KFRI Research Report 117

STATISTICAL ANALYSIS PACKAGE FOR FOREST MENSURATION

K. Jayaraman

KERALA FOREST RESEARCH INSTITUTE
PEECHI, THRISSUR

October 1996

Pages: 38

CONTENTS

ABSTRACT

A software package dealing with certain statistical analyses in forest mensuration was developed. The package consists of three modules, viz., SAMPLE, TREEVOL and TREEBIOM.

SAMPLE deals with parameter estimation for sampling from finite populations. The module covers a wide array of sampling plans commonly used in forestry. The module allows the user to specify the sampling plan used and provides a frame for entering and editing data. The estimates of population parameters like mean, total and ratio are made available by the programme with corresponding standard errors.

TREEVOL is for developing prediction equations for commercial volume at tree level based on measurements of diameter at breast-height (dbh) and height or dbh alone. The programme accepts billet level or treelevel data, selects the best fitting regression function from a set of candidate functions for predicting commercial volume and provides the user a volume table for the range specified.

TREEBIOM is useful for developing biomass prediction equations for the whole tree or its components like stem, branches and leaves based on diameter at breast-height (dbh) and height or dbh alone. The programme offers facilities for entering and editing measurements on dry weight and fresh weight of samples, develops estimates of dry weight at tree level, selects the best fitting regression function from a candidate set of functions for predicting biomass and provides a biomass table for the range specified.

The package fills in the need for a self contained user-friendly software for executing certain important statistical applications in forest mensuration and is conceived to be useful for a wide array of professionals in forestry and allied fields. The present version of the package is PC based, runs under DOS 3.0 and above and requires a minimum of 384 KB run time memory.

1. INTRODUCTION

Forest mensuration is one of the many fields of forestry, where statistical applications abound. Forest mensuration is concerned with measurements in forests at widely different resolution levels like forests, stands, trees or leaves. Estimation of population parameters and prediction of system attributes using certain easily measured characteristics are two major problems at any of the levels mentioned above. Naturally, sampling and regression techniques have found many application in forest mensuration. Successful execution of the related statistical analyses requires a fairly high level expertise and availability of appropriate computer programmes. General statistical analysis packages like SAS, SPSS, BMDP etc., demand a high level of knowledge and familiarity with the use of such software and are not in the reach of many of the professionals in forestry. The present project was undertaken to fill in the need for a user-friendly package dealing with some of the commonly occurring statistical analyses in forest mensuration. This was also conceived as part of the technology development in view of the Government of India's policy decision on self-financing of research institutions in the country. Consequent on the directives issued to work torwards research results suitable for commercialization, software development for statistical applications was considered as one of the means of generating funds.

Three main modules were developed under the present project. These were named as SAMPLE, TREEVOL and TREEBIOM respectively. The first one deals with parameter estimation for sampling from finite populations. Most of the commonly used sampling plans in forestry are covered by the programme. Sampling is an important activity in forestry and many related fields like agriculture and fisheries. The software developed will be useful in estimating population parameters related to sample surveys conducted in all these fields. The next module is on tree volume prediction and is thereby more specific to forestry. Construction of volume table is an age old practice in forestry. In the recent past, regression techniques have been widely used for the purpose. The programme developed has the capacity to accept data at the billet or tree level, choose the best fitting function and produce a volume table for specified range of input variables. Recent advances in forestry in general have put more emphasis on biomass production. Biomass is measurable only destructively and accordingly prediction of biomass through nondestructive means has gained importance. A computer programme which is useful in developing biomass prediction equations will be of value to researchers in the concerned fields. More detailed description of each of the above programmes can be found in the following sections.

2. GENERAL REMARKS ON THE PROGRAMMES

The three modules that are developed presently are incorporated in the following executable files, viz, SAMFLE.EXE, TV.EXE and TB.EXE. The above files can be copied on to hard drive or can be used from the floppy diskettes as such. The programmes are invoked by the names of the executable files at the DOS prompt. Further instructions on proceeding with the programmes appear on the screen. Help on running the programmes can be obtained while executing the programmes by pressing the F1 key. Sample data sets for running the different programmes are provided in the programme diskette under the subdirectory by name, DATA. The programmes require a minimum of 384 KB memory but may allocate higher levels of run time memory to process large data sets. The programme runs under DOS Ver. 3.0 and above.

Functions of some of the keys during programme operation are shown below.

Key	Function
↑ → ← ↓	Edit keys
F1	Help
F2	Save
Esc	Previous screen
Ins	Insert record
Del	Delete record

3. DESCRIPTION OF THE PROGRAMME SAMPLE

SAMPLE is a collection of programmes useful for estimating population parameters like mean, total and ratio for different sampling schemes. The package covers a wide array of sampling schemes involving stratification, multistage and multiphase sampling with systematic or random sampling patterns at the final stage. It also takes cam of ratio or regression estimation with an auxiliary variable and performs computations associated with probability proportional to size (PPS) sampling. The programme can handle up to 3 response (main) variables at a time and has built-in data editing facilities.

At the initiation, the programme tries to understand the sampling scheme of the user by requesting for information about the same. After preliminary queries, the data entry screen appears according to the scheme specified. The data can be entered through key board or can be read from an ASCII file containing the properly ordered data set. Format having a width of 9 columns with 2 decimals has to be used for each column of the input data file. The output can be diverted to the screen, printer or a disk file.

All input/output values are set to 2 decimals. However the internal computations are not subjected to this restriction. If large values are present in the input data set, the whole data are to be properly scaled to fit in with the format specified by the SAMPLE. The results will have to be transformed back to the original scale with appropriate computations.

The sampling schemes covered by the programme SAMPLE are discernible from Table 1. The formulae used correspond to those given by Chacko (1964) and Sukhatme and Sukhatme (1970). The exact set of formulae used is given in Appendix 1.

The following points are to be noted about the data entry and the computations.

Sampling is assumed to have been done without replacement from a finite population for all schemes covered by SAMPLE except for'the PPS scheme.

Estimation of variance under systematic sampling utilizes first order differencing of the successive values. Hence the values are to be entered in correspondence with the order of observations in the field. For two dimensional systematic sampling, linearization by dividing the area into a series of strips and joining them hypothetically to form a single continuous strip for one-way systematic sampling is suggested. For this purpose, the end of strip will be joined with the end of the next strip and beginning of that strip with the beginning of the next strip and so on.

PPS sampling is assumed to have been done with replacement. The observations for a unit are to be entered as many times as that unit has been selected in the repeated sampling.

The response (main) variables in the programme are indicated as Y1, Y2, Y3. The phase is indicated in brackets like YI(Pl), Yl(P2) etc. In the case of two-phase sampling, enter zero for Yl(P2) in places where observations are not available for the same but are available for Yl(P1).

A value of ' 1 ' has to be given for number of strata for sampling schemes not involving stratification.

When regression estimates are requested for, with an auxiliary variable or in the case of two-phase sampling, the results include an estimate of the population ratio also, obtained through regression method of estimation.

For two-stage sampling schemes, systematic sampling is understood to be carried out at the second-stage. Similarly for two-stage sampling involving PPS sampling, the latter is understood to be done at the first-stage.

Illustrations on running the programme SAMPLE for the case of ratio estimation under simple random sampling with auxiliary variate is given in the following.

A :\>SAMPLE ↵

```
                Welcome
                  to

         SAMPLE       1.0

   Developed by Division of Statistics
    Kerala Forest Rersearch Institute
         Trichur,  Kerala,  India

   (Unauthorized duplication prohibited)
```

```
Press ,any key to continue...
```

```
           ═SAMPLING SCHEME═
  NUMBER OF STRATA               :1
  NUMBER OF STAGES               :1
  NUMBER OF PHASES               :1
  NO. OF RESFOHSE VARIABLES      :1
  SYSTEMATIC/ RANDOM (S/R)       :r
  AUX ILIARY VARIABLE(Y/N)       ;y
  FROB. PROP. TO SIZE(Y/N)       :n
```

```
F1 = Help ▌ Esc = Exit to Dos ▌
```

```
 ST ═    TOTAL-SIZE    SAMP-S IZE      AUX SIZE
  1          20             5             100
```

```
All entries OK(Y/N)
F1 = Help ▌ Esc = Previous screen ▌ Arrow keys to rdit ▌
```

```
ENTER FILE NAME: a:\data\ratio.dat
(For source of data/For saving data)
```

ST	UNIT	Y1(P1)	AUX VAR
1	1	5.50	10.00
1	2	3.20	6.00
1	3	6.80	15.00
1	4	2.00	5.00
1	5	4.60	6.00

```
All entries OK(Y/N):
F1 = Help ▌ Esc = Previous screen ▌ F2 = Save ▌ Arrow keys to edit ▌
```

DO YOU REQUIRE RATIO OR REGRESSION ESTIMATE(R/E):

SFECIFY THE DESTINATION FOR OUTPUT
Screen/Printer/Filer

* * * OUTFUT OF PROGRAM SAMFLE * * *

	MEAN	SE(MEAN)	TOTAL	SE(TOTAL)	RATIO	SE(RATIO)
STRATUM 1						
Y1 ::	2.63	0.37	52.62	7.48	0.53	0.07
POPULATION						
Y1 ::	2.63	0.37	52.62	7.48	0.53	0.07

Press any key to continue...

Table 1. Sampling schemes covered by the programme S A M P L E

Schem	Strata	Stage	Phase	Aux. var	PPS	systematic/ Random	included (Y/N)
1	No	Smgle	Single	No	No	Random	Y
2	No	Single	Single	No	No	Systematic	Y
3	No	Single	Single	No	Yes	Random	Y
4	No	Smgle	Single	No	Yes	Systematic	N
5	No	Single	Single	Yes	No	Random	Y
6	No	Single	Single	Yes	No	Systemtic	N
7	No	Single	Single	Yes	Yes	Random	N
8	NO	Single	Single	Yes	Yes	systematic	N
9	No	Single	TWO	No	No	Random	Y
10	No	Single	TWO	No	No	Systematic	N
11	No	Single	TWO	No	Yes	Random	N
12	No	Single	TWO	No	Yes	systematic	N
13	No	Single	Two	Yes	No	Random	N
14	No	Single	TWO	Yes	No	Systematic	N
15	No	Single	TWO	Yes	Yes	Random	N
16	No	Single	Two	Yes	Yes	systematic	N
17	No	TWO	Single	No	No	Random	Y
18	NO	Two	Single	No	No	Systematic	Y
19	No	TWO	Single	No	Yes	Random	Y
20	No	Two	Single	No	Yes	systematic	N
21	No	Two	Single	Yes	No	Random	Y
22	NO	TWO	Single	Yes	No	Systematic	N
23	NO	TWO	Single	Yes	Yes	Random	N
24	No	Two	Single	Yes	Yes	Systematic	N
25	No	TWO	Two	No	No	Random	Y
26	No	TWO	TWO	No	No	Systematic	N
27	No	TWO	TWO	No	Yes	Random	N
28	No	Two	Two	No	Yes	Systematic	N
29	No	Two	Two	Yes	No	Random	N
30	No	Two	TWO	Yes	No	systematic	N
31	No	TWO	Two	Yes	Yes	Random	N
32	No	TWO	TWO	Yes	Yes	Systematic	N

Cont ...

Table 1. cont...

Scheme	Strata	Stage	Phase	Aux. var	PPS	Systematic/ Random	Included (Y/N)
33	Yes	Single	Single	No	No	Random	Y
34	Yes	Single	Single	No	No	Systematic	Y
35	Yes	Single	Single	No	Yes	Random	Y
36	Yes	Single	Single	No	Yes	systematic	N
37	Yes	Single	Single	Yes	No	Random	Y
38	Yes	Single	Single	Yes	No	Systematic	N
39	Yes	§ingle	Single	Yes	Yes	Random	N
40	Yes	Single	Single	Yes	Yes	systematic	N
41	Yes	Single	Two	No	No	Random	Y
42	Yes	Single	Two	No	No	Systematic	N
43	Yes	Single	Two	No	Yes	Random	N
44	Yes	Single	Two	No	Yes	Systematic	N
45	Yes	Single	Two	Yes	No	Random	N
46	Yes	Single	Two	Yes	No	Systematic	N
47	Yes	Single	Two	Yes	Yes	Random	N
48	Yes	Single	Two	Yes	Yes	Systematic	N
43	Yes	Two	Single	No	No	Random	Y
50	Yes	Two	Single	No	No	Systematic	Y
51	Yes	Two	Single	No	Yes	Random	Y
52	Yes	Two	Single	No	Yes	Systematic	N
53	Yes	Two	Single	Yes	No	Random	Y
54	Yes	Two	Single	Yes	No	Systematic	N
55	Yes	Two	Single	Yes	Yes	Random	N
56	Yes	Two	Single	Yes	Yes	Systematic	N
57	Yes	Two	Two	No	No	Random	Y
58	Yes	TWO	Two	No	No	systematic	N
59	Yes	Two	Two	No	Yes	Random	N
60	Yes	Two	Two	No	Yes	Systematic	N
61	Yes	Two	Two	Yes	No	Random	N
62	Yes	Two	Two	Yes	No	Systematic	N
63	Yes	TWO	Two	Yes	Yes	Random	N
64	Yes	Two	Two	Yes	Yes	Systematic	N

4. DESCRIPTION OF THE PROGRAMMAE TREEVOL

TREEVOL is a set of programmes useful for developing prediction equations for commercial volume at tree level based on measuremcnts of diameter at breast-height (dbh) and height or dbh alone from a sample set of trees. The programme can accept billet level measurements and compute tree volume or accept tree level data on commercial volume directly. Different regression functions are then fitted and the best fitting function is identified. A tabular output of the individual tree volume (volume table) for specified ranges of dbh and height is also producible by the programme.

At the initiation, the programme tries to understand the prediction scheme of the user by requesting for information about the same. After preliminary queries, the data entry screen appears according to the scheme specified. The data can be entered through key board or can be read from an ASCII file containing the properly ordered data set. The formats for the input data files, when supplied as external ASCII files, are given below.

Data entry level	Type of measurement	Format
Billet	Tree number	6.0
"	Billet number	6.0
"	Basal girth	9.2
"	Mid-girth	9.2
"	Tip girth	9.2
"	Length	9.2
Tree	Diameter	10.3
"	Height	10.3
"	Volume	10.3

Note: Tree number need not be entered for tree level data.

Certain details of the computations performed by the TREEVOL are mentioned below.

Calculation of the billet volume is made using Newton's formula if basal girth, mid-girth, tip girth and length of the billets are available. The programme uses Muber's formula if only mid-girth and length are fed in. All billet level measurements are to be supplied in centimetres. Although, the above formulae compute the full circular volume, there is a provision available to convert it to the quarter girth volume. Quarter girth volume is obtained by multiplying the full circular volume by 0.785.

Newton's formula for computation of tree volume is given below.

$$\text{Tree volume} = \sum_{i=1}^{n} \left[\frac{l_i}{24\pi}\right] [b_i^2 + 4m_i^2 + t_i^2] \tag{1}$$

where l_i = length of the i *th* billet from a tree

b_i = basal girth of the i *th* billet from a tree

m_i = mid-girth of the i *th* billet from a tree

t_i = tip-girth of the i *th* billet from a tree.

Since the girth and length values are entered in cm, the volume is divided by 1000000 to convert the unit to m^3.

Huber's formula for computation of tree volume is shown below.

$$\text{Tree volume} = \sum_{i=1}^{n} l_i \left[\frac{m_i^2}{4\pi}\right] \tag{2}$$

where l_i and m_i as are defined earlier.

Tree volume is computed in m^3 from the billet level measurements. If provided directly, tree volume is to be entered in m^3 as well. Tree diameter (dbh) and tree height (total height) may also be entered in metric units.

The prediction equations fitted are either based on dbh alone or dbh and height depending on the user's specification. The forms of the regression functions included are given below.

Based on dbh

$$b_0 + b_1 D \qquad (3)$$

$$\ln V = b_0 + b_1 \ln D \qquad (4)$$

$$V^{1/2} = b_0 + b_1 D \qquad (5)$$

$$V = b_0 + b_1 D + b_2 D^2 \qquad (6)$$

Based on dbh and height

$$V = b_0 + b_1 D^2 H \qquad (7)$$

$$\ln V = b_0 + b_1 D^2 H \qquad (8)$$

$$V^{1/2} = b_0 + b_1 D^2 H \qquad (9)$$

$$\ln V = b_0 + b_1 \ln D + b_2 \ln H \qquad (10)$$

$$V^{1/2} = b_0 + b_1 D + b_2 H \qquad (11)$$

$$V^{1/2} \quad b_0 + b_1 D^2 + b_2 H + b_3 D^2 H \qquad (12)$$

Where **V = Tree volume**
D = Diameter at breast-height of the tree
H = Total height of the tree
bi's are parameters
ln indicates natural logrithm

Furnival index (Furnival, 1961) is used for choosing the best fitting function The estimates of the coefficients, standard errors of the estimated coefficients, analysis of variance;, adjusted coefficient of determination and Furnival index are displayed for each fitted function.

All the computations related to the regression are by Ordinary Least Squares (OLS) performed through matrix operations. (Montgomery and Peck, 1982). The details are furnished in Appendix 2. In the volume table, values outside the range of data are shown in asteriks

Illustration on running the programme TRREVOL for the case of billet level data is given in the following.

A:\>TV↵

```
                 Welcome
                    to

               TREEVOL  1.0

    Developed  by Division  of Statistics
     Kerala   Forest Research Institute
          Trichur. Kerala.   India

   (Unauthorized   duplication prohibited)
```

```
Press any     to continue.
```

PREDICTION SCHEME

BILLET LEVEL/TREE LEVEL DATA(B/T) : b

NUMBER OF TREES : 15

F1 = Help ‖ Esc = Exit to Dos ‖

INDICATE THE DATA AVAILABLE FOR BILLETS

BASAL GIRTH(Y/N) : y

MID GIRTH (Y/N) : y

TIP GIRTH (Y/N) : y

F1 = Help ‖ Esc = Previous screen ‖ Arrow keys to edit

ENTER FILE NAME: a:\data\bmt.dat
(For source of data/For saving data):Data on billets

TREE	BILLET	BASAL GIRTH	MID GIRTH	TIP GIRTH	LENGTH
1	1	129.00	99.00	89.00	570.00
1	2	89.00	90.10	91.00	630.00
1	3	64.00	60.00	54.90	68.00
1	4	76.00	85.00	84.60	102.00
1	5	84.90	80.10	76.[illegible]	111.00
2	1	164.20	123.20	108.50	510.00
2	2	108.50	102.50	95.00	565.00
2	3	95.00	87.50	81.00	538.00
2	4	76.50	71.00	76.00	94.00
2	5	72.00	73.00	64.00	94.00
2	6	73.00	73.00	73.60	104.00
3	1	145.00	119.30	109.20	434.00
3	2	107.00	100.70	97.00	549.00
3	3	97.00	79.70	68.00	546.00
3	4	68.00	57.80	61.00	600.00

All entries OK(Y/N):
F1 = Help ‖ Esc = Previous screen ‖ F2 = Save ‖ Arrow keys to edit ‖

```
Do you consider full circular volume/quarter girth volume (C/G):
```

```
INDICATE THE PREDICTOR VAR  BLES TO BE USED

        DBH    (Y/N)     :     y
        HEIGHT(Y/N)      :     y

F1 = Help | Esc = Previous screen | Arrow keys to edit
```

```
ENTER FILE NAME: a:\data\bdh.dat
(For source of data/For saving data): Data on dbh/height
```

TREE	DBH	HEIGHT
1	0.380	26.500
2	0.470	26.900
3	0.440	29.000
4	0.620	29.900
5	0.540	32.100
6	0.380	27.700
7	0.500	30.700
8	0.430	29.900
9	0.480	28.400
10	0.480	27.200
11	0.520	29.500
12	0.400	29.500
13	0.440	28.200
14	0.260	20.900
15	0.440	31.300

```
All entries OK(Y/N):
F1 = Help | Esc = Previous screen | F2 = Save | Arrow keys to edit |
```

SPECIFY THE DESTINATION FOR OUTPUT
Screen/Printer/File:

RANGE OF VARIABLES IN THE REGRESSION

Variable	Minimum	Maximum
DBH(m)	0.2600	0.6200
Height(m)	20.9000	32.1000
Volume(cu.m)	0.3933	3.7299

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : Y = b0 + b1D^2H

COEFF	EST.COEFF	SE(EST.COEFF)
b0	-0.2467	0.2022
b1	0.3103	0.0308

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	7.7432	7.7432	101.8097
Residual	13	0.9887	0.0761	

Adjusted R square : 0.8781
Furnival index : 0.2758

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : ln(Y) = b0 + b1D^2H

COEFF	EST.COEFF	SE(EST.COEFF)
b0	-0.8351	0.1358
b1	0.2004	0.0208

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	3.2308	3.2308	94.2551
Residual	13	0.4456	0.0343	

Adjusted R square : 0.8695
Furnival index : 0.2758

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : Sqrt(Y) = b0 + b1D^2H

COEFF	EST.COEFF	SE(EST.COEFF)
b0	0.5153	0.0668
b1	0.1204	0.0102

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	1.1656	1.1656	140.[illegible]49
Residual	13	0.1078	0.0083	

Adjusted R square : 0.9089
Furnival index : 0.2223

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : ln(Y) = b0 + b1D^2H

COEFF	EST.COEFF	SE(EST.COEFF)
b0	-0.8351	0.1358
b1	0.2004	0.0206

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	3.2308	3.2308	94.2551
Residual	13	0.4456	0.0343	

Adjusted R square : 0.8895
Furnival index : 0.2758

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : Sqrt(Y) = b0 + b1D^2H

COEFF	EST.COEFF	SE(EST.COEFF)
b0	0.5153	0.0668
b1	0.1204	0.0102

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	1.1656	1.1656	140.[illegible]49
Residual	13	0.1078	0.0083	

Adjusted R square : 0.9089
Furnival index : 0.2223

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : Sqrt(Y) = b0 + b1D^2 + b2H + b3D^2H

COEFF	EST.COEFF	SE(EST.COEFF)
b0	-0.1066	0.3837
b1	9.1971	2.7697
b2	0.0174	0.0140
b3	-0.1742	0.0936

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	3	1.2226	0.4075	88.3848
Residual	11	0.0507	0.0046	

Adjusted R square : 0.9493

Furnival index : 0.1658

Press any key to continue...

THE BEST FITTING MODEL

Function : ln(Y) = b0 + b1ln(D) + b2ln(H)

Proceed with volume table construction (Y/N):

ENTER VALUES FOR CONSTRUCTING THE VOLUME TABLE

Enter minimum dbh in metres for volume table :.3
Enter maximum dbh in metres for volume table :.5
Enter interval for dbh in metres :.1
Enter minimum height in metres for volume table :25
Enter maximum height in metres for volume table :30
Enter interval for height in metres :1

F1 = Help ▌ Esc = Previous screen ▌ Arrow keys to edit

VOLUME TABLE * * * VOLUME (cu.m)

DBH(metre)	HEIGHT(metre)					
	25.00	26.00	27.00	28.00	29.00	30.00
0.30	0.539	0.525	0.511	0.498	0.486	0.474
0.40	1.215	1.182	1.151	1.122	1.095	1.069
0.50	2.283	2.221	2.162	2.108	2.058	2.007

Press any key to continue...

5. DESCRIPTION OF THE PROGRAMME TREEBIOM

TREBIOM is a get of programmes useful for dcveloping prediction equations for biomass at tree level based on measurements of diameter at breast-height(dbh) and height or dbh done from a sample set of trees. The programme can accept measurements of fresh and dry weight of samples from different parts of trees and compute the total dry weight of trees based on additional data on the total fresh weight of trees. The programme can also accept tree level data on dry weight directly. Different regression functions are then fitted and the best fitting function is identified. A tabular output of the individual tree biomass (biomass table) for specified ranges of dbh and height is also producible by the programme. Without loosing generality, the above algorithms can be used for developing individual prediction equations for any component of tree biomass like stem, branches or leaves.

At the initiation, the programme tries to understand the prediction schemc of the user by requesting for information about the same. After preliminary queries, the data entry screen appears according to the scheme specified. The data can be entered through kcy board or can bc read from an ASCII file containing the properly ordered data act. The formats for the input data files, when supplied as external ASCII files, are given below.

Data entry level	Type of	Format
Disc	Tree number	6.0
"	Disc number	6.0
"	Fresh weight	9.2
"	Dry weight	9.2
Tree	Diameter	10.3
"	Height	10.3
	Dry weight	10.3

Note : Tree number need not be entered for tree level data.

Certain details of the computations performed by the TREEBIOM are mentioned below.

Tree biomass (dry weight) is computed in tonnes utilizing the disc level mea surements and also data on total fresh weight of trees in tonnes. A pooled estimate of the ratio of fresh weight to dry weight is developed from the disc level measurements from each tree and then multiplied by the fresh weight of the tree to get an estimate of the dry weight of the tree. When provided directly, tree biomass either in fresh or dry weight terms is to be entered in tomes. Tree diameter (dbh) and tree height (total height) may also be entered in metric units.

The prediction equations fitted are either based on dbh alone or dbh and height depending on the user's specification. The forms of the regression functions included are given below.

Based on dbh

$$\ln W = b_0+b_1 D \quad (13)$$

$$\ln W = b_0+b_1 \ln D \quad (14)$$

$$W^{1/2} = b_0+b_1 D \quad (15)$$

$$W = b_0+b_1 D+b_2 D^2 \quad (16)$$

Based on dbh and height

$$W = b_0+b_1 D^2 H \quad (17)$$

$$\ln W = b_0+b_1 D^2 H \quad (18)$$

$$W^{1/2} = b_0+b_1 D^2 H \quad (19)$$

$$\ln W = b_0+b_1 \ln D+b_2 \ln H \quad (20)$$

$$W^{1/2} = b_0+b_1 \ln D+b_2 H \quad (21)$$

$$W^{1/2} = b_0+b_1 D^2+b_2 H+b_3 D^2 H \quad (22)$$

Where W = Tree biomass
D = Diameter at breast-heat
H = Total height
b_1's are parameters
In indicates natural logrithm

Furnival index (Furnival, 1961) is used for choosing the best litting function. The estimates of the coefficients, standard errors of the estimated coefficients, analysis of variance, adjusted coefficient of determination and Furnival index are displayed for each fitted function.

All the computations related to the regression are by Ordinary Least Squares (OLS) performed through matrix operations. (Montgomery and Peck. 1982). The details are furnished in Appendix 2. In the biomass table, values outside the range of data are shown in asterisk.

Illustration, on running the programme TREEBIOM for the case of tree level data is given below.

A : \> TB <┘

```
                    Welcome
                      to

                TREEBIOM    1.0

      Developed  by Division of Statistics
       Kerala Forest Research Institute
            Trichur.  Kerala,  India

     (Unauthorized duplication prohibited)
```

```
Press any key to continue.. .
```

```
=PREDICTION SCHEME=

DISC LEVEL/TREE  LEVEL DATA(D/T)     :   t

number OF TREES                   .  :   15
```

```
F1 = Help ▌ Esc = Exit to Dos ▌
```

INDICATE THE PREDICTOR VARIABLES TO BE USED

DBH (Y/N) : y

HEIGHT(Y/N) : n

F1 = Help | Esc = Previous screen | Arrow keys to edit

ENTER FILE NAME: a:\data\md.dat
(For source of data/For saving drta):Data on dry weight/dbh

TREE	DRY WT	DBH
1	0.404	0.389
2	0.799	0.479
3	0.710	0.446
4	1.868	0.628
5	1.198	0.540
8	9.514	0.380
7	1.448	0.590
8	0.628	0.430
9	0.830	0.488
10	1.199	0.469
11	I.938	0.524
12	0.610	0.499
13	0.608	0.440
14	0.289	0.260
15	0.668	0.440

All entries OK(Y/N) :
F1 = Help | Esc = Previous screen | F2 = Save | Arrow keys to edit |

SPECIFY THE DESTINATION FOR OUTPUT
Screen/Printer/Filer

RANGE OF VARIABLES IN THE REGRESSION

Variable	Minimum	Maximum
DBH(m)	0.2600	0.6200
Height(m)	0.0000	0.0000
Dry Weight(t)	0.2000	1.8600

<<< Press any key For next page

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : W = b0 + b1D + b2D^2

COEFF	EST.COEFF	SE(EST.COEFF)
b0	0.5952	0.4810
b1	-3.9307	2.1724
b2	9.5316	2.4356

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	2	2.0683	1.0341	105.6616
Residual	12	0.1174	0.0098	

Adjusted R square : 0.9373
Furnival index : 0.0989

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : ln(W) = b0 + b1ln(D)

COEFF	EST.COEFF	SE(EST.COEFF)
b0	1.7542	0.1138
b1	2.5334	0.1365

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	3.5381	3.5381	344.2948
Residual	13	0.1336	0.0103	

Adjusted R square : 0.9608
Furnival index : 0.0750

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : Sqrt(V) = b0 + b1D

COEFF	EST.COEFF	SE(EST.COEFF)
b0	-0.2501	0.0769
b1	2.5123	0.1676

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	0.6037	0.6037	224.6865
Residual	13	0.0349	0.0027	

Adjusted R square : 0.9411

Furnival index : 0.0892

<<< Press any key for next page >>>

*****RESULTS OF FITTING DIFFERENT FUNCTIONS*****

FUNCTION : ln(V) = b0 + b1D

COEFF	EST.COEFF	SE(EST.COEFF)
b0	-3.0383	0.1670
b1	6.0555	0.3639

ANALYSIS OF VARIANCE

SOURCE	DF	SS	MS	F
Regression	1	3.5071	3.5071	276.9150
Residual	13	0.1646	0.0127	

Adjusted R square : 0.9517

Furnival index : 0.1125

Press any key to continue...

THE BEST FITTING MODEL

Function (In(Y) = b0 + blln(D)

Proceed with weight table construction (Y/N):

ENTER VALUES FOR CONSTRUCTION THE WEIGHT TABLE

Enter minimum dbh in metres for weight table : .5
Enter maximum dbh in metres for weight table : 1.0
Enter interval for dbh in metres : .l

F1 = Help | Esc = Previous screen | Arrow keys to

*** * * * * WEIGHT TABLE * * * * ***

DBH(metre)	WEIGHT(tonnes)
0.500	0. 990
0.600	I.584
0.700	*
0.800	
0.900	•

Press any key to continue..

Note:* indicates values outside the range of prediction.

References

Chacko, V. J. 1964. A Manual on Sampling Techniques for Forest Surveys. The Manager of Publications, Forest Research Institute, Dehra Dun. 172 p.

Furnival, G.M. 1961. An index for comparing equations used m constructing volume table. Forest Science, 7(4) :337-341.

Montgomery, D.C. and Peck, E.A. 1982. Introduction to Linear Regression Analysis. John Wilcy and Sons. New York. 504 p.

Sukhatme, P.V. and Sukhatme, B.V. 1970. Sampling Theory of Surveys with Applications. Iowa State University Press, Ames, Iowa, U.S.A. 452 p.

Appendix 1 Formulae used in SAMPLE

SAMPLE uses the following general formulae for combining estimates over different strata and hence the formulae enlisted under individual sampling schemes are shown without such an operation.

$$\hat{\bar{Y}} = \sum_{t=1}^{K} (N_t/N)\,\hat{\bar{Y}}_t$$

$$V(\bar{Y}) = \sum_{t=1}^{K} (N_t/N)^2\,\hat{V}(\hat{\bar{Y}}_t)$$

where $\hat{\bar{Y}}$ = Estimate of the population mean

$\hat{\bar{Y}}_t$ = Estimate of the t th stratum mean

N = Number of units in the population

N_t = Number of units in the t th stratum

$\hat{V}(\hat{\bar{Y}})$ = Estimate of the variance of $\hat{\bar{Y}}$

$\hat{V}(\hat{\bar{Y}}_t)$ = Estimate of the variance of $\hat{\bar{Y}}_t$

Σ = Summation sign

K = Total number of strata

Similarly the following formulae also apply generally unless indicated otherwise.

$$\hat{Y} = N\,\hat{\bar{Y}}$$

$$\hat{V}(\hat{Y}) = N^2\,\hat{V}(\hat{\bar{Y}})$$

$$\hat{R} = \hat{Y}/X$$

$$\hat{V}(\hat{R}) = \hat{V}(\hat{Y})/X^2$$

where $\hat{Y}$ = Estimate of the population total

R Estimate of the population ratio

X = Total size of the auxiliary variable

Each sampling scheme is assigned a number, the complete specification of which is given in Table 1.

Single stage sampling schemes

Notation

N	=	Number of sampling units in the population
n	=	Number of sampling units in the sample
y_i	=	Value of some character y on the i *th* unit
x_i	=	Value of an auxiliary character x on the i *th* unit
Y	=	Total of y in the population
$\bar{Y}$	=	Mean of **y** in the population
X	=	Total of x in the population

Scheme I

$$\hat{\bar{Y}} = \bar{y}$$

$$\hat{V}(\hat{\bar{Y}}) = [(1/n) - (1/N)]\, s_y^2$$

where

$$s_y^2 = 1/(n-1) \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$\bar{y} = (1/n) \sum_{i=1}^{n} y_i$$

Scheme 2

$$\bar{Y} = \bar{y}$$

$$\hat{V}(\hat{\bar{Y}}) = \frac{1}{2n(n-1)} \sum_{i=1}^{n-1} [d(y_i)]^2$$

where

$$d(y_i) = y_{i+1} - y_i$$

$$\bar{y} = (1/n) \sum_{i=1}^{n} y_i$$

Scheme 3

$$\hat{\bar{Y}} = (1/nN) \sum_{i=1}^{n} y_i / p_i$$

$$\hat{V}(\hat{\bar{Y}}) = \frac{1}{n(n-1)N^2} \left[\sum_{i=1}^{n} (y_i/p_i)^2 - n\, \hat{Y}^2 \right]$$

where

$$p_i = \frac{X_i}{X}$$

Scheme 5

(a) Ratio estimate

$$\hat{\bar{Y}} = \frac{\bar{y}}{\bar{x}} \frac{X}{N}$$

$$\hat{V}(\hat{\bar{Y}}) = \frac{(1/n) - (1/N)}{n-1} \left[\sum_{i=1}^{n} y_i^2 - 2\hat{R}\sum_{i=1}^{n} y_i x_i + \hat{R}^2 \sum_{i=1}^{n} x_i^2\right]$$

where $\hat{R} = \bar{y} / \bar{x}$

$$\bar{y} = (1/n) \sum_{i=1}^{n} y_i$$

$$\bar{x} = (1/n) \sum_{i=1}^{n} x_i$$

(b) Regression estimate

$$\hat{\bar{Y}} = \bar{y} + b\ (\bar{X} - \bar{x})$$

$$\hat{V}(\hat{\bar{Y}}) = [(1/n) - (1/N)]\ s_e^2$$

where
$$b = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$s_e^2 = \frac{1}{n-2} \left[\sum_{i=1}^{n} (y_i - \bar{y})^2 - b \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})\right]$$

$$\bar{X} = (1/N) \sum_{i=1}^{N} x_i$$

Scheme 9

(a) Ratio estimate

$$\hat{\bar{Y}} = \bar{y}\ \bar{x}_{(n)} / \bar{x}_{(m)}$$

$$\hat{V}(\hat{\bar{Y}}) = \frac{s_y^2 - 2\hat{R}s_{yx} + \hat{R}^2 s_x^2}{m} + \frac{2\hat{R}s_{yx} - \hat{R}^2 s_x^2}{n}$$

where $\hat{R} = \bar{y}/\bar{x}_{(m)}$

$$\bar{y} = \sum_{i=1}^{m} y_i /m \quad , \quad \bar{x}_{(m)} = \sum_{i=1}^{m} x_i /m \quad , \quad \bar{x}_{(n)} = \sum_{i=1}^{n} x_i /n$$

$$s_y^2 = \sum_{i=1}^{m} (y_i - \bar{y})^2 / (m-1) \quad , \quad s_x^2 = \sum_{i=1}^{m} (x_i - \bar{x}_{(m)})^2 / (m-1)$$

$$s_{yx} = \sum_{i=1}^{m} (y_i - \bar{y})(x_i - \bar{x}_{(m)}) / (m-1)$$

*(b) **Regression estimate***

$$\hat{\bar{Y}} = \bar{y} + b(\bar{x}_{(n)} - \bar{x}_{(m)})$$

$$\hat{V}(\hat{\bar{Y}}) = \frac{s_{yx}^2}{m} + \frac{s_y^2 - s_{yx}^2}{n}$$

where $s_{yx}^2 = \frac{1}{m-2}\left[\sum_{i=1}^{m} (y_i - \bar{y})^2 - b^2 \sum_{i=1}^{m} (x_i - \bar{x}_{(m)})^2\right]$

$$s_Y^2 = \sum_{i=1}^{m} (y_i - \bar{y})^2 / (m-1)$$

$$b = \sum_{i=1}^{m} (x_i - \bar{x}_{(m)})(y_i - \bar{y}) / \sum_{i=1}^{m} (x_i - \bar{x}_{(m)})^2$$

Two-stage sampling schemes

Notation

N = Number of first-stage units

M_i = Number of second-stage units in the i ***th*** first-stage unit

$\bar{M}$ = Average number of second-stage units per first stage unit

n = Number of first-stage units in the sample

m_i = Number of second-stage units selected from the selected i th first-stage unit

y_{ij} = Value of the character y for the j ***th*** second-stage unit in the i ***th*** first-stage unit

Y = Total of y in the population

$\bar{Y}$ = Mean of y in the population

X = Total of x in the

Scheme 17

$$\hat{\bar{Y}} = (1/n\bar{M}) \sum_{i=1}^{n} M_i \bar{y}_{i(m_i)}$$

$$\hat{V}(\hat{\bar{Y}}) = \left[\frac{1}{n} - \frac{1}{N}\right] s'^2_b + \frac{1}{nN} \sum_{i=1}^{n} u_i^2 \left[\frac{1}{m_i} - \frac{1}{M_i}\right] s_i^2$$

where $\bar{y}_{i(m_i)} = \sum_{j=1}^{m_i} y_{ij} / m_i$

$$s'^2_b = \sum_{i=1}^{n} \left[u_i \bar{y}_{i(m_i)} - \hat{\bar{Y}}\right]^2 / (n - 1)$$

$$u_i = M_i / \bar{M}$$

$$s_i^2 = \sum_{j=1}^{m_i} \left[y_{ij} - \bar{y}_{i(m_i)}\right]^2 / (m_i - 1)$$

Scheme 18

$$\hat{\bar{Y}} = \frac{1}{n\bar{M}} \sum_{i=1}^{n} M_i \bar{y}_{i(m_i)}$$

$$\hat{V}(\hat{\bar{Y}}) = \left[\frac{1}{n} - \frac{1}{M}\right] s'^2_b + \frac{1}{nN} \sum_{i=1}^{n} u_i^2 \sum_{j=1}^{m_i-1} \frac{(y_{i,j+1} - y_{ij})^2}{2m_i(m_i-1)}$$

where $\bar{y}_{i(m_i)} = \sum_{j=1}^{m_i} y_{ij} / m_i$

$$s'^2_b = \sum_{i=1}^{n} \left[u_i \bar{y}_{i(m_i)} - \hat{\bar{Y}}\right]^2 /(n-1)$$

$$u_i = M_i / \bar{M}$$

Scheme 19

$$\hat{\bar{Y}} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{m_i} \sum_{j=1}^{m_i} z_{ij}$$

$$\hat{V}(\hat{\bar{Y}}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} (\bar{z}_{i(m_i)} - \hat{\bar{Y}})^2$$

where $z_{ij} = M_i \, y_{ij} / (M_0 p_i)$

$$p_i = X_i / X, \qquad M_0 = N\bar{M}$$

$$\bar{z}_{i(m_i)} = \sum_{j=1}^{m_i} z_{ij} / m_i$$

Scheme 21

(a) Ratio estimate

$$\hat{Y} = \frac{\bar{y}'_{s_2}}{\bar{x}'_{s_2}} \; \bar{X}$$

$$\hat{V}(\hat{Y}) = \left[\frac{(1/n)-(1/N)}{n-1}\right] \sum_{i=1}^{n} u_i^2 \left[(\bar{y}_{i(m_i)} - \hat{R}_{s_2} \bar{x}_{i(m_i)}\right]^2$$

$$+ \; (1/nN) \sum_{i=1}^{n} u_i^2 [(1/m_i) - (1/M_i)] \; d_i^2$$

where

$$d_i^2 = s_{iy}^2 - 2\hat{R}_{s_2} s_{ixy} + \hat{R}_{s_2}^2 \, s_{ix}^2$$

$$(1/n\bar{M}) \sum_{i=1}^{n} M_i \, \bar{y}_{i(m_i)} \quad , \quad \bar{x}'_{s_2} = (1/n\bar{M}) \sum_{i=1}^{n} M_i \, \bar{x}_{i(m_i)}$$

$$(1/N\bar{M}) \sum_{i=1}^{N} \sum_{j=1}^{M_i} x_{ij} \quad , \quad \hat{R}_{s_2} = \bar{y}'_{s_2} / \bar{x}'_{s_2}$$

$$\frac{1}{m_i-1} \sum_{j=1}^{m_i} (x_{ij} - \bar{x}_{i(m_i)}) (y_{ij} - \bar{y}_{i(m_i)})$$

$$\frac{1}{m_i-1} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_{i(m_i)})^2$$

$$\frac{1}{m_i-1} \sum_{j=1}^{m_i} (x_{ij} - \bar{x}_{i(m_i)})^2$$

Scheme 25

Ratio estimate

$$\hat{Y} = (1/n\bar{M}) \sum_{i=1}^{n} M_i \, \hat{Y}_{i(m_i)}$$

$$\hat{V}(\hat{Y}) = [(1/n) - (1/N)] \; s'^2_b + (1/nN) \sum_{i=1}^{n} u_i^2 \, v_i^2$$

where

$$\hat{Y}_{i(m_i)} = \bar{y}_{i(m_i)} \, \bar{x}_{i(n_i)} / \bar{x}_{i(m_i)}$$

$$\bar{y}_{i(m_i)} = (1/m_i) \sum_{j=1}^{m_i} y_{ij}$$

$$\bar{x}_{i(m_i)} = (1/m_i) \sum_{j=1}^{m_i} x_{ij}$$

$$\bar{x}_{i(n_i)} = (1/n_i) \sum_{j=1}^{n_i} x_{ij}$$

$$s'^2_b = \frac{1}{n-1} \sum_{i=1}^{n} [u_i \hat{\bar{Y}}_{i(m_i)} - \hat{\bar{Y}}]^2$$

$$u_i = \frac{M_i}{\bar{M}}$$

$$v_i^2 = \frac{s_{iy}^2 - 2\hat{R}_i s_{iyx} + \hat{R}_i^2 s_{ix}^2}{m_i} + \frac{2k_i s_{iyx} - \hat{R}_i^2 s_{ix}^2}{n_i}$$

$$s_{iy}^2 = \frac{1}{m_i-1} \sum_{j=1}^{m_i} [y_{ij} - \bar{y}_{i(m_i)}]^2$$

$$s_{ix}^2 = \frac{1}{m_i-1} \sum_{j=1}^{m_i} [x_{ij} - \bar{x}_{i(m_i)}]^2$$

$$s_{iyx} = \frac{1}{m_i-1} \sum_{j=1}^{m_i} [y_{ij} - \bar{y}_{i(m_i)}] [x_{ij} - \bar{x}_{i(m_i)}]$$

$$\hat{R}_i = \frac{\bar{y}_{i(m_i)}}{\bar{x}_{i(m_i)}}$$

Appendix 2. The Computations related to regression used in TREEVOL and TREEBIOM

The general multiple linear regression equation has the following form,

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots \beta_p x_{pi} + \varepsilon_i \quad , \qquad i = 1, 2, \ldots, n$$

From the general model, different specific model forms can be obtained by changing the nature of the dependent and independent variables. The above general form of the equation can be cast into the matrix form.

$$\underset{\sim}{Y} = \underset{\sim}{X} \underline{\beta} + \underline{\varepsilon}$$

where

$$\underset{\sim}{y}' = (y_1, y_2, \ldots, y_n)$$

$$\underset{\sim}{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \ldots & x_{p1} \\ 1 & x_{12} & x_{22} & \ldots & x_{p2} \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ 1 & x_{1n} & x_{2n} & \ldots & x_{pn} \end{bmatrix}$$

$$\underline{\beta}' = (\beta_0, \beta_1, \ldots, \beta_p)$$

$$\underline{\varepsilon}' = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$$

The assumptions of the Ordinary Least Squares are

$$E(\underline{\varepsilon}) = \underline{0}$$

$$E(\underline{\varepsilon}\, \underline{\varepsilon}') = \sigma^2 \underset{\sim}{I}$$

where E indicates the expectation operator

The estimates are obtained through the following formulae.

$$\hat{\underline{\beta}} = (\underset{\sim}{X}' \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{y}$$

$$\hat{D}(\hat{\underline{\beta}}) = \hat{\sigma}^2 (\underset{\sim}{X}' \underset{\sim}{X})^{-1}$$

where D indicates the dispersion operator

An estimate of σ^2 is obtained through analysis variance as shown below.

Analysis of Variance

Source	Degrees of freedom	Sum of squares	Mean square	F ratio
Regression	**P**	**SSR**	**MSR**	**MSR / MSE**
Residual	n-p-1	**SSE**	**MSE**	
Total	**n-1**	**SST**		

$$
\begin{aligned}
\text{where} \quad SSR &= \hat{\beta}' X' y - (\underline{1}' y)^2/n \\
SST &= \underline{y}' \underline{y} - (\underline{1}' y)^2/n \\
SSE &= SST - SSR \\
MSR &= SSR/p \\
MSE &= SSE/ln\text{-}p\text{-}1 = \hat{\sigma}^2
\end{aligned}
$$

The coefficient of multiple determination (R^2) is obtained as the ratio of SSR to SST. Additionally Adjusted coefficient of multiple determination is obtained as

$$
\text{Adj: } R^2 = 1 - \frac{(n-1)}{n-p-1}(1-R^2)
$$

The computation of Furnival index involves the following steps.

1. The value of the square root of MSE is obtained for each model under consideration, through analysis of variance.
2. The geometric mean of the derivative of the dependent variable with respect to 'y' is obtained for each model, from the observation. Note that the dependent variable in each model is some function of the variable y.
3. The Furnival index for each model is then obtained by multiplying the corresponding values of the square root of MSE with the inverse of the geometric mean.