African Journal of Mathematics and Computer Science Research Vol. 3(6), pp. 79-82, June 2010
Available online at http://www.academicjournals.org/AJMCSR
ISSN 2006-9731©2010 Academic Journals

Full Length Research Paper

Computing leaf rectangularity index: An estimation problem when the parameter is a norm of a vector

Isaac Didi Essi

Statistics and Econometrics Modeling Group, P. O. Box 11518 (Main GPO) Near Rivers State Judiciary, Port Harcourt, Rivers State, Nigeria. E-mail: ddmetra_utibe@yahoo.com.

Accepted 21 April, 2010

A measure known as leaf rectangularity index (LRI) is estimated by means of bootstrap regression. The index, it is envisaged, will assist in discussing the geometry of leaf surfaces, if possible among different plants and across species. The study shows that one cannot obtain the point estimate of LRI before its interval estimate; rather, the interval estimate comes first before the point estimate. This paper gives the formal steps of obtaining both the point and interval estimate when the parameter is a norm of a vector.

Key words: Leaf rectangularity index, bootstrap regression, norm.

INTRODUCTION

The index and its relevance

Leaf area index (LAI), defined as leaf area per unit ground area, is an important canopy parameter that has attracted the attention of scientists for several decades now. Another parameter, leaf rectangularity index (LRI), is also useful in the study of plant and their species. The invention and eventual construction was first mentioned by Essi (2005). In Essi (2009), more work was done on the index including estimation under alternative error specifications by using multiplicative error model (MEM) and the additive error model (AEM). LRI is a measurable criterion for determining the extent to which a leaf surface departs from being a perfect rectangle. This study focuses on a formal step-by-step procedure for estimating LRI. There are a lot of problems and expectations arising from possible variation of the index over time and space. Can this index be affected by time? Can the same plant in different regions and climatic zones produce different indices? Finally can environmental degradation as we have in the Niger Delta of Nigeria affect the value of this index for a plant of interest? Finally, what is the most probable range within which LRI is expected to lie? Attempts for these questions are themselves foci for research and their answers may then constitute the importance of LRI.

Abbreviations: LRI, Leaf rectangularity index; **LAI,** leaf area index; **AEM,** additive error model; **MEM,** multiplicative error model.

THEORETICAL FRAMEWORK

The area A, of leaf surface with length L, and breadth B (omitting the error term) is given by

$$A = \theta_0 L^{\theta_1} B^{\theta_2} \quad (1)$$

Where θ_0, θ_1 and θ_2 are parameters to be determined. These parameters are the ones that will dominate the discussion for the rest of this paper. For a perfect rectangular surface with sides L and B, its area is

$$A^* = LB \quad (2)$$

On comparing Equations (1) and (2) and using dimensional analysis of elementary physics, a leaf surface reduces to a perfect rectangle surface if simultaneously we have

$$\theta_0 = 1, \quad \theta_1 + \theta_2 = 2 \quad (3)$$

The set of conditions in Equation (3) may be referred to as necessary conditions of rectangularity for an arbitrary surface A. We will keep this notion until a much better definition for leaf surface rectangularity is stated. Since any two numbers θ_1 and θ_2 can add to 2, we insist that the numbers must be such that θ_1/θ_2 = 1. Hence, we have a modified set of conditions as follows:

$$\theta_0 = 1,\ \theta_1 + \theta_2 = 2,\ \theta_1/\theta_2 = 1 \quad (4)$$

Consider the vector parameter η defined as

$$\eta = \begin{bmatrix} \alpha \\ \lambda \\ \gamma \end{bmatrix} = \begin{bmatrix} \theta_0 - 1 \\ 2-(\theta_1+\theta_2) \\ (\theta_1/\theta_2)-1 \end{bmatrix} \quad (5)$$

The quantities α , λ and γ are all dimensionless since θ_0 , θ_1 and θ_2 are dimensionless. Now the condition η = 0 is equivalent to the Euclidean norm $\|\eta\|$ = 0. Using Equation (5) it is easy to see that for a perfect rectangle, $\|\eta\|$ = 0. This implies that, a surface A tends to a rectangular surface A^* as $\|\eta\| \to 0$. $\|\eta\|$ can now be used to define a criterion for rectangularity which we choose to call leaf rectangularity index (LRI). LRI is defined as:

$$\text{LRI} = \|\eta\| \quad (6)$$

The problem is to find an estimate for $\|\eta\|$. Regression analysis can furnish estimates for θ_0 , θ_1 and θ_2 and their corresponding standard errors. However, the estimates of $\|\eta\|$ and its associated confidence interval are not possible by direct regression. One way to estimate it is by bootstrap resampling. Bootstrap involves resampling from a sample of size n with replacement and computing the estimate of parameter θ of interest using the replicated samples. The series of bootstrap estimate of θ , $\theta_1, \theta_2, ..., \theta_m$ can be used in estimating θ and finding its confidence interval, bias and variance or calibrating hypothesis tests involving θ. Since its invention by Efron (1979) and subsequent development by Efron (1982), Efron and Tibshirani (1986, 1993) bootstrap has witnessed great expansion in theory and application over the years. Some applications of the bootstrap include toxicology (Bailer and Oris, 1994), fisheries survey (Schmoyer et al., 1996); groundwater and air pollution modeling (Archer and Giovannoni, 1998; Cooley, 1997), chemometrics (Wehrens and Van der Linden, 1997), hydrology (Fortin et al., 1997), phylogenetics Newton (1996), spatial point patterns Solow (1989), ecological indices Dixon (2001), and multivariate summarizations (Pillar, 1999; Yu et al., 1998). Literature on the bootstrap is very wide, covering concepts, applications and theory. While Manly (1997) emphasizes applications, Chernick (1999), Davison and Hinkley (1997) and Efron and Tibshirani (1993) contain comprehensive coverage. The theory is well treated in Efron (1982), Hall (1992), LePage and Billard (1992) and Shao and Tu (1995). A wide range of papers stating applications are cited in Manly (1997) and Chernick (1999). The particular bootstrap method used in this work is bootstrap regression. There are generally two approaches of data for this type of regression. They are: (1) bootstrapping the observations, also called paired resampling and; (2) bootstrapping the residuals, also called residual resampling (Efron and Tibshirani, 1993; Chernick, 1999; Davison and Hinkley, 1997; Mittelhammer et al., 2000).

Consider a linear regression model

$$Y = XB + \varepsilon \quad (7)$$

Bootstrapping observations means drawing n rows from the rows of (Y, X) with replacement. Residual bootstrapping essentially involves four steps and is the one used in this study. Details are furnished in the methodology of this paper. Bootstrapping the residuals and the observations are not equivalent in small samples but are asymptotic (Efron and Tibshirani, 1986). The choice of bootstrap depends on both the goal and the type of regressor set x. If x is fixed, we use residual resampling and if it is stochastic we apply observation resampling.

METHODOLOGY AND DATA

The data on area A, length L and breadth B used in the study are that of sycamore leaves and they are found in Please (1987). We assume first, a multiplicative error model (MEM) of the form;

$$A = \theta_0 L^{\theta_1} B^{\theta_2} e^{u} \quad (8)$$

By setting lnA = Y, lnL = X_1 and lnB = X_2 we have

$$Y = XB + u \quad (10)$$

While Please (1987) used a sample of size 12, the simulated sample in the work is of size n = 30. The steps for the residual bootstrap regression are as follows:

1. The least squares (LS) estimate of $\mathrm{B}, \widehat{\mathrm{B}}$, and estimate of $u, \hat{u}$ are computed.
2. n errors from $\hat{u}$ with replacement was drawn and called $u_{(1)}$ it

Table 1. Bootstrap estimates of parameters α, λ, γ .

Parameter	Model	Pre-resampling estimate	Post-resampling estimate	Bias
θ_0	AEM	1.2842	1.0241 (0.1297)*	-0.2601
	MEM	1.0204	1.0331 (0.1344)	0.0127
θ_1	AEM	1.0805	1.0191 (0.0717)	-0.0614
	MEM	1.0438	1.0242 (0.0663)	-0.0196
θ_2	AEM	0.7318	0.8851 (0.1079)	0.1533
	MEM	0.8609	0.8770 (0.1053)	0.0161
α	AEM	0.2842	0.0241 (0.1297)	-0.2601
	MEM	0.0204	0.0331 (0.1344)	0.0127
λ	AEM	0.1877	0.0958 (0.0500)	-0.0919
	MEM	0.0953	0.0988 (0.0512)	0.0005
γ	AEM	0.4765	0.1766 (0.2231)	-0.2999
	MEM	0.2125	0.1924 (0.2207)	-0.0201

*Every number in brackets is a standard deviation for estimate before it.

was used to form a vector $Y_{(1)}$ of first n bootstrap observations using the equation;

$$Y_{(1)} = X\hat{B} + u_{(1)}$$

3. Using the replicated sample ($Y_{(1)}$, X) the LS estimate $\hat{B}_{(1)}$ was computed.
4. Steps 2 and 3 was repeated for m times to yield a set of m bootstrap estimates
$\hat{B}_{(1)}, \hat{B}_{(2)}, \dots, \hat{B}_{(m)}$. What follows is the use of the set of LS estimates from the m replicated samples to approximate the sampling distribution of $\hat{B}$ and hence that of $\hat{\theta}_0, \hat{\theta}_1 \, and \, \hat{\theta}_2$. Finally, estimate for η was obtained as defined in Equation (5) and a value for $\|\eta\|$ was also obtained. Following similar steps outlined above, consideration was given next to bootstrap regression with an additive error model in the form.

$$A = \theta_0 L^{\theta_1} B^{\theta_2} + u \qquad (10)$$

Estimation of $\|\eta\|$

1. We first of all estimate α , λ and γ and confidence intervals
$\hat{\alpha} \pm$ a, $\hat{\lambda} \pm$ b and $\hat{\gamma} \pm$ c

Where each of a, b and c is a constant derivable from the level of significance and of the experiment and standard deviation of the associated estimate.

2. Define vectors of lower and upper limits in the form

$$V_L = \left[\hat{\alpha} - a, \quad \hat{\lambda} - b, \quad \hat{\gamma} - c \right]$$

$$V_U = \left[\hat{\alpha} + a, \quad \hat{\lambda} + b, \quad \hat{\gamma} - c \right]$$

3. If the length of the vectors V_U and V_L are respectively l_1, l_2, then $\|\eta\|$ lies in the interval $\left[l_1, l_2 \right]$ with probability, 1 - α, where α is the level of significance of the computation.

4. The next problem is, ''How do we get a point estimate for $\|\eta\|$ and its standard deviation''?

5. The interval $\left[l_1, l_2 \right]$ is equivalent to $\|\hat{\eta}\| \pm \varepsilon$ where $\|\hat{\eta}\| = (l_1 + l_2)/2$ and

$$\varepsilon = \left| \frac{(l_1 + l_2)}{2} - l_1 \right| = \left| l_2 - \frac{(l_1 + l_2)}{2} \right| = \frac{1}{2} \left| l_2 - l_1 \right| > 0$$

6. If $z_{\alpha/2}$ is the standard normal variate corresponding to the significance level α then

$$\varepsilon = s z_{\alpha/2}$$

Where s is the standard deviation of $\|\hat{\eta}\|$

EMPIRICAL RESULTS

The results of bootstrap estimation are reported in Tables 1 and 2. From Table 1, a 95% confidence interval for the

Table 2. Estimates of leaf rectangularity for AEM and MEM.

Model	Estimated LRI	95% confidence interval for LRI	Standard deviation for estimated LRI
AEM	0.5602	0.5602 ± 0.2894 (0.2708 , 0.8495)	0.1447
MEM	0.5380	0.5342 ± 0.1962 (0.3380, 0.7303)	0.0981

set $(\alpha, \lambda, \gamma)$ for AEM is (0.2842 ± 0.2594, 0.0958 ± 0.1000, 0.1766 ± 0.4462) which gives the vectors of lower and upper bounds of V_L = (0.0248, -0.0042, -0.2696) and V_U = (0.5436, 0.1958, 0.6228). The length of the vectors are respectively l_1 = 0.2708 and l_2 =0.8495. Therefore $\|\eta\|$ lies in the interval $\left[l_1, l_2\right]$ = (0.2708, 0.8495) with probability of 0.95. Similarly, a 95% confidence interval for the set $(\alpha, \lambda, \gamma)$ for MEM is (0.0331 ± 0.2688, 0.0988 ± 0.1024, 0.1924 ± 0.4414) which gives the vectors of lower and upper bounds of V_L = (-0.2357, -0.0036, -0.2422) and V_U = (0.3019, 0.2012, 0.6338). The length of the vectors are respectively l_1 = 0.3380 and l_2 = 0.7303. Therefore $\|\eta\|$ lies in the interval $\left[l_1, l_2\right]$ = (0.3380, 0.7303) with probability of 0.95. LRI estimates with a 95% confidence interval are shown in Table 2.

Conclusion

Leaf rectangularity index (LRI) is defined as a norm of a vector. The confidence interval for LRI is calculated first before its point estimate and standard deviation. This is by backward computation. The common experience is having point estimate and standard deviation before confidence interval. This paper submits that the normal order of procedure breaks down when the parameter estimated is a norm.

REFERENCES

Archer G, Giovannoni JM (1998). .Statistical analysis with bootstrap diagnostics of atmospheric pollutants predicted in the APSIS experiment. Water, Air Soil Pollu., 106: 43-81.

Bailer AJ, Oris JT (1994). Assessing toxicity of pollutants in aquatic systems. pp. 5-9.

Chernick MR (1999). Bootstrap Methods: A Practioner's Guide, Wiley,New York. pp. 76-78.

Davison AC, Hinkley DV (1997). Bootstrap Methods and their Applications, Cambridge University, Cambridge. pp. 261-266.

Essi ID (2005). "On Constructing Leaf Rectangularity Index by Bootstrap Regression: AMSE J. 66(6): 13-22.

Essi ID (2009). "Computing Leaf Rectangularity Index under Alternative Error Specifications, AMSE J. C. 70(1): 67-79.

Efron B (1979). Bootstrap Methods: Another Look at the Jackknife, Ann. Stat. 7(1): 1-26.

Efron B (1982). The Jacknife, the Bootstrap and other Resampling Plans, SIAM, Philadelphia.

Efron B, Tibshirani R (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. Stat. Sci. 1: 54-77.

Efron B, Tibshirani RJ (1993). An Introduction to the Bootstrap, Chapman and Hall, New York. pp. 113-115.

Fortin V, Bernier J, Bobee B (1970). Simulation, Bayes, and bootstrap in statistical hydrology. Water Resources Res. 3: 439-448.

Hall P (1992). *The Bootstrap and Edgeworh Expansion,* Springer-Verlag, New York, pp. 86-100.

LePage R, Billard L (1992). Exploring the Limits of Bootstrap, Wiley, New York. pp. 25-50.

Manly BFJ (1997). Randomization, Bootstrap and Monte Carlo Methods in Biology, 2nd Edition, Chapman and Hall, London. pp. 22-68.

Mittelhammer RC, Judges GD, Miller DJ (2000). Econometric Foundations, Cambridge University Press. pp. 724-730.

Newton MA (1996). Bootstrapping Phylogenies: Large Deviations and Dispersion Effects *Biometrika.* 83: 315-328,

Pillar VD (1999). The bootstrap ordination examined. J. Veg. Sci. 10: 895-902.

Please NW (1987). Advanced Statistical Analysis on BBC Microcomputers, Edward Arnold, London. pp. 38.

Schmoyer RL, Beauchamp JJ, Brandt CC, Hoffman FO JR (1996). Difficulties with Lognormal model in mean estimation and testing, Environ. Ecol. Stat. 3: 81-97.

Shao J, Tu D (1995). The Jackknife and Bootstrap Springer, New York. pp. 77-101.

Solow AR (1989). Bootstrapping Sparsely Sampled Spatial Points Patterns. Ecol. 70: 379-382.

Wehrens R, Van Der Linden WE (1997). Bootstrapping principal component regression models. J. Chemom. 11: 157-171.

Yu CC, Quinn JT, Dufournaud CM, Harrington JJ, Rogers PP, Lohani BN (1998). Effective dimensionality of environmental indicators: A principal components analysis with bootstrap confidence intervals, J. Environ. Manage. 53: 101-119.