# ABOVE-GROUND WEIGHT OF FOREST PLOTS — COMPARISON OF SEVEN METHODS OF ESTIMATION

H. A. I. MADGWICK

Forest Research Institute, New Zealand Forest Service,
Private Bag, Rotorua, New Zealand

(Received for publication 16 April 1982; revision 14 October 1982)

ABSTRACT

The basal area ratio method, unweighted regression of weight on tree diameter squared, and methods based on logarithmic regressions with a variety of published correction factors were compared using simulated sampling of nine forest plots. The bias and variability of estimates derived using the basal area ratio method, unweighted regression, and logarithmic regression without correcting for expected bias were less than when logarithmic regression was used with correction factors. Logarithmic regression appeared most affected by the inclusion of unrepresentative sample trees. The basal area ratio method yielded the most estimates closest to the measured plot weights and has the added advantage of being the simplest to apply.

## INTRODUCTION

Pressure on the forest resource to provide greater yields of fibre and, potentially, of fuel is leading to an expanding literature on forest biomass (Pardé 1980). One problem common to many research studies of forest biomass is the selection of a suitable method for estimating the weight of forest plots. The three methods most frequently used in the past have been the basal area ratio method, unweighted linear regressions of weight on diameter squared or its equivalent basal area, and regressions of weight on tree size (usually diameter) after logarithmic transformation (Satoo 1973). Of these the log-log regression method has received the greatest attention and recently Flewelling & Pienaar (1981) have summarised the considerable statistical literature relevant to this method. Each method for estimating plot weight embodies assumptions on error distributions and the like.

A number of authors have published comparisons of estimates based on a limited sample (e.g., Ovington & Madgwick 1959; Ando 1962; Crow 1971; Swank & Schreuder 1974; Egunjobi 1976). While such studies provide relative ranking of estimates they lack the control of a known plot weight. In other studies single methods have been investigated using replicated sampling of data from forest plots in which each individual tree had been weighed (Madgwick 1971, 1981; Madgwick & Satoo 1975).

This paper compares the three most commonly used methods of estimating stand weights, including five different ways of applying the log-log method as outlined by Flewelling & Pienaar (1981).

## METHODS

Data from nine sample plots used in earlier studies (Madgwick & Satoo 1975; Madgwick 1981) form the basis of this work (Table 1). The trees in each plot were ranked according to increasing diameter and then subdivided into five size classes containing as nearly as possible equal numbers of stems. Sets of 100 samples were then selected from each plot's data by taking one tree per diameter-class at random. Separate sets of samples were taken for each of the three components – stems, branches, and foliage. It should be noted that in the plot with fewest trees there were only 108 different samples possible whereas in the plot with most trees there were potentially over 14 million different samples of five trees. Such samples are not independent.

Each sample of five trees was used to calculate stand weight by each of the methods detailed in Table 2. For each set of 100 replicates the following information was recorded for each calculation procedure: minimum, mean, and maximum plot estimates, and variance of the estimates. Since the sample plots varied both in number of stems and in total weight, the estimates were scaled relative to the actual plot weights. Variances were adjusted using "finite population corrections" (Cochran 1963) to allow for differences in the number of stems per plot. For each sample the methods were ranked in order of the closeness of their estimate to the measured plot value and a tally was kept of the number of first, second, . . ., seventh placings. For each method the total sum of squares of scaled actual minus expected values was recorded.

## RESULTS

The average differences between actual and estimated weights of the three tree components were less than 1.6% of plot values using either the basal area ratio method, unweighted regression on $d^2$, or the uncorrected log-log regression (Table 3). The four correction values for the log-log procedure all tended to increase the bias, though in only two cases was the mean bias more than 4%. The coefficients of variation of estimates differed little among methods but increased in magnitude in the order stems, foliage, branches. Mean minimum estimates differed little among methods but, as expected, mean maximum values were almost always larger using methods based on log-log regressions. As a consequence, the range of estimates was smaller for the basal area ratio and unweighted regression methods and was more equally distributed about the mean than were those obtained using log-log regression.

The low average bias and relatively low coefficient of variation of the basal area ratio method are reflected in the over-all ranking of the seven methods (Table 4). The basal area ratio method gave many more "best" estimates than the other six methods. The sum of observed minus expected values squared was lowest for the basal area ratio method, followed closely by the estimates using unweighted regression.

## DISCUSSION

The relatively poor performance of the various correction factors when applied to the log-log estimates was unexpected. In an effort to identify the cause of this poor performance the regressions were studied in greater detail. In particular the distribution of error mean squares was tabulated (Fig. 1). For any one sample plot these errors tended

TABLE 1—Summary of sample plot data (from Madgwick & Satoo 1975)

| Plot No. | Species | Age (yr) | Plot area (m2) | Stems (N) | Mean diameter (cm) | Mean height (m) | Weight (tonnes/ha) Stems | Branches | Foliage | Origin |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **Abies sachalinensis*** | 9–30 | 1.5 | 45 | 1.66† | 1.12 | 22.6 | 8.5 | 15.1 | Natural regeneration |
| 2 | **Abies sachalinensis*** | 17–30 | 2 | 34 | 2.27† | 1.38 | 27.5 | 6.5 | 13.1 | Natural regeneration |
| 3 | **Betula ermanii*** | 18 | 24 | 25 | 4.93 | 7.00 | 51.5 | 8.9 | 3.6 | Natural regeneration |
| 4 | **Cryptomeria japonica*** | 10 | 37.2 | 16 | 7.97 | 5.24 | 24.0 | 6.4 | 14.7 | Plantation |
| 5 | **Cryptomeria japonica*** | 43 | 32 | 14 | 15.18 | 14.85 | 245.3 | 10.3 | 17.8 | Plantation |
| 6 | **Larix leptolepis**** | 18 | 180 | 14 | 11.05 | 9.11 | 13.6 | 6.1 | 2.0 | Plantation |
| 7 | **Pinus densiflora** | 15 | 20 | 13 | 7.13 | 6.61 | 41.6 | 6.4 | 4.7 | Natural regeneration |
| 8 | **Pinus radiata***** | 8 | 810 | 100 | 13.28 | 7.91 | 26.7 | 16.8 | 10.4 | Plantation |
| 9 | **Pinus virginiana** | 19 | 237 | 136 | 7.54 | 8.65 | 58.0 | 12.7 | 4.7 | Natural regeneration |

* Data collected by the joint study group on forest productivity of four universities, Japan
** Data collected by the joint study group on forest productivity of five universities, Japan
*** Data of Ovington **et al.** (1968) supplied by Dr J. D. Ovington and Dr W. G. Forrest
† At base of stem

TABLE 2—Methods of estimating the weights of tree plots

| | Method | |
|---|---|---|
| 1. | Basal area ratio method | $\dfrac{\sum^{n} w_i}{\sum^{n} d_i^2} \cdot \sum^{N} D_j^2$ |
| 2. | Unweighted regression on diameter squared | $\sum^{N} (a_1 + b_1 \cdot D_j^2)$ |
| | Log-log regression | $\sum^{N} \theta_i (\exp (a_2 + b_2 \cdot \log_e D_j))$ |
| 3. | where $\theta_1 = 1$ | |
| 4. | $\theta_2 = \exp (\frac{1}{2} s^2)$ | |
| 5. | $\theta_3 = g_m (\frac{1}{2} s^2)$ | |
| 6. | $\theta_4 = g_m (\frac{m+1}{2m} (1 - x_j) s^2)$ | |
| 7. | $\theta_5 = \exp ((1 - 3x_j) \frac{s^2}{2})$ for $x_j > \frac{1}{3}$ | |
| | $\theta_6 = \exp ((1 - 3x_j) s^2 \frac{m}{2(m+2)})$ for $x_j < \frac{1}{3}$ | |

and $a_k$ and $b_k$ are regression constants
N is the number of trees in the plot
n is the number of sample trees
$m = n - 2$
$d_i$ and $w_i$ are the diameter and weight of the $i^{th}$ sample tree
$D_j$ is the diameter of the $j^{th}$ tree in the plot
$s^2$ is the error mean square from the regression of $\log_e$ w on $\log_e$ d

$x_j = \frac{1}{n} + \dfrac{(D_j - d)}{\sum^{n} (d_i - \bar{d})^2}$ and $g_m$* is a tabulated value based on $m = n - 2$

* Bradu & Mundlak 1970

TABLE 3—Mean bias, coefficient of variation, minimum and maximum estimated plot weights of stems, branches, and foliage based on 100 replicated samples for each of nine sample plots

| Component | Method | Mean bias (%) | Coeff. of variation (%) | Mean minimum (%) | Mean maximum (%) |
|---|---|---|---|---|---|
| Stems | B.A. ratio | -0.7 | 5.4 | -12 | 12 |
| | Regression on $d^2$ | -0.5 | 5.4 | -12 | 13 |
| | Log-log regression | | | | |
| | $\theta_1$ | 1.4 | 6.7 | -13 | 21 |
| | $\theta_2$ | 2.9 | 6.8 | -12 | 24 |
| | $\theta_3$ | 2.7 | 6.7 | -12 | 23 |
| | $\theta_4$ | 2.2 | 6.6 | -12 | 22 |
| | $\theta_{5,6}$ | 0.8 | 6.6 | -13 | 19 |
| Branches | B.A. ratio | -1.4 | 18.7 | -32 | 41 |
| | Regression on $d^2$ | -1.4 | 17.9 | -30 | 42 |
| | Log-log regression | | | | |
| | $\theta_1$ | -1.6 | 17.6 | -32 | 44 |
| | $\theta_2$ | 1.9 | 17.5 | -29 | 53 |
| | $\theta_3$ | 1.5 | 17.4 | -29 | 51 |
| | $\theta_4$ | 0.2 | 17.3 | -30 | 48 |
| | $\theta_{5,6}$ | -3.1 | 17.5 | -33 | 40 |
| Foliage | B.A. ratio | -0.6 | 13.0 | -25 | 30 |
| | Regression on $d^2$ | -0.5 | 12.6 | -25 | 31 |
| | Log-log regression | | | | |
| | $\theta_1$ | 0.5 | 13.9 | -28 | 44 |
| | $\theta_2$ | 7.1 | 15.0 | -21 | 60 |
| | $\theta_3$ | 6.0 | 14.4 | -22 | 55 |
| | $\theta_4$ | 3.5 | 13.8 | -24 | 49 |
| | $\theta_{5,6}$ | -2.0 | 14.3 | -31 | 36 |

TABLE 4—The relative number of times each of seven methods gave the best estimate of plot weight and the relative sum of errors squared. In each case the "best" method is given a score of 100

| | | Best estimate | Sum of error squared |
|---|---|---|---|
| Basal area ratio | | 100 | 100 |
| Unweighted regression | | 53 | 115 |
| Log-log regression | | | |
| Uncorrected ($\theta_1$) | | 18 | 382 |
| Corrected | $\theta_2$ | 51 | 820 |
| | $\theta_3$ | 58 | 637 |
| | $\theta_4$ | 13 | 425 |
| | $\theta_{5,6}$ | 61 | 515 |

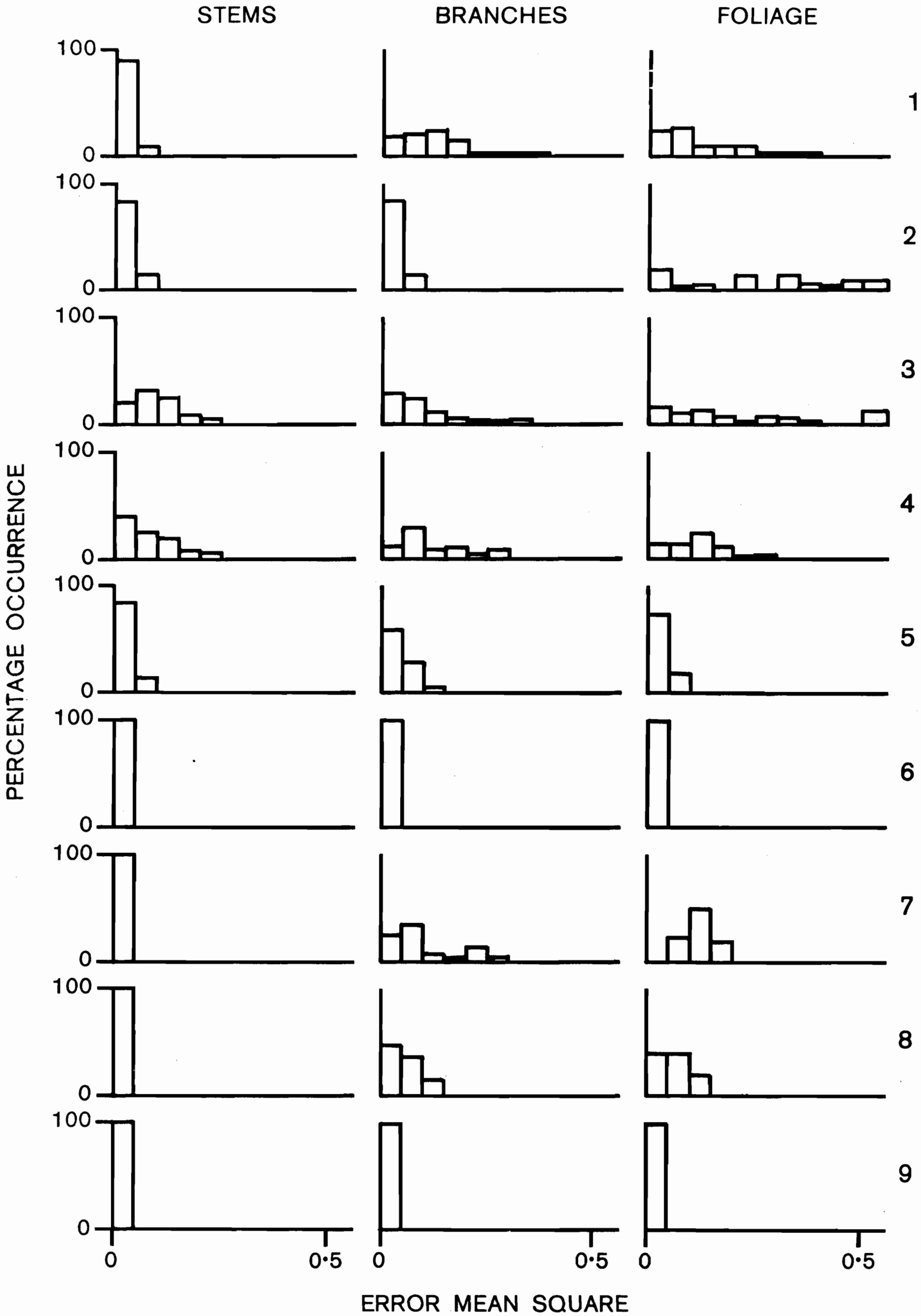


FIG. 1—The distribution of error mean square of log-log regressions based on 100 samples from each of nine plots. Numbers refer to plots as in Table 1.

to increase in the order stems, branches, and foliage and the distribution of values varied considerably from plot to plot. Examination of selected samples with relatively large error mean squares using the log-log regression indicated that these samples tended to be associated with large over-estimates especially of foliage weight using $\theta_2$ and $\theta_3$. It would appear that the "corrected" log-log procedures are particularly sensitive to "odd" sample trees. This problem would increase in importance when estimation was attempted of components such as dead branches or cones which tend to be poorly related to stem diameter. The failure of the various correction factors to improve estimates suggests that the assumptions underlying their derivation are not met.

Egunjobi (1976) has emphasised previously the relative differences in computational complexity among various methods of estimating stand weight. While noting that the basal area method is computationally the simplest he incorrectly stated that no confidence interval could be calculated. Calculation of the confidence interval is much more demanding than calculating the plot weight and, with the small numbers of sample trees often used in practice, the confidence intervals are biased (Madgwick 1981). Few authors publish confidence intervals but those based on log-log regressions appear unreliable (Madgwick & Satoo 1975).

In conclusion, it would appear that the basal area ratio method, the unweighted regression based on $d^2$, or the uncorrected log-log regression method all give similar estimates of plot weight when used in the context of research with a small number of sample trees per plot. The log-log method would appear to be more susceptible to the inclusion of unrepresentative sample trees. For computational simplicity the basal area ratio method would be best, followed by unweighted regression of weight on diameter squared. When using the basal area ratio method the work of Royall & Cumberland (1981) would suggest that a balanced sampling design is desirable.

## ACKNOWLEDGMENTS

Thanks are extended to those people who have allowed the use of their data for this study and to the reviewers for their helpful comments.

## REFERENCES

ANDO, T. 1962: Growth analysis on the natural stands of Japanese red pine (**Pinus densiflora** Sieb. et Zucc.). II. Analysis of stand density and growth. **Government Forest Experiment Station Bulletin 147:** 45–77.

BRADU, D.; MUNDLAK, Y. 1970: Estimation in normal linear models. **Journal of the American Statistical Association 65(329):** 198–211.

COCHRAN, W. G. 1963: "Sampling Techniques". John Wiley & Sons, New York. 413 p.

CROW, T. R. 1971: Estimation of biomass in an even-aged stand – Regression and mean tree techniques. Pp. 35–48 **in** Young, H. E. (Ed.) "Forest Biomass Studies". University of Maine Press, Orono, United States.

EGUNJOBI, J. K. 1976: An evaluation of five methods for estimating biomass of an even-aged plantation of **Pinus caribaea** L. **Oecologia Plantarum 11(2):** 109–16.

FLEWELLING, J. W.; PIENAAR, L. V. 1981: Multiplicative regression with lognormal errors. **Forest Science 27(2):** 281–9.

MADGWICK, H. A. I. 1971: The accuracy and precision of estimates of the dry matter in stems, branches and foliage in an old field **Pinus virginiana** stand. Pp. 105–12 **in** Young, H. E. (Ed.) "Forest Biomass Studies". University of Maine Press, Orono, United States.

——— 1981: Estimating the above-ground weight of forest plots using the basal area ratio method. **New Zealand Journal of Forestry Science 11(3):** 278–86.

MADGWICK, H. A. I.; SATOO, T. 1975: On estimating the above-ground weights of tree stands. **Ecology 56(6):** 1446–50.

OVINGTON, J. D.; MADGWICK, H. A. I. 1959: Distribution of organic matter and plant nutrients in a plantation of Scots pine. **Forest Science 5(4):** 344–55.

OVINGTON, J. D.; FORREST, W. G.; ARMSTRONG, J. E. 1968: Tree biomass estimation. Pp. 4–31 **in** Young, H. E. (Ed.) "Symposium on Primary Productivity and Mineral Cycling in Natural Ecosystems". University of Maine Press, Orono, United States.

PARDÉ, J. 1980: Forest biomass. **Forestry Abstracts 41(8):** 343–62.

ROYALL, R. M.; CUMBERLAND, W. G. 1981: An empirical study of the ratio estimator and estimators of its variance. **Journal of the American Statistical Association 76(373):** 66–77.

SATOO, T. 1973: Matter production of terrestrial communities (forests). Ecology lectures - 5A. Kuritsu Publishing Co. [In Japanese].

SWANK, W. T.; SCHREUDER, H. T. 1974: Comparison of three methods of estimating surface area and biomass for a forest of young eastern white pine. **Forest Science 20(1):** 91–100.