence.

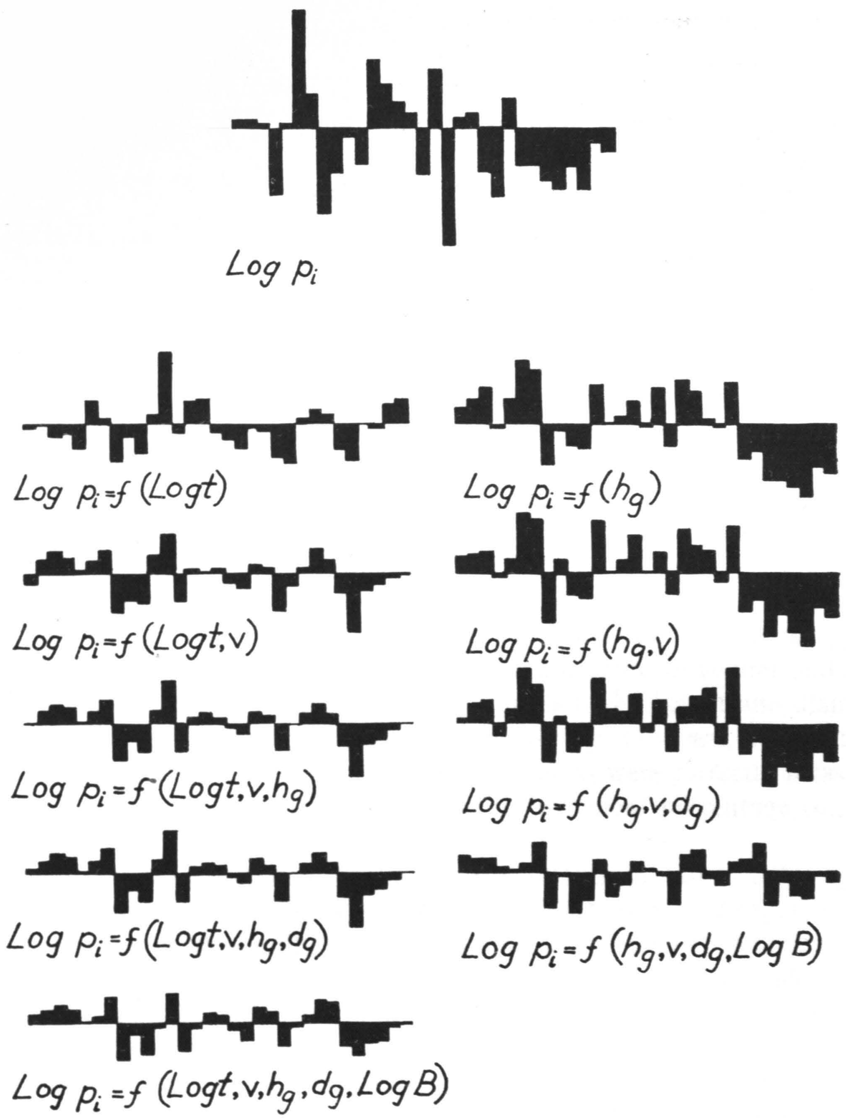

Fig. 10. Deviations of individual log p_i values and computed values of it from the mean log p_i in a systematic sub-sample taken from the total sample of 190 stands.

Kuva 10. Log p_i:n ja sen arvion vaihtelu log p_i:n keskiarvoon nähden systemaattisesti valitussa osanäytteessä, joka on otettu 190 koealametsikön kokonaisnäytteestä.

Constants and regression coefficients and their standard errors for the regression equations described above, multiple correlation coefficients and the proportional standard errors of the estimate calculated by an electronic computor are presented in Table 3.

The first equation might have been calculated without h_g (or without d_g) because of the high positive correlation between height and diameter. For the same reason d_g could have been omitted from the third equation. With the combination of log t and v in the second equation log B is insignificant and may be omitted.

The multiple correlation coefficients in Table 3 are not exactly equal to the corresponding coefficients in Table 2 because the method of the step-by-step analysis is not identical with the method used in calculating by hand computor.

Fig. 10 shows how the variation of the individual log p_i decreases in a systematic sub-sample taken from the total sample when a new independent variable is added to the regression equation. The decreasing variation corresponds to the decreasing error of estimate in Table 3. In the lowest figure of each sequence there is a remainder variation which could not be explained. It is caused, at least in part, by experimental errors due to defects in measurement and the limited number of increment sample trees.

Importance of the characteristics as independent variables

In this sample of pine stands age and volume are the most important independent variables with a marked significance. The function log $p_i = f$ (log t, v) explains 60 per cent of the variation of log p_i. The site index is non-significant as on additional variable to this combination. Age and volume determine the increment of pine stands within a comparatively wide range of site quality.

If mean height and mean diameter (cf. p. 9) had been measured at the beginning of the increment measurement-period, or were determined exactly in the middle of the increment measurement-period in cases where there had been cutting during the period, the best result could obviously be attained either by function log $p_i = f$ (log t, v, h_g, log B) or by function log $p_i = f$ (log t, v, d_g, log B). Because the bole area (or the area of living cambium layer) is an approximate function of volume and mean diameter (or a function of volume and mean height, because of the great positive correlation between height and diameter) the volume of the increment is determined mainly by stand age, bole area and site quality (compare LEXEN 1943). If these variables were correctly measured, obviusly 70—80 per cent of the variation in increment percentage could be explained.

If age is omitted a comparatively good result may be obtained either by the function log $p_i = f$ (h_g, v, log B) or by the function log $p_i = f$ (d_g, v, log B). Thus bole area and site quality determine a considerable part of increment.

For many practical purposes age and volume may be used as independent variables without the site index. If age is omitted from the regression equation, however, the site index becomes indispensable.

The effect of age on the increment percentage is most significant. Volume comes second in importance. Only the correlation between site index and increment percentage is positive. The negative correlation between increment percentage and mean height or mean diameter is also of importance because it shows that if age and volume are constant, a stand comprising a mixture of large and small trees gives a slightly higher increment than does a stand consisting only of large trees.

The standard error in the increment percentage cannot be inferred directly from the standard error in the estimate of log p_i (JEFFERS 1960, p. 127). When a preliminary function with age and volume as combined independent variable was used (p. 12), the standard error of the estimate was $\pm$ 21 % of the mean percentage while the standard deviation of the observed percentages was $\pm$ 61 %. Obviously the best regression equations in the final analysis give better results than this preliminary function. E.g., the standard error of the estimate of the increment percentage given by the function $\log p_i = f (\log t, v)$ is 19.0 % of the mean percentage.

In some other experiments, with the increment in volume measure as the dependent variable, about 1/3 of the variation could be explained (e.g. SPURR 1952; VUOKILA 1960). The same ratio in the experiments of this study is less than 2/3. Although these ratios are greatly dependent on the variation in sample, especially on the range of age, it seems that regression analysis can be based more advantageously on the increment percentage than on the increment in volume measure. In any case the increment percentage functions are less complicated. Because volume (growing capital) is a very important independent variable and one that is usually known where increment is wanted, the estimation of the inceement can be based on the percentage if this facilitates the calculation of regression equations.

Mensurational aspects of the independent variables, and conclusions

If increment is calculated as a function of the independent variables measured at the beginning of the measurement period, the mean height, mean diameter and volume can be estimated easily and accurately. Determination of age requires borings, but in mainly coniferous stands of more or less even age the determination of age is neither particularly difficult nor laborious. On the other hand, if the stands consist of a variety of tree species, of broad-leaved species or if they are uneven in age, it would be wise to omit age as an independent variable.

Site quality is at present the most difficult variable to be determined objectively and expressed numerically. In conditions where the original dominant trees exist in all stands the dominant height or the mean increment of the dominant height against age is a workable site index. Unfortunately very many stands are treated by selective cuttings or other unsilvicultural methods, the trees grow for at least some time dominated or as undergrowth, etc. In Finnish forest conditions the forest site types may be used as site-quality classes but their exact determination and the numerical expression of their value is an unsolved problem thus far.

Because the increment during a measurement or forecast period is more a function of the average variables during the period than of the variables at the beginning of the period, and because cuttings very markedly affect the variables during a 5—10 year period, the determination of the average mean height and diameter during the period is difficult. The problem is not insoluble in cases of stands treated silviculturally, but in the very circumstances when the increment functions are most urgently needed the stands are treated by cuttings varying considerably in intensity and quality. For a cutting budget over a large forest area it is particularly difficult to predict the average mean height or mean diameter over any budget period.

In present Finnish forest conditions age and volume can be most easily determined for increment forecast purposes. The average age during a forecast period is roughly the age at the beginning of the period plus half the number of years in the period. The average volume can be calculated as the arithmetic mean of the volumes at the beginning and at the end of the period. In planned forestry the volume at the beginning of an increment forecast-period is estimated by an inventory and the volume at the end of the period is the optional quantity needed for the sustained production of timber. The average volume described above can thus be calculated.

For forest mensuration and for the purposes of forest management, the most useful independent variables in Finnish forest conditions are age and volume. In conjunction with these, the site index is dispensable if the variation in site is not great. The importance of mean height and mean diameter increases with the number of silviculturally-treated stands. In a forest of silviculturally-treated stands the increment function can possibly be based on the mean height (or mean diameter), volume and site index as independent variables.

Applications of the increment percentage functions

The increment percentage function

$$p_i = a\ t^b \cdot 10^{cv}$$

can be expressed as an increment function where the increment in volume measure is:

$$i_v = a\ t^b \cdot 10^{cv} \cdot \frac{v}{100}$$

For practical purposes the annual increment in cubic measure can be calculated and presented in tabular form as a function of age and volume class (Table 4). The table may be used directly in the estimation of the increment, e.g. for cutting budget purposes:

The initial volume of a 55-year-old pine stand is 130 m³/ha, the desirable volume at the end of the 10 years' management plan period is 170 m³/ha, and a figure for the allowable cut is wanted. The average age of the stand during the period is 60 years and the average growing stock volume (130 + 170) : 2 = 150 m³/ha. The average annual increment from Table 5 is 7.1 m³/ha. The allowable drain (D), including cut and mortality, is

$$D = 130 + 10 \cdot 7.1 - 170 = 31 \text{ m}^3/\text{ha}$$

One form of the increment percentage is the compound interest factor, $1.0p^n = K/k$, where p is the increment percentage, k the initial volume, K the final volume and n the number of years in the calculation period (e.g. KUUSELA and NYYSSÖNEN 1962). If deducted drain occurs during the period and the drain is supposed to be deducted from the live growing stock in the middle of the n-year period ($n = 2\,m$), the increment drain process has the form

$$(1.0p^m\, k - D)\ 1.0p^m = K$$

Initial volume, final volume, compound interest factor and drain can be solved from the equation when the remaining three constituents are known. If $1.0p^m$ is solved and calculated from sample-plot or yield-table data, the remainder ($1.0p^m - 1$) can be studied, by regression analysis, as the dependent variable. Once there is a regression equation for ($1.0p^m - 1$) the estimation of the allowable drain can be calculated with the formula

Table 4. Increment (m³/ha, excluding bark) as a function of age (t = years) and volume (v = m³/ha, excluding bark).

Taulukko 4. Kasvu (m³/ha kuoretonta puuta) iän (t = vuosia) ja kuutiomäärän (v = m³/ha, kuoretonta puuta) funktiona.

v \ t	30	40	50	60	70	80	90	100	110	120	130
30	4.8	3.3	2.5	2.0	1.7	1.4	1.2	1.1	1.0	0.9	0.8
40	6.2	4.3	3.3	2.6	2.2	1.8	1.6	1.4	1.3	1.1	1.0
50	7.5	5.3	4.0	3.2	2.6	2.2	1.9	1.7	1.5	1.4	1.2
60	8.7	6.1	4.6	3.7	3.1	2.6	2.3	2.0	1.8	1.6	1.4
70	9.9	6.9	5.2	4.2	3.5	2.9	2.6	2.2	2.0	1.8	1.6
80	10.9	7.7	5.8	4.7	3.8	3.3	2.8	2.5	2.2	2.0	1.8
90	11.9	8.4	6.4	5.1	4.2	3.6	3.1	2.7	2.4	2.2	2.0
100	12.8	9.0	6.9	5.5	4.5	3.8	3.3	2.9	2.6	2.3	2.1
110	—	9.6	7.3	5.8	4.8	4.1	3.6	3.1	2.8	2.5	2.3
120	—	10.2	7.7	6.2	5.1	4.3	3.8	3.3	2.9	2.6	2.4
130	—	10.7	8.1	6.5	5.4	4.6	4.0	3.5	3.1	2.8	2.5
140	—	11.2	8.5	6.8	5.6	4.8	4.1	3.6	3.2	2.9	2.6
150	—	11.6	8.8	7.1	5.9	5.0	4.3	3.8	3.4	3.0	2.7
160	—	—	9.1	7.3	6.0	5.1	4.4	3.9	3.5	3.1	2.8
170	—	—	9.4	7.5	6.2	5.3	4.6	4.0	3.6	3.2	2.9
180	—	—	9.7	7.8	6.4	5.4	4.7	4.1	3.7	3.3	3.0
190	—	—	9.9	7.9	6.6	5.6	4.8	4.2	3.8	3.4	3.1
200	—	—	10.1	8.1	6.7	5.7	4.9	4.3	3.8	3.5	3.1
210	—	—	10.3	8.3	6.8	5.8	5.0	4.4	3.9	3.5	3.2
220	—	—	10.5	8.4	6.9	5.9	5.1	4.5	4.0	3.6	3.2
230	—	—	10.7	8.5	7.0	6.0	5.2	4.5	4.1	3.6	3.2
240	—	—	10.8	8.6	7.1	6.0	5.2	4.6	4.1	3.7	3.3
250	—	—	10.9	8.7	7.2	6.1	5.3	4.7	4.1	3.7	3.4
270	—	—	—	8.9	7.3	6.2	5.4	4.7	4.2	3.8	3.4
290	—	—	—	9.0	7.4	6.3	5.5	4.8	4.3	3.8	3.5
310	—	—	—	9.0	7.5	6.3	5.5	4.8	4.3	3.8	3.5
330	—	—	—	9.0	7.5	6.4	5.5	4.8	4.3	3.9	3.5

$$D = \text{l.op}^m \cdot k - \frac{K}{\text{l.op}^m}$$

This calculation procedure is applied in a cutting-budget method for a desirable growing stock prepared by Kuusela and Nyyssönen (1962).

If the first derivative of the increment function with regard to volume (v) equals zero, the increment is at a maximum when

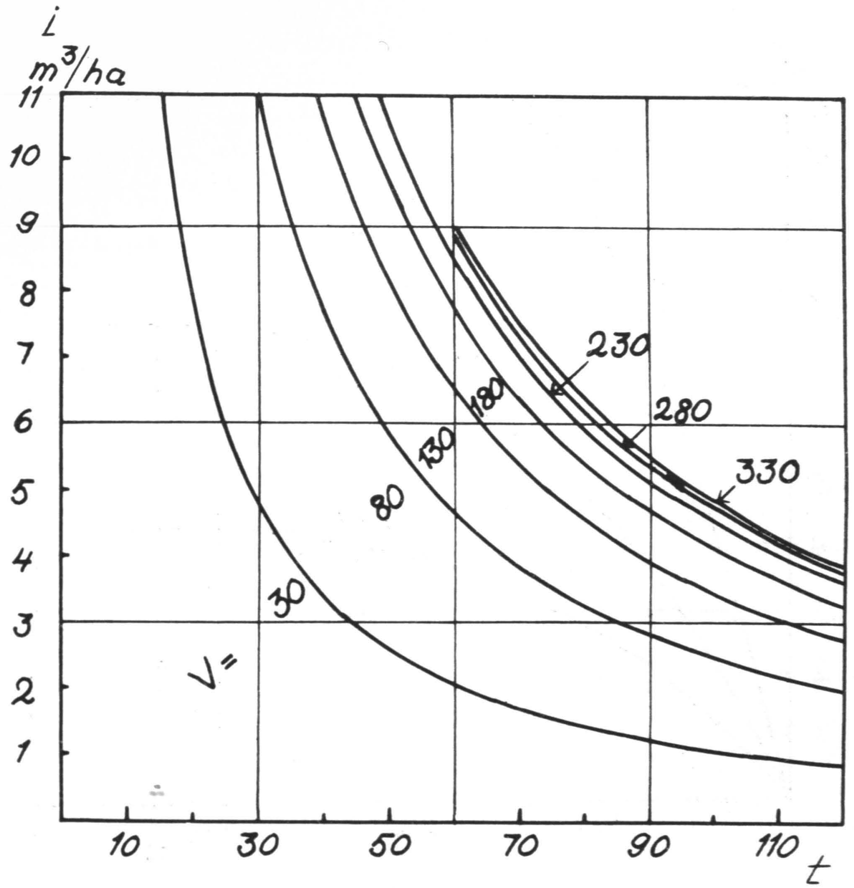

Fig. 11. Annual increment, excl. bark, as a function of age.

Kuva 11. Vuotuinen kuoreton kasvu iän funktiona.

$$v = -\frac{1}{c \ln 10}$$

Calculated from the regression equation in which age and volume are the independent variables gives $v = 332$ m³/ha. In other words, the growing stock volume of 332 m³/ha gives the greatest possible increment.

The effect of age and volume on increment is illustrated in Figures 11 and 12. The curves in Fig. 12 show particularly clearly that within a large volume range around the value of 332 m³/ha the deviation of the increment from the maximum is very small. Thus the decrease in volume from 332 m³ to 232 m³ reduces the increment by about 6 %. Because the reduction in volume by thinning from below increases the proportional increment of large size timber, the most eco-

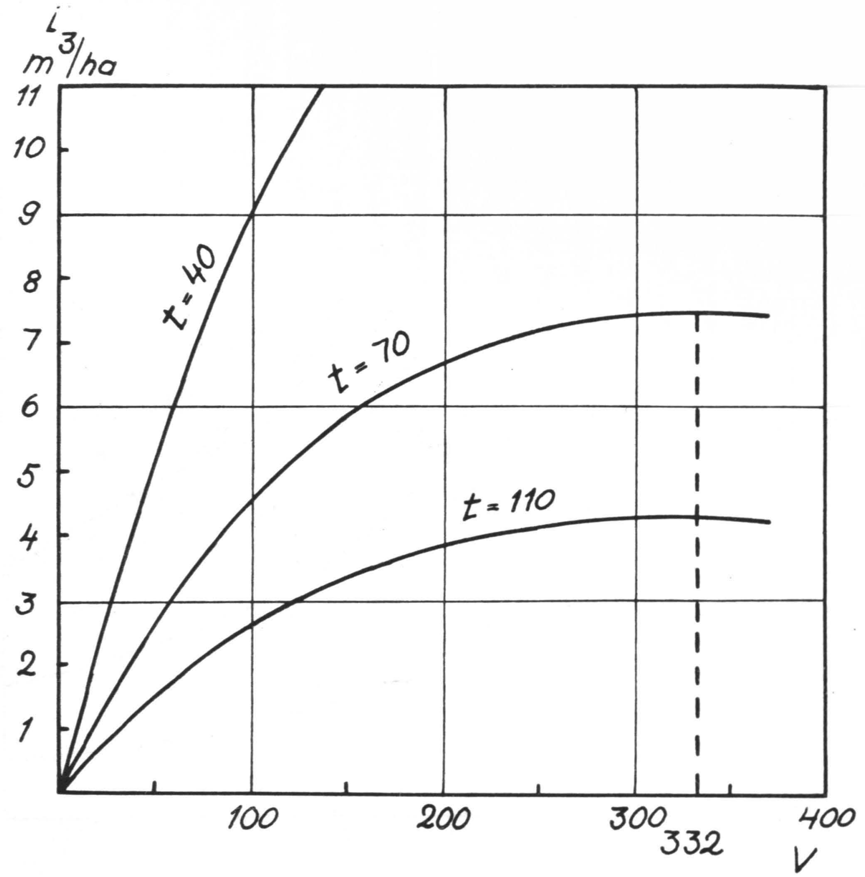

Fig. 12. Annual increment, excl. bark, as a function of volume.
Kuva 12. Vuotuinen kuoreton kasvu kuutiomäärän funktiona.

nomic volume is obviously between 200—250 m³/ha. A further conclusion is that an understocked condition in a stand is more severe at an early age than later on from the standpoint of wood production (compare MÖLLER 1945; 1954; etc).

Although conclusions about a biological process on the basis of regression equations should be made with caution, the above examples show what a powerful tool regression analysis is in yield studies.

Discussion

Increment functions are of great importance in the increment forecast for a cutting budget. Because 60—80 per cent of the variation in the increment percentage can be explained by stand characteristics in circumstances where the stand age ranges from 40 to 130 years and the volume varies with a coefficient of variation 0.6—0.7, regression equations for increment percentage may be based on a number of sample plots smaller than the number of plots needed for a growing stock inventory in the same conditions. And because of this comparatively small number of increment sample plots the necessary measurements and estimation of the cut and mortality during the measurement period can be carried out more accurately than similar measurements and estimates on ordinary inventory plots.

The smaller number of increment sample plots makes it possible to develop measurement techniques. In Finland, volume increment is based on the diameter increment at breast height, height increment and empirical data concerning the average change of tree form. Obviously much of the inexplicable variation in increment percentage is caused by the variation of diameter increment near the stump buttress in stands treated with cuttings. If the diameter increment could be bored higher on the bole the explainable variation might be expected to increase, and the number of sample plots necessary for required precision could be reduced still further.

Once reliable and accurate increment percentage functions are available they can be used easily and cheaply in increment forecasts.

In Finnish forest conditions the most suitable and efficient independent variables (besides tree species) are age, bole area (which is approximately a function of volume and mean diameter) and site. The site index is not very significant within a large variation in site quality. Further experiments are needed to measure and express the mean height and diameter in such a way that these variables are in correlation with the average growing conditions during an increment-measurement period in stands treated with cuttings, because it seems possible to base a regression equation on characteristics exclusive of age.

Where age and bole area are taken as independent variables the unimportance of site may seem a paradox; obviously site quality is of primary importance to increment. Yet soon after the seedlings and saplings have fully occupied the site the age and bole area reach a state of mutual equilibrium which is a func-

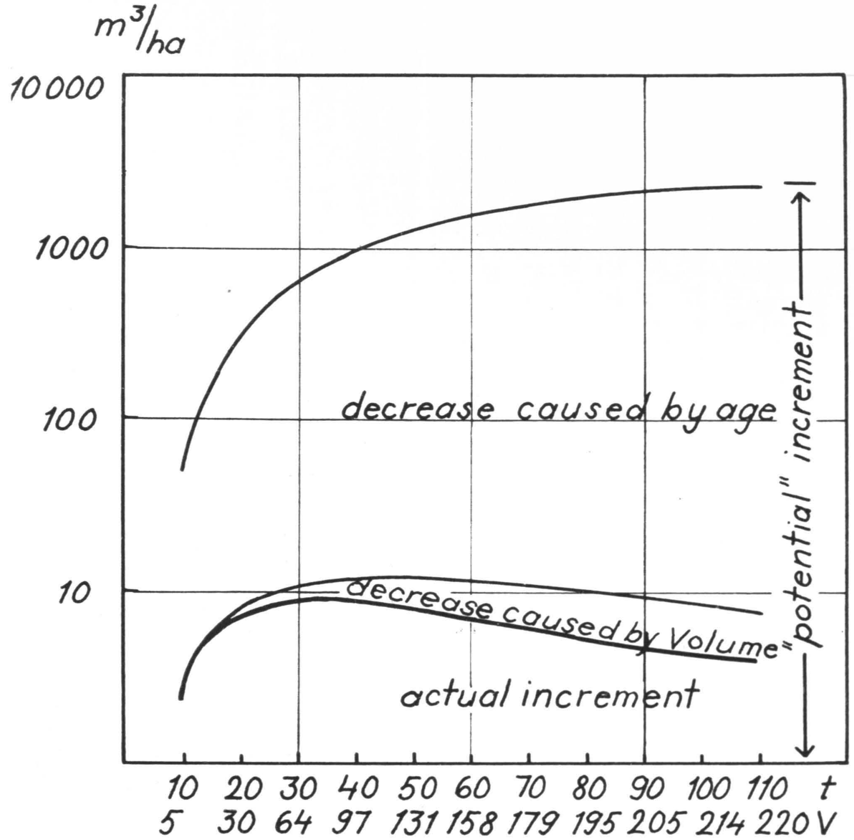

Fig. 13. An illustration of »potential» increment without the reducing effect of age and volue.

Kuva 13. »Potentiaalinen» kasvu ilman iän ja kuutiomäärän alentavaa vaikutusta (»potential increment» = »potentiaalinen kasvu», decrease caused by age = iän alentava vaikutus, decrease caused by volume = kuutiomäärän alentava vaikutus, actual increment = todellinen kasvu).

tion of site and which largely determines the amount of increment. Furthermore, the small significance of site in regression equations which have age and bole area as independent variables is a great practical advantage because an exact determination of site quality is more difficult than the determination of any other stand characteristic.

It is worth-while emphasizing the vigorous growing energy of young stands and the notable importance of volume as growing capital. If a stand is old or if its volume is small the production of timber is small even on the best sites.

Fig. 13 shows a speculative illustration (based on a regression equation of increment percentage, age and volume) of how great the increment might be without any reduction in increment caused by increasing age and volume. The

actual increment is calculated by the regression equation for a series of growing stock volume of average yield-table pine stands treated with cuttings during a rotation of 110 years. The second line above actual increment shows the increment calculated without the reducing effect of increasing volume, and the top line shows the »potential» increment without the reducing effect of increasing age and volume. If a stand of volume 214 m³/ha, excluding bark, were to have an increment persentage as great as that of a very young stand of very small volume, the annual increment would be more than 1000 m³/ha. Unfortunately, in an old and fully stocked stand most of energy is spent in maintaining the old and large-size living organisms and in continuing their struggle for existence on a crowded site.

References — Kirjallisuus

ASSMAN, ERNST. 1961. Waldertragskunde. München.

CHAPMAN, HERMAN H. and WALTER H. MEYER. 1949. Forest mensuration. First edition. New York, Toronto, London.

EIDE, E. og. A. LANGSAETER. 1941. Produksjonundersökelser i granskog. Referat: Produktionsuntersuchungen von Fichtenwald. Medd. f. det norske Skogforsöksvesen, 7: 356—500.

ERTELD, W. 1957. Einiges über die ertragskundliche Forschungsmethodik der Gegenwart. Summary: Some thoughts about the methods of research in modern forest mensuration. Archiv Forstw., p. 421—431.

HAGBERG, NILS. 1943. Skogsindelning. Stockholm.

ILVESSALO, YRJÖ. 1920. Kasvu- ja tuottotaulut Suomen eteläpuoliskon mänty-, kuusi- ja koivumetsille. Referat: Ertragstafeln für die Kiefern-, Fichten- und Birkenbestände in der Südhälfte von Finnland. AFF 15.

—»— 1942. Suomen metsävarat ja metsien tila. II valtakunnan metsien arviointi. Summary: The forest resources and the condition of the forests of Finland. The second National Forest Survey. MTJ 30.

—»— 1942a. Pystypuiden kuutioimis- ja kasvunlaskentataulukot. Helsinki.

—»— 1947. Pystypuiden kuutioimistaulukot. Summary: Volume tables for standing trees. MTJ 34.

JEFFERS, J. N. R. 1960. Experimental design and analysis in forest research. Stockholm.

JONSSON, BENGT. 1961. Om barrblandskogens volymproduktion. Yield of mixed coniferous forests. Medd. fr. statens skogsforskn. inst. Bd 50: 8.

KUUSELA, KULLERVO. 1953. Zur Theorie der forstlichen Zuwachsberechnung auf Grund der periodischen Messung. AFF 60.

KUUSELA, KULLERVO ja AARNE NYYSSÖNEN. 1962. Tavoitehakkuulaskelma. Summary: The cutting budget for a desirable growing stock. AFF 74.

LEXEN, BERT. 1943. Bole area as an expression of growing stock. Journal of Forestry 41.

MICHAJLOW, J. 1952. Mathematische Formulierung des Gesetzes für Wachstum und Zuwachs der Waldbäume und Bestände. Schw. Zeitschr. Forstw.

MÖLLER, CARL MAR. 1945. Untersuchungen über Laubmenge, Stoffverlust und Stoffproduktion des Waldes. Det forstl. Forsögsv. i Danmark, Bd. XVII.

—»— 1954. The influence of thinning on volume increment. Thinning problems and practices in Denmark. Technical Publication Number 76. College of Forestry at Syracuse, New York.

NYYSSÖNEN, AARNE. 1954. Hakkuilla käsiteltyjen männiköiden rakenteesta ja kehityksestä. Summary: On the structure and development of Finnish pine stands treated with different cuttings. AFF 60.

—»— 1958. Kiertoaika ja sen määrittäminen. Summary: Rotation and its determination. MTJ 49.

PETRINI, S. 1948. Skogsuppskattning och skogsindelning. Stockholm.

—»— 1957. Beräkning av skogens tillväxt och avverkning. Norrtälje.

PETTERSON, HENRIK. 1955. Barrskogens volymproduktion. Medd. fr. statens skogsforskn. inst. Bd. 45: IA.

SCHLETTER, A. 1954. Betrachtung zur Weckschen Methode der Wachstumsdiagnose. Archiv Forstw.

SIRÈN, GUSTAF. 1961. Skogsgränstallen som indikator för klimafluktuationerna i Norra Fennoskandien under historisk tid. Summary in English. MTJ 54.

SPURR, STEPHEN H. 1952. Forest inventory. The Ronald Press Co, New York.

TÖRNQVIST, LEO. 1957. A method for calculating changes in regression coefficients and inverse matrices corresponding to changes in the set of available data. Skandinavisk Aktuarietidskrift N:o 3—4.

VUOKILA, YRJÖ. 1960. Siperialaisten lehtikuusikoiden kehityksestä ja merkityksestä maamme metsätaloudessa. Summary: On the development of Siberian Larch stands and their importance to forestry in Finland. MTJ 52.

WECK, J. 1948. Über die Brauchbarkeit von Wachstumsgesetzen als diagnostisches Hilfsmittel der Waldwachstumskunde. Forstw. Centralbl.

Abbreviations — Lyhennykset

AFF = Acta forestalia fennica
MTJ = Communicationes Instituti Forestalia Fenniae

SELOSTE:

KASVUPROSENTIN JA MUIDEN METSIKKÖTUNNUSTEN VÄLINEN YHTEISKORRELAATIO MÄNNIKÖISSÄ

Johdanto

Metsiköiden kasvu ilmoitetaan kasvutaulukoissa tavallisesti kuutiometreinä sekä usein myös prosenttina, kun taas metsien inventointien tuloksissa on sekä Suomessa että muissa Pohjoismaissa käytetty paljon kasvuprosenttia. Prosentin etuna on pidetty sen vähäistä riippuvuutta boniteetista ja puulajista.

Kasvututkimusten regressioanalyyseissä on selitettävänä muuttujana ollut useimmiten kuutiokasvu. Kasvuprosenttiakin on käytetty, mutta pääasiassa silloin, kun on tutkittu jonkin kuutiokasvun kompponentin kuten pohjapinta-alan, läpimitan ja pituuden kehittymistä. Kasvuprosenttiin selitettävänä muuttujana liittyy se regressioanalyysiä helpottava etu, että funktion kuvaaja on useiden selittävien tunnusten kuten iän ja pituuden suhteen yhteen suuntaan kupera käyrä, kun taas absoluuttisen kasvun kuvaaja on muodoltaan monimutkaisempi.

Käsillä olevan työn tarkoituksena on tutkia niitä periaatteita, joiden mukaan kasvufunktioita voidaan käyttää tehtäessä kasvun ennuste laajalle metsäalueelle. Työ on kohdistunut niihin metsikkötunnuksiin, jotka sopivat selittäviksi muuttujiksi.

Aineisto ja sen alustava käsittely

Koska kysymyksessä on laajempaan ohjelmaan liittyvä alustava tarkastelu, ei sitä varten katsottu tarpeelliseksi uuden aineiston keräämistä. Inventointikoealat edustavat parhaiten metsiköiden rakennevaihtelua, joten ne olisivat muuten sopivia, mutta kotimaisilla inventointikoealoilla on se puute, että niitä mitattaessa ei ole tutkittu jakson aikana hakattujen puiden kasvua. Työ onkin tehty käyttäen Nyyssösen (1954) julkaisemaa männikköaineistoa, jonka tärkeimmät ominaisuudet on esitetty taulukossa 1 ja kuvassa 2. Aineiston alustavassa käsittelyssä on laskettu eräitä analyysissä tarpeelliseksi katsottuja selittäviä tunnuksia.

Selitettävänä muuttujana on kuorettoman kuutiokasvun prosentti (p_i), joka on mittausjakson keskimääräinen vuotuiskasvun suhde saman jakson keskimääräiseen metsikön kuutioon. Ikä (t) on biologinen ikä jakson keskellä. Boniteetin kvantitatiiviseksi ilmaisemiseksi useampana kuin metsätyyppien mukaisena kolmena luokkana valmistettiin kuvassa 2 esitetty käyrästö luonnontilaisten männiköiden kasvutaulukoiden perusteella. Siitä saadaan boniteettitunnus (B) kaikille niille metsiköille, joissa alkuperäistä vallitsevaa puustoa on jäljellä. Harsituille metsiköille on käytetty asianomaisen metsätyypin keskimääräistä arvoa.

Muina selittävinä muuttujina on metsikön kuoreton kuutiomäärä ($v = m^3/ha$) jakson keskellä, pohjapinta-ala jakson lopussa ($g_e = m^2/ha$) ja jakson keskellä ($g_m = m^2/ha$), keskiläpimitta ($d_g = cm$) ja keskipituus ($h_g = m$), molemmat viimeksi mainitut jakson lopussa. Pohjapinta-ala, keskiläpimitta ja -pituus eivät ole moitteettomia selittäviä muuttujia, koska ne eivät osoita ko. muuttajien arvoa jakson keskellä eikä niissä ilmene jakson aikana suoritetun hakkuun vaikutus. Näiden tunnusten osalta päätelmät ovatkin vain suuntaa antavia.

Tutkimusmenetelmä

Alustava analyysi on sekä graafinen että laskennallinen. Kasvuprosentit sijoitettiin akselistoon jonkin muuttujen funktiona, pisteparvi tasoitettiin sopivalla matemaattisella käyrällä, luettiin havaintojen poikkeama käyrästä ja näin saatujen »jäännösarvojen» riippuvuutta tarkasteltiin jälleen jonkin muun muuttujan suhteen graafisesti. Kuvat 3 ja 4 ovat esimerkki yhdestä tämän vaiheen tarkastelusta.

Havaintojen riippuvuutta tarkasteltiin myös useammasta kuin yhdestä tunnuksesta tehdyn yhdistelmämuuttujan funktiona. Eräs tällainen on muotoa

$$p_i = \frac{1}{t\,(v + b)} + c$$

Vakioista ratkaistaan b kokeilemalla ja c pienimmän neliösumman menetelmällä. Aineiston 190 metsikön kasvuprosenttien suhteellinen hajonta keskiarvoon nähden on 61 %. Yhtälöllä saadun estimaatin keskivirheen suhteellinen arvo on 21 %, joten havaittujen arvojen vaihtelusta selittyy noin $^2/_3$ (kuva 5).

Alustava vaihe antoi käsityksen funktioiden muodosta ja tunnusten arvosta selittävinä muuttujina. Sen jälkeen suoritettiin kokeita hyperbolisella funktiolla

$$y = a\,x^b$$

ja eksponentiaalisella funktiolla

$$y = a\;10^{bx}$$

joiden logaritmiset muodot ovat

$$\log y = \log a + b \log x$$
$$\log y = \log a + b\ x$$

Kokeiden tarkoituksena oli löytää kullekin muuttujalle sopiva muoto. Päättelyn perusteena käytettiin funktion eri muotojen antamia yhteiskorrelaatiokertoimia, estimaatin keskivirheitä ja graafista tarkastelua, joista viimeksi mainittu osoitti, miten funktio kulkee pisteparvessa ja ennenmuuta pisteparven äärialueilla (kuvat 6—9).

Esitettyjen työvaiheiden jälkeen valittiin ne funktioiden muodot, joille laskettiin tietokoneella regressioyhtälöt. Valintaa tuki vielä regressiokertoimien taulukko (taulukko 2). Regressioyhtälöille saatiin vakiot ja kertoimet keskivirheineen, korrelaatiokertoimet ja estimaatin keskivirhe vaiheittain (TÖRNQUIST 1957) aina uuden muuttujan lisäämisen jälkeen.

Analyysissä käytettiin pääasiassa kahta yhtälöä. Ensinnäkin pidettiin ikää ensisijaisena selittävänä muuttujana ja pyrittiin löytämään kaikki ne muut muuttujat, jotka parantavat yhtälöä. Toisena vaihtoehtona pyrittiin löytämään funktio, jossa ikä ei ole ollenkaan selittävänä muuttujana, sillä, jos näin saataisiin käyttökelpoinen funktio, päästäisiin vaivalloisesta iän arvioimisesta.

Laskennan tulokset ovat taulukossa 3. Tuloksia havainnollistaa myös kuva 10, joka osoittaa havaittujen ja laskettujen kasvuprosenttien logaritmien vaihtelun keskiarvoon nähden.

Metsikkötunnusten arvo selittävinä muuttujina

Ikä ja kuutiomäärä ovat tärkeimmät selittävät muuttujat. Niiden lisäksi ei boniteettitunnuksella ole aineiston puitteissa merkitsevää vaikutusta. Jos yhtälöön otetaan lisäksi keskipituus tai keskiläpimitta, niin tällöin on boniteettitunnuksesta hyötyä ja jäännösvaihtelusta voidaan vielä osa selittää.

Kasvuprosenttien suhteellinen hajonta keskiarvoon nähden on 61 %. Funktiolla $\log p_i = f(\log t, v)$ lasketun estimaatin suhteellinen keskivirhe 19 %. Keskipituus ja boniteetti lisätunnuksina antavat vielä jonkin verran paremman tuloksen.

Kun ikä jätetään käyttämättä, on pituus tärkein selittävä tunnus. Tällöin on kuitenkin boniteettitunnus välttämätön ja ellei sitä voida arvioida kvantitatiivisesti ilman ikää, tulee ikä tällöinkin mukaan. Ilman ikää on parhaalla yhtälöllä saatu tulos jonkin verran huonompi, kuin jos ikä on mukana. Tarpeelliset tunnukset ovat pituus, kuutiomäärä ja boniteetti sekä myös keskiläpimitta.

Metsikön kasvu on pääasiassa riippuvainen iästä, kuutiomäärästä, keskipituudesta tai keskiläpimitasta ja boniteetista. Koska puuston kasvuvaippojen pinta-ala on lähinnä kuutiomäärän ja keskipituuden (tai keskiläpimitan funktio), voidaan sanoa, että kasvu riippuu iästä, kasvuvaippojen pinta-alasta ja kasvupaikan laadusta.

Kasvupaikan laadulla on tosin regressiyhtälöissä näennäisesti pieni merkitys, mutta se johtuu siitä, että muiden metsikkötunnusten yhdistelmä on jo sinänsä kasvupaikan funktio. Käytännössä tällä on suuri merkitys sen vuoksi, että kasvuprosenttifunktio voidaan usein laskea ottamatta boniteettitunnusta selittäväksi muuttujaksi.

Selittävien tunnusten mittausteknilliset ominaisuudet ja päätelmät

Jos kasvu lasketaan mittausjakson alussa todettavien metsikkötunnusten funktiona, voidaan kuutiomäärän ja pohjapinta-alan lisäksi arvioida sekä keskipituus että keskiläpimitta nopeasti ja luotettavasti. Jos taas käytetään jakson keskikohdan tunnuksia ja jos jakson aikana on poistumaa, tuottaa keskipituuden ja keskiläpimitan täsmällinen määrittäminen vaikeuksia.

Iän määrittäminen on likimain tasaikäisissä ja ennenmuuta havupuuvaltaisissa metsiköissä varmaa, joskin kairausten suorittaminen vie enemmän aikaa kuin useimpien muiden tunnusten arvioiminen. Eri-ikäisissä ja lehtipuuvaltaisissa metsiköissä on eduksi, jos ikää ei tarvittaisi kasvufunktiossa.

Kasvupaikan laadun arvioiminen objektiivisesti ja sen ilmoittaminen kvantitatiivisesti on vaikeinta. Silloin kun metsikön alkuperäiset vallitsevat puut ovat jäljellä ja kun ne ovat alusta pitäen kasvaneet vapaasti, voidaan käyttää pituuden ja iän perusteella selvitettyä boniteettitunnusta. Keskimääräisissä olosuhteissa on boniteetin määrittäminen niin epätäsmällistä, että se vaikeuttaa tuntuvasti sellaisten kasvufunktioiden käyttämistä, joissa tarvitaan boniteettitunnusta.

Sanotun perusteella ja ottaen huomioon regressioanalyysien tulokset, ikä ja kuutiomäärä ovat suomalaisissa olosuhteissa tärkeimmät kasvua selittävät metsikkötunnukset, joita käyttäen voidaan laskea käyttökelpoisia kasvuprosenttifunktioita puulajivaltaisuuden mukaisille metsiköluokille. Suurella metsäalueella voidaan erottaa myös laaja-alaisia boniteettiluokkia. Esim. kasvufunktio voidaan laskea erikseen tuoreille ja kuiville kankaille. Kun pääosa metsistä on hyvin hoidettuja, lisääntyy keskipituuden ja keskiläpimitan merkitys. Vaikka aineisto ei annakaan mahdollisuuksia tehdä päätelmiä pohjapinta-alan merkityksestä, voitaneen sitä käyttää kuutiomäärän asemesta.

Kasvuprosenttifunktioiden sovellutuksia

Metsäalueen puustolle selvitetyllä kasvuprosenttifunktiolla voidaan laskea taulukko, josta saadaan absoluuttinen kuutiokasvu selittävien muuttujien funktiona. Sellainen on esim. taulukko 4.

Jos jakson kasvu ilmoitetaan koronkorkolaskun jälkiarvotekijänä, voidaan sen prosenttiosalle laskea samanlaisia funktioita kuin edellä on laskettu vuotuiskasvun prosentille. Tällaiset funktiot ovat sellaisenaan käytettävissä tavoitehakkuulaskelman kasvunennusteessa (Kuusela ja Nyyssönen 1962).

Ikään ja kuutiomäärään perustuvalla funktiolla voidaan laskea se kuutiomäärä, joka antaa suurimman absoluuttisen kasvun. Tässä käytetyn aineiston puitteissa se on 332 m³/ha. Kuvista 11 ja 12 nähdään, miten kasvu on funktion mukaan riippuvainen iästä ja kuutiomäärästä. Tulos on kuutiomäärän suhteen kutakuinkin sama, mikä on saatu keski- ja pohjois-eurooppalaisissa harvennuskokeissa.

Eräitä näkökohtia

Kasvuprosenttifunktioiden merkitys on huomattava kasvututkimuksissa ja ilmoitettaessa inventointien kasvutuloksia hakkuulaskelmia varten ennen muuta siksi, että kun kasvu tutkitaan sopivien metsikkötunnusten funktiona, saadaan luotettavia tuloksia suhteellisen pienellä koealojen määrällä.

Funktiot tarjoavat myös mahdollisuuden tutkia kasvutapahtumaa biologisena ilmiönä. Tällöin on kuitenkin korostettava varovaisuuden ja suuren kriitillisyyden merkitystä, sillä regressioanalyysi ei aina osoita jonkin selittävän muuttujan vaikutusta sinänsä, koska jokainen muuttaja on puolestaan muiden muuttujien funktio. Tässä lasketut funktiot osoittavat, miten voimakkaasti suureneva ikä ja kuutiomäärä pienentävät metsikön »potentiaalista» kasvukykyä (kuva 13). Samoin jo aikaisemmin esitetty päätelmä suurimman kasvun antavasta kuutiomäärästä sekä pienenevän kuutiomäärän vaikutuksesta kasvun suuruuteen on merkityksellinen tutkittaessa kasvua biologisena ja ekologisena tapahtumana.