Leaflet No. 19/ 81-82

Government of the Union of Myanmar
Ministry of Forestry
Forest Department

Multivariate Statistical Analysis on Existing Forest Inventory Data

U Htun Lynn
Forest Research Institute
January – 1982

လက်ရှိသစ်တောပေါက်ရောက်မှုစာရင်းများအားစာရင်းအင်း ပညာနည်းအရလေ့လာခြင်း

ဦးထွန်းလင်း
သစ်တောစာရင်းအင်းဌာနစိတ်
သစ်တောသုတေသနဌာနဗိဟာန်

စာတမ်းအကျဉ်းချုပ်

မြန်မာပြည်တွင် ၁၉၆၃ခုနှစ်မှစတင်ပြီး ၁၉၇၄ ခုနှစ်အထိပြုလုပ်ခဲ့သော သစ်တောသယံဇာတ စာရင်းကောက်ယူမှုများမှ အချက်အလက်များအပေါ် အခြေခံပြီး၊ အရေးပါသောသစ်မျိုးများ၏ ပါဝင်ပင် အရေအတွက်၊ ရောနှောပေါက်ရောက်ပုံနှင့် နှိုင်းယှဉ်ပြန့်နှံ့မှုများကို ICL ၁၉၀၂ သင်္ချာတွက်စက်အား အသုံးပြု၍ လုံးပတ်အတန်းအစားနှင့် သစ်တောနယ်များအလိုက် ဇယားများကိုပြန်လည် ဖွဲ့စည်းမှုပြုပါသည်။ အရေးပါသောသစ်မျိုး အချို့အတွင်းဆက်စပ်မှုပြကိန်း လေ့လာချက်နှင့် နယ်ပယ်အလိုက် စံထုထည် ဇယားများနှင့် ဒေသအလိုက် ရောင်းတမ်းဝင်ထုထည်ဇယားများကို ကိန်းရှင်ထုအား အသုံးပြုတွက်ချက်ပြီး နှိုင်းယှဉ်တင်ပြ ဆွေးနွေးထားပါသည်။

Multivariate Statistical Analysis on Existing Forest Inventory Data

U Htun Lynn
Forest Biometrics and Statistics Section
Forest Research Institute

Abstracts

Forest inventories conducted in Burma from 1963 to 1974 have resulted in a large accumulation of data. Based upon these data, stock, composition and relative density of important species by girth classes and forest divisions were retabulated by the use of ICL1902 computer. Analysis of correlation coefficients were made for some important species. By using dummy variables, regional standard volume tables, local standard volume tables and local merchantable volume tables of some important species were determined.

Contents

1. Introduction

Ever since 1963, the Forest Department of the Union of Burma has conducted forest inventory every year with the object of collecting data for management planning and of developing practical, economical and statistically sound methods of sampling and analysis. From these surveys an accumulation of quite a considerable amount to important statistics such as growing stocks of trees species composition, growth rates, moralities etc. have resulted for each surveyed area. If sufficiently detailed, they are, in actuality, the basic tools without which the scientific working if the forests is impossible.

The present paper deals with

(1) retabulation of stocks, stock per acre and composition of 90 species of trees by girth classes and Forest Divisions

(2) construction of volume tables by using dummy variables .

2. Data Processing

The data were compiled from the forest inventory survey reports. The following tables shows the forest divisions inventoried from the year 1963 to 1975.

Table 2.1 Forest Divisions inventoried since 1963-64

Serial No.	Years	Forest Division	Sampling method
1.	1963-64	Pyinmana (Binbyin)	stratified two stage topographical -plots
2.	1964-65	North Taungoo	"
3.	1965-66	Pyinmana	"
4.	1966-67	Yamethin	"
5.	1967-68	Mandalay / Maymo	Stratified two stage topographical-plots, strip and circular plots
6.	1968-69	East Katha, Mongmit and Bhamo	Stratified replicated
7.	1969-70	Upper Chindwin/Myittha	"
8.	1970-71	Shwebo	"
9.	1971-72	Upper Chindwin/Myittha and West Katha	"
10.	1972-73	Shwebo	"
11.	1974-75	Prome	Stratified two stage

Excluded Bamboo Surveys

After coding (codings were made according to the Field Instruction Manual by Kyaw Tint (1980)), checking and editing, these data were sent to Universities' Computer Center (UCC) that the data could be punched on to cards. The data cards have the following layout ;

Card column	Characters	Description
1- 2	2	Forest division code
3- 5	3	Forest reserve or stratum code
6- 7	2	Year
8-13	6	Reserve or stratum area in acre
14-16	3	Tree species code
17-80	63	Stock of trees by 1 ft. girth classes

After punching, the card were run on the computer for data cleaning i.e. to see whether impossible data were on the cards. Manual checking were made for coding errors and digital position and wrong data were then rectified. The clean data file was established on tape file under the name DATA-FORESTS for use in the tabulation programme.

3. Tabulation

A FORTRAN program was used to produce (a) stock of trees (b) trees per acre (c) species composition by girth classes and forest divisions. Two printouts were made on duplicating paper so that four sets of tables were available. The tables show an absoulate zero by a dash. Zeros are printed where the per-acre value is too small for even four decimals. There are 90 species in the stand table which show at least a non-zero row in the stock table. These tabulated species are listed below :

Serial No.	Tree species	Serial No.	Tree species	Serial No.	Tree species
1.	Teak	32.	Myaukngo	63.	Thingadu
2.	Pyinkado	33.	Nabe	64.	Thingan
3.	Akyaw	34.	Ngu	65.	Thinwin
4.	Anan-bo	35.	Nyaung	66.	Thit-E
5.	Auchinsani	36.	Odein	67.	Thitka
6.	Baing	37.	Padauk	68.	Thitkado
7.	Binga	38.	Panga	69.	Thitkya
8.	Bonmeza	39.	Peinnebo	70.	Thitkyabo
9.	Chinyoke	40.	Pat-wun	71.	Thitmagyi
10.	Dahat	41.	Pinlekwe	72.	Thitsi
11.	Didu	42.	Polthinma	73.	Thitto
12.	Gwe	43.	Pyauksaik	74.	Thitya
13.	Hnaw	44.	Pyilonchantha	75.	Wetshaw
14.	In	45.	Pyinma	76.	Yamane
15.	Ingyin	46.	Sagawa	77.	Yindaik

16.	Kanyaung	47.	Sandwa	78.	Yinma
17.	Kanyin	48.	Sawbya	79.	Yon
18.	Karawe	49.	Setkadon	80.	Zinbyun
19.	Kashit	50.	Sha	81.	Myaukloak
20.	Kaungmu	51.	Sit	82.	Linlum
21.	Kokhe	52.	Tamalan	83.	Taung-meok
22.	Kokko	53.	Taukkyan	84.	Thitpayoung
23.	Kuthan	54.	Taung peinne	85.	Mini-auga
24.	Kyi	55.	Taung tangyi	86.	Yingatgyi
25.	Laukya	56.	Taung thayet	87.	Petshat
26.	Letpan	57.	Twa thayet	88.	Box wood
27.	Leza	58.	Thabye	89.	Match wood
28.	Leza-byu	59.	Thabyu	90.	Others
29.	Lunbo	60.	Thadi		
30.	Maukadone	61.	Than		
31.	Mauletanshe	62.	Thayet		

A sample table for Teak species is presented in Appendix I. Tables for 90 species by girth classes and forest divisions are available at the libarary of Forest Research Institute. In doing research work on a certain species of tree, the researcher can easily find the growing stocks, stock per acre and species composition from these tables.

4. Construction of Volume Tables

4.1 Volume table for in

Volume table equations for standard, local and merchantable volumes for in species were prepared for various geographical localities since 1969. Due to the limitation of the calculation works, the volume equations were mostly based on the linear regression of volume on basal area within each height class. One of the main objective of the present section is to introduce the method of selecting the best regression equation in conctructing the tree volume table. Instead of describing the detail procedure, we prefer illustrating its application to an actual set of some 320 sample trees of in species selected from the some reserved and public forests in Shwebo forest division 1970-71.

Volume data were preliminarily tabulated with reference to two variables 10-foot height classes and 6-inch girth at breast height (gbh). The first main step in the Analysis is the selection of an appropriate basic regression model to approximate the true but unknown regression function of y= volume on X_1= diameter and X_2= height.

Table 4.1 The frequency table of in species

Girth class (ft)	Height class in feet								
	46-55	**56-65**	**66-67**	**76-85**	**86-95**	**96-105**	**106-115**	**116-125**	**Total**
4' 6" – 4' 11"	4	2	5	2	2	3			18
5' - 5' 6"	2	11	16	7	6	1	1		44
5' 6" - 5' 11"		2	8	12	7	4			33
6' - 6' 5"	2	3	12	18	7	6	1		49
6' 6" - 6' 11"	1	1	8	7	7	5	1		30
7' - 7' 5"	1		10	10	9	9	6		45
7' 6" - 7' 11"		1	1	2	7	3	4		18
8' - 8' 5"			2	6	7	10	2	3	30
8' 6" - 8' 11"				1	4	10		2	17
9' - 9' 5"					1	7	7	2	17
9' 6" - 9' 11"				1	1	4	2	2	10
10' - 10' 5"							3		3
10' 6"-10' 11"						1	1	3	5
11'-11' 5"						1			1
Total	**10**	**20**	**62**	**66**	**58**	**64**	**28**	**12**	**320**

Looking at the above table the basic data information is adequate for most girth and height classes. The following models were used as the basic models.

Model I $y = B_o + B_1 D + B_2 D^2 + B_3 D^3 + B_4 D^4$
Model II $y = B_o + B_1 D^2 + B_2 D^2 H + B_3 H$
Model III $y = B_o \; D^{B_1} H^{B_2}$
where y = volume
D = diameter at breast height
H = height of the tree

The regression equations for each reserve were derived, using the method of stepwise multiple regression. To determine which of the basic models provides the best fit, we have used the residual mean square error as a pivot value. The estimated regression equations are show in table 4.2.

Table 4.2 Volume table equations

Reserve	Model	N	no.of varia -bles	Best equation with K variables	Residual mean square error	R^2
Kyathin and Kyathin extn reserves	I	38	1	$-32.46 + 44.29\,D^2$	43.2876	.832
			2	$3.68 + 32.96\,D^2 - 2.33D^3$	40.3395	.838
			3	$-120.75 + 200.66D^2 - 125.64D^3 + 20.68D^4$	41.5693	.834
			4	$-523.6 + 1287.9D + 3284.8D^2 - 6759.9D^3 + 58.99D^4$	43.9645	.834
	II		1	$2.64 + .07\,D^2\,H$	25.4231	.932
			2	$22.07 - 14.42\,D^2 + .4889\,D^2\,H$	24.9567	.936
			3	$131.67 - 36.36\,D^2 + .7763D^2\,H - 1.4837\,H$	24.42	.941
	III		1	$22.5206\,D^{2.3955}$	28.7467	.860
			2	$.3597\,D^{1.6019}\,H^{1.0733}$	23.1681	.914
Naung – U public forest	I	85	1	$80.81 + 3.55D^4$	50.2676	.892
			2	$10.2 + 40.76\,D + 2.8925D^4$	50.2927	.893
			3	$-188.71 + 195.99\,D^2 - 101.\,D^3 + 17.84\,D^4$	50.3009	.895
			4	$1278.6 - 2434.7D + 1750.95\,D^2 - 532.74\,D^3 + 61.86\,D^4$	50.5267	.895
Naung – U public forest	II	85	1	$4.5359 + .3651\,D^2\,H$	47.76	.903
			2	$35.06 + .3787\,D^2\,H - .3925\,H$	47.93	.904
			3	$137.08 - 22.18D^2 + .5898D^2\,H - 1.3839\,H$	47.94	.905
	III		1	$29.5568\,D^{2.2302}$	42.3253	.900
			2	$.8724\,D^{1.9037}\,H^{.8298}$	39.3552	.919

Table. 4.2(contd.)

Reserve	Model	n	no.of varia -bles	Best equation with K variables	Residual mean square error	R^2
Ponhmwa proposed reserve	I	44	1	$38.0062 + 11.4055\ D^3$	54.4949	.889
			2	$-57.2871 + 77.01D + 2.1098\ D^4$	54.7136	.891
			3	$-393.11 + 463.55\ D^2 – 238.34\ D^3 + 36.834\ D^4$	53.2216	.900
			4	$6523.6 – 12072\ D + 8184\ D^2 – 2381\ D^3 + 255\ D^4$	50.9875	.911
	II		1	$18.1217 + .3381\ D^2\ H$	49.4196	910
			2	$85.3869 + .3846\ D^2\ H – 1.0185\ H$	48.9110	.914
			3	$171.4165 – 17.3734\ D^2 + .552\ D^{2H} – 1.8781\ H$	48.9900	.916
	III		1	$23.8588\ D^{2.4084}$	44.4953	.914
			2	$4.3322\ D^{2.1503}\ H^{.4273}$	42.9529	.922
Paungkadaun g reserve	I	15 3	1	$-34.9255 + 39.6166\ D^2$	45.6448	.879
			2	$2.6430 + 20.8807\ D^3 – 2.99\ D^4$	45.7616	.879
			3	$6.0364 – 3.8711\ D^2 + 23.016\ D^3 – 3.3079\ D^4$	45.9158	.879
			4	$-954.28 + 1689.26\ D – 1090.1\ D^2 + 325.77\ D^3 – 34.1918\ D^4$	46.0183	.879
Paungkadaun g reserve	II	15 3	1	$-34.9255 + 39.6166\ D^2$	45.6448	.879
			2	$-56.0572 + 37.866\ D^2 + .3739\ H$	45.4313	.881
			3	$-54.6082 + 37.5139\ D^2 + .0039\ D^2\ H + 3574\ H$	45.5842	.881
	III		1	$21.1203\ D^{2.4574}$	42.9850	.856
			2	$10.8439\ D^{2.3353}\ H^{.1749}$	42.4022	.861
All reserves	I	32 0	1	$-40.9267 + 42{,}3877\ D^2$	48.2848	.888
			2	$-97.977 + 99.0093\ D + 1.766\ D^4$	48.0632	.889
			3	$-41.75865 + 63.753\ D^2 – 17.7987\ D^3 + 3.585\ D^4$	48.1380	.889
			4	$1356 + 345\ D – 4566\ D^2 – 76\ D^3 + 24455\ D^4$	49.8876	.889
	II		1	$16.1755 + .3439\ D^2\ H$	46.9731	.894
			2	$-11.0464 + 17.91\ D^2 + .2054\ D^2 + .2054\ D^2\ H$	45.7313	.900
			3	$7.1699 + 14.722\ D^2 + .2414\ D^2\ H - .2169\ H$	48.2844	.901
	III		1	$22.3627\ D^{2.4531}$	42.6029	.882
			2	$6.3719\ D^{2.2339}\ H^{.3222}$	40.9557	.893

In comparing the residual mean square errors of model I, II and III, it may be conclude that model III consisting of the allometric function $y = B_o D^{B}{}_{1} H^{B}{}_{2}$ is the best eqation in constructing the volume table for in species.

4.2 Use of dummy variables in tree volume table construction

We have calculated four individual, all different regressions, one for each forest reserve and a pooled regression for all reserves. More reasonable, however, is to expect some similarities among some of the regression function so that the volume tables could be constructed from parallel regressions. It is the purpose of this section to present the procedure by which regressions are compared for similarities and statistical test are constructed and carried out to justify the decisions made.

From Table 4.2, the best estimate of the giant size regression is

$$\begin{aligned} y =: \quad & -1.0226\, X_{11} + 1.6019\, X_{12} + 1.0733\, X_{13} \\ & -.1365\, X_{21} + 1.9037\, X_{22} + 0.8298\, X_{23} \\ & + 1.4661\, X_{31} + 2.1503\, X_{32} + 0.4278\, X_{33} \\ & + 2.3836\, X_{41} + 2.3353\, X_{42} + 0.1749\, X_{43} \end{aligned}$$

where

$X_{11}=1$, $X_{12}=\text{Ln (D)}$, $X_{13}= \text{Ln (H)}$	if the tree is from Kyatthin reserves, or zero otherwise.
$X_{21}=1$, $X_{22}=\text{Ln (D)}$, $X_{23} = \text{Ln (H)}$	if the tree is from Naung-U Public forest, or zero otherwise
$X_{31}=1$, $X_{32}=\text{Ln (D)}$, $X_{33} = \text{LN (H)}$	if the tree is from Pohawa proposed reserve, or zero otherwise
$X_{41}=1$,$X_{42}=\text{Ln (D)}$, $X_{43}=\text{LN (H)}$	if the tree is from Paungkdaund reserve,or zero otherwise

Null Hypothesis 1: The slopes of diameter and height of the four regression lines are parallel

H_{01}; $B_{12} = B_{22} = B_{32} = B_{42} = B_2$

$B_{13} = B_{23} = B_{33} = B_{43} = B_3$

The sample regression function is

$$\begin{aligned} y \quad = \quad & 2.0587X_{11} + 2.2204\, \text{Ln (D)} + .2698\, \text{Ln (H)} \\ & 2.1621\, X_{21} + 2.1160\, X_{31} + 2.0582\, X_{41} \end{aligned}$$

Under Null Hypothesis 1 the slope differences are equal to zero and test statistic for significance is $F = 2.5478$ with 6 and 307 degree of freedom. This value is significant at the 5% level so that the Null Hypothesis 1 is rejected.

Null Hypothesis 2; The slope of height are parallel

H_{02}; $B_{13} = B_{23} = B_{33} = B_{43} = B_3$

The restricted sample regression function is

$$y = 3.0913\ X_{11} + 2.3911\ X_{12} + 3.3608\ X_{21}$$
$$2.2279\ X_{22} + 3.1482\ X_{31} + 2.4048\ X_{32}$$
$$3.0274\ X_{41} + 2.4532\ X_{42} + 0.006\ \mathrm{Ln}\ (H)$$

The sample F value is 11.1646 with 3 and 307 degree of freedom. Because this is a value which is highly significant, Null Hypothesis 2 is also rejected.

Null Hypothesis 3; The slopes of diameter are parallel

$$H_{03};\ B_{12} = B_{22} = B_{32} = B_{42} = B_2$$

To test this null hypothesis, the restricted regression is calculated as

$$Y = 0.1228\ X_{11} + 2.1586\ \mathrm{Ln}\ (D) + .7277\ X_{13}$$
$$+ 0.6321\ X_{21} + 0.6158\ X_{23} + 1.4834\ X_{31}$$
$$+ .4218\ X_{33} + 2.3298\ X_{41} + 0.2175\ X_{43}$$

Performing the required calculations we find that the sample value of F is 2.5216 with 3 and 307 degrees of freedom. As this value is not significant, the Null Hypothesis 3 is accepted. we are now in a position where we feel entitled to use parallel slopes for diameter. We still do not know whether there are differences among the intercepts of the various regressions. Then, we have to test the Null Hypothesis as follows.

Null Hypothesis 4; The differences in intercept terms of the four regressions are equal to zero, that is

$$H_{04};\ B_{11} = B_{21} = B_{31} = B_{41} = B_1$$

Under the decision of the previous testing, the restricted sample regression is

$$y = 2.0731 + 2.2186\ \mathrm{Ln}\ (D) + 0.2675\ X_{13}$$
$$+ 0.2898\ X_{23} + 0.2799\ X_{33} + 0.2662\ X_{43}$$

Because of the F value is not significant at 5% level, Null Hypothesis is accepted.

Consequently, the final regression for constructing tree volume table for in the following .

$V = 7.9494\ D^{2.2186}\ H^{.2675}$ for Kyatthin reserve

$= 7.9494\ D^{2.2186}\ H^{.2898}$ for Naung-U publicforest

$= 7.9494\ D^{2.2186}\ H^{.2799}$ for Pohawa proposed reserve

$= 7.9494\ D^{2.2186}\ H^{.2662}$ for Paungkadung reserve

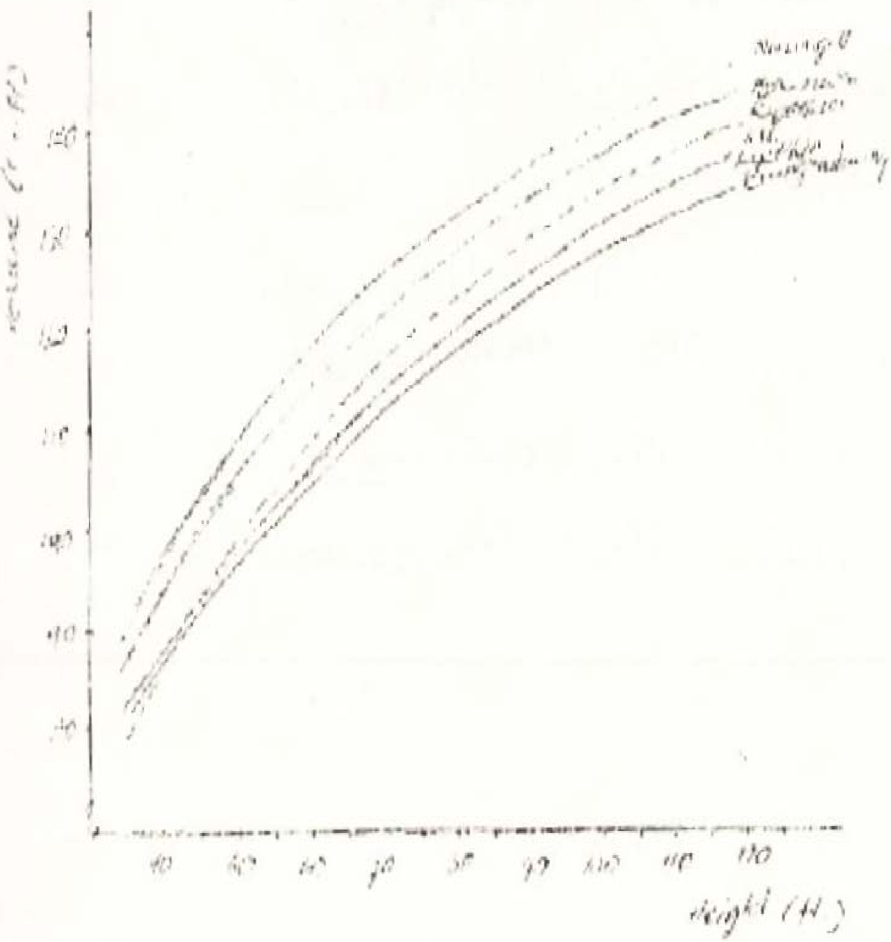

Fig. 4.1 Estimated regressions for in

5. Merchantable Volume Tables for Species other Tran in

Merchantable volume tables for Ingyin, Taukkyan, Nabe, Padauk, Thinwin, Pyinkado, and Thitya were constructed in Mandalay / Maymyo deport. The method of constructing merchantable volumes were the same as those applied for in as explained in the proceeding section. Volume table equations for merchantable volumes for different species are shown in table 5.1. The results of the significance tests are highly significant. and hence the Null Hypothesis related to parallelism, identity and equidistance among regressions are rejected.

Table 5.1 Volume table equations

Species	No. of trees	Volume table equations	Residual error	R^2
Ingyin	17	V= e .3451 $D^{1.7589}$ $H^{1.1186}$	7.7896	0.98
Taukkyan	22	V= e .8732 $D^{1.9945}$ $H^{0.4577}$	8.8015	0.87
Nabe	26	V= e .79 $D^{.4157}$ $H^{0.6578}$	10.3346	0.92
Padauk	26	V= e .2826 $D^{1.5667}$ $H^{.7912}$	21.4533	0.88
Thinwin	18	V= e .0223 $D^{2.0014}$ $H^{.8512}$	17.8862	0.94
Pyinkado	46	V= e .2072 $D^{2.0514}$ $H^{.75}$	12.8799	0.95
Thitya	21	V= e .034 $D^{2.0514}$ $H^{.75}$	18.4494	0.96

6. Conclusion

We have described how to apply the technique of dummy variables in constructing the volume table equation. The results indicate that in choosing the best regression equations the allometric form $y = B_1 D^{B}{}_2 H^{B}{}_3$ is satisfactory for most of the tree species. We also tried to devide the forest divisions into groups or clusters on the basis of the stock per acre of some tree species. Of eight forest divisions, Pyinmana and Yamethin were taken as Group I, North Taungoo and Prome as Group II, Katha, Upper Chindwin and Shwebo as Group III and the last Mandalay/Maymyo as Group IV on the basic of geographical location. When discriminatory analysis was applied on the species Teak, Pyinkado, Padauk, In, Ingyin and Yamane using Statistical Analysis Mark II on the computer ICL 1902 S, the various tests revealed that these groups discriminated one another fairly well. Division wise tests showed that membership probabilities of belonging to their own group were very high and to the other groups were extremely small except grap IV. The following table shows these probabilities.

Average probabilities of 'Hits and Misses'

Forecast Group	Actual Group			
	I	II	III	IV
I	0.78	0.08	0.04	0.10
II	0.07	0.87	0.01	0.05
III	0.03	0.04	0.91	0.02
IV	0.12	0.26	0.06	0.56

This study should be extended by adding more tree species. For ton trees pecies, it needs 45 computer runs. The major problem is that it repuires considerable core storage and computer running time.

Appendix I

Teak Growing Stoke

G.B.H (ft.)	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9+	TOTAL
Pyinmana (63-64)	123035	71231	62092	59919	50227	30525	9679	-	486708
N-Toungoo (64-65)	330727	197178	136158	103878	77061	43215	7980	3366	899561
Pyinmana (65-66)	434608	261+26	201838	180327	151992	86696	22601	18541	1358229
Yemethin (66-67)	153713	135641	759308	155946	123326	55154	8666	3969	1395723
Mand/Maym (67-68)	115676	122045	122924	99849	65474	32022	8149	4842	570981
East-Katha (68-69)	252598	167649	106731	52817	22241	6795	1644	1192	611167
U-Chindwin (69-70)	115876	120743	108687	92831	64672	38712	11614	11788	564923
Shwebo (70-71)	31326	24603	18351	11781	6629	3597	851	1254	98392
West-Katha (71-72)	151139	170876	141361	68405	23396	6342	1764	649	563923
U-Chindwin (71-72)	9615	6121	5299	3163	1481	985	288	461	27413
Shwebo (72-72)	134626	150001	148166	114885	77725	36814	8501	5336	676052
Prome (74-75)	252150	363370	383217	292017	168125	80552	24848	13760	1578039

References

1. Cunia, T. (1973) Dummy variables and some of their uses in regression analysis, proceedings of the june 1973 meeting of IUFRO Subject Group 84.02, Vol.I.
2. Kyaw Tint (1980)Field Instructions for the National Forest Inventory of Burma, Forest Dept.
3. ICL Statistical Analysis mark 2