Jl. Rubb. Res. Inst. Sri Lanka (1978) **55**, 29—33

A GENERAL REGRESSION EQUATION FOR THE ESTIMATION OF LEAF AREAS

G. A. J. P. R. GOONASEKERA

SUMMARY

A general equation $A = kL^{\alpha} B^{\beta}$ *is suggested for the estimation of leaf area A using the length L and the maximum breadth B of the leaf. For a sample of Jak leaves, it is shown that the equations (i)* $A = kL$ *(ii)* $A = kB$ *(iii)* $A = kLB$ *and (iv)* $A = kL^n$ *which are the commonly used particular forms of the above equation, are inappropriate.*

INTRODUCTION

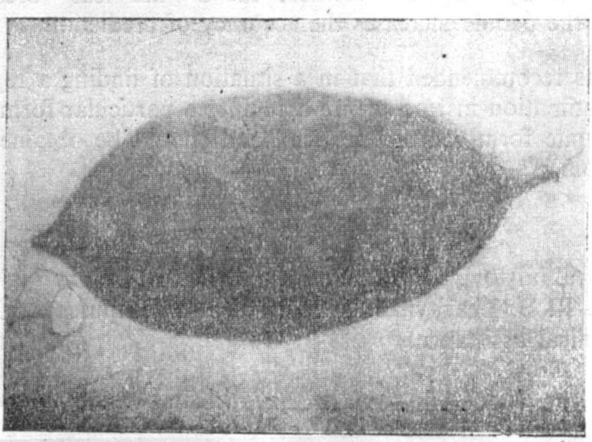
The uses of leaf area are manifold in biological research. Its importance may be as an index of growth or in studies of dry matter accumulation, carbohydrate metabolism, nutritional uptake, yield, quality, competition and disease.

The derivation of regression equations for the estimation of leaf areas has been studied for several decades (Baten, 1941). Since then many other authors (Owen, 1957; Reynolds, 1971; Lim & Narayan, 1972; Gupta & Dass, 1976) have presented different forms of equations, linear and curvilinear, for estimating leaf areas of various trees and grasses. Baten obtained two equations for bean leaflets using the length L and the maximum breadth B separately as the independent variables and tested for the significantly better predictor. Owen assumed an equation of the form area $A = kLB$ for sugar-beet leaves. This equation has been very extensively used in the literature. Reynold's assumption was $A = kX^n$. He estimated three different equations for cocoa leaves using the leaf length, breadth and dry matter weight as predictors. Least squares estimation procedure had been applied in all cases, after linearising by a logarithmic transformation in the case of Reynold's equations.

The purpose of this paper is to suggest a general equation $A = kL^{\alpha} B^{\beta}$ for leaf area prediction. Here, k, α and β are constants to be estimated. When (i) $\alpha = 1$ and $\beta = 0$, (ii) $\alpha = 0$ and $\beta = 1$, the equations used by Baten are obtained; (iii) $\alpha = \beta = 1$ gives the form assumed by Owen and (iv) $\alpha = n$, $\beta = 0$ gives Reynold's assumption.

MATERIALS & METHODS

Jak (*Artocarpus indicus*) leaves which are rather elliptic in shape, (see Figure) were randomly selected from two trees on different occasions and the pooled sample



of 300 leaves was used in the analysis. A translucent paper was placed on a leaf and its trace was marked. The length and the maximum breadth of this trace was measured using a ruler. A planimeter was used to obtain the area outlined on the paper. This procedure was repeated for each leaf. A single person was employed for this work.

The suggested equation was reduced to a linear form by a logarithmic transformation and the multiple regression computing procedures given by Rao (1952) were employed to estimate the constants (see appendix).

RESULTS

The units of measurements were centimeters for the lengths and breadths. The planimeter gave the area in square inches. The prediction equation in C.G.S. units was obtained to be $A = 0.47342\ L^{1.46186}\ B^{0.54697}$.

It is seen that the sample estimates of α and β add to 2.01 giving a dimensionally balanced equation. The coefficient of multiple correlation R = 0.9283. The standard error of the residuals (i.e. planimetered area — estimated area) was 1.4685. A test of the hypothesis $\alpha = \beta = 0$ (see appendix) gave an F value of 924.9 (for 2 and 297 degrees of freedom) showing that the variables L and B are useful in predicting A. Then the hypothesis that the coefficients of L and B are equal (*i.e.* $\alpha = \beta$) was tested and rejected at the very highly significant value of $F_{1,297} = 54.53$. A test was also carried out to find whether B is necessary when L is considered i.e. the hypothesis $\beta = 0$. The usefulness of B was indicated by the value 47.19 for the F-ratio. Similarly, the necessity of L was shown. These test results indicate that the forms of equations used by Baten, Owen and Reynolds should be statistically rejected for the sample of Jak leaf data considered here.

CONCLUSIONS

The equation introduced above has shown, for a sample of 300 Jak leaves, that the most appropriate equation is different from any of the normally used leaf area predicting regression equations. However, smaller samples of sizes, 60 and 120 subjected to the above analyses, indicated appropriateness of the widely used form A = kLB. This would be due to the reason indicated by Rao (1952) in his study of human skills, that a large collection of measurements may be necessary before anything definite can be said about the differences in the powers and B of the predictors. By a simple extension of the calculations, it would also be possible to test whether the inclusion of any extra variables, for example an equation of the form $A = kL\ B_1^{\beta_1}\ B_2^{\beta_2} \ldots\ldots$ where the B's are leaf-breadths at different distances from the petiole increases the accuracy of prediction.

It is thus recommended that in a situation of finding a regression equation for leaf area estimation instead of pre-assuming a particular form of the equation, the most adequate form for the data in hand should be obtained by assuming a general equation.

ACKNOWLEDGEMENTS

I am thankful to Dr. O. S. Peries and Dr. A. Perera for critically reading the paper, Mr. R. B. Ekanayake for assistance in computer work and Mr. L. T. Peiris for technical assistance.

REFERENCES

BATEN, W. D. (1941). How to determine which of two variables is better for predicting a third variable. *J.Am. Soc. Agron.* **33**.

GUPTA, S. K. and DASS, A. D. (1976). Determination of leaf area in Jute, *Indian J. Agron,* **21**, 160 - 161.

LIM, T. M. and NARAYAN, R. (1972). Estimation of the area of rubber leaves (*Hevea brasiliensis* Muell Arg.) using two leaflet parameters. *Exp. Agric.*, **8**, 311 - 313.

OWEN, P. C. (1957). Rapid estimation of the areas of the leaves of crop plants. *Nature,* **180**, 611.

RAO, C. R. (1952). Advanced Statistical Methods in Biometric Research. New York: Wiley & Sons, p. 103.

REYNOLDS, S. G. (1971). A note on the estimations of leaf areas of cocoa (*Theobroma cacao L*) from three leaf parameters. *Trop. Agric. Trinidad*, **48**, 177.

APPENDIX

The assumed regression equation

$$A = k_1 L^{\alpha} B^{\beta}$$

can be written as

$$Y = k + \alpha X_1 + \beta X_2$$

by the transformations

$$Y = \log_{10} A$$
$$X_1 = \log_{10} L \text{ and}$$
$$X_2 = \log_{10} B$$

For the leaf sample under consideration, the mean values are

$$\bar{Y} = 2.393986 \text{ sq. inches}$$
$$\bar{X}_1 = 2.646248 \text{ cms}$$
$$\bar{X}_2 = 2.079925 \text{ cms}$$

and the number of observations n = 300.

The corrected sums of products matrix for X_1, X_2 is

$$(S_{ij}) = \begin{pmatrix} 11.304030 & 5.319742 \\ 5.319742 & 5.427244 \end{pmatrix}$$

$$\text{where } S_{ij} = \sum X_i X_j - \frac{\sum X_i \sum X_j}{n}$$

The corrected sums of products of Y with X_1 and X_2 are, respectively,

$$Q_1 = 19.434599 \text{ and}$$
$$Q_2 = 10.745229$$

The inverse* of (S_{ij}) is

$$(C_{ij}) = \begin{pmatrix} 0.164213 & -0.160960 \\ -0.160960 & 0.342027 \end{pmatrix}$$

The least squares estimates of the regression coefficients and the constant term are

$$\hat{\alpha} = \sum_{j=1}^{2} C_{1j}\, Q_j = 0.164213 \times 19.434599 - 0.160960 \times 10.745229 = 1.461856$$

$$\beta = \sum_{j=1}^{2} C_{2j}\, Q_j = 0.160960 \times 19.434599 - 0.342027 \times 10.745229 = 0.546968$$

and

$$\hat{k} = \bar{y} - \hat{\alpha}\,\bar{X}_1 - \hat{\beta}\,\bar{X}_2 = -2.612104$$

The formula for the prediction of leaf area in C.G.S. units is

$$A = 0.47342\, L^{1.46186}\ \ B^{0.54697}$$

Test of the Hypothesis $\alpha = \beta = 0$

	D. F.	S.S.		M.S.	F
Regression	2	$\alpha Q_1 + \beta Q_2 =$	34.28789	17.143945	924.99
Residual	n–3=297	$\left(\sum y^2 - n\bar{y}^2\right) - \alpha Q_1 - \beta Q_2 =$	5.504689	0.0185343	
Total	n–1=299	$\left(\sum y^2 - n\bar{y}^2\right) =$	39.792579		

Test of the Hypothesis $\alpha = \beta = \theta$ (say)

	D.F.	S.S.		M.S.	F.
Deviation from equality	1	$\alpha Q_1 + \beta Q_2 - \theta Q =$	1.010696	1.010696	54.53
Residual	n–3=297	$\left(\sum y^2 - n\bar{y}^2\right) - \alpha Q_1 - \beta Q_2 =$	5.504689	0.0185348	
Total	n–2=299	$\left(\sum y^2 - n\bar{y}^2\right) - \theta Q =$	6.515384		

Where $Q = Q_1 + Q_2$ and

$$\theta = \frac{Q}{S_{11} + S_{22} + S_{12} + S_{21}}$$

Test of the Hypothesis $\beta = 0$

Variance $\beta = C_{22}\, \epsilon^2$

Where

$$\epsilon^2 = \frac{\left(\sum y^2 - n\bar{y}^2\right) - \alpha Q_1 - \beta Q_2}{n-3}$$

$$= 0.0185343$$

Then $$F_{1,\,n-3} = \frac{\beta^2}{C_{22}\, \epsilon^2}$$

i. e. $F_{1,297} = 47.194091.$

Test of the Hypothesis $\alpha = 0$

$$F_{1,297} = \frac{\alpha^2}{C_{11}\,\sigma^2} = \frac{1.4618556^2}{0.164213 \times 0.0185343} = 702.1435$$

* The inverse of a2 X 2 matrix $(M) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

is easily obtained as

$$(M^{-1}) = \frac{1}{(ad - bc)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$$