Comparison of Techniques Used for Adjusting Biomass Estimates by Double Sampling

JAVED AHMED, CHARLES D. BONHAM, AND WILLIAM A. LAYCOCK

Abstract

This paper compares the ratio and regression estimator procedures for adjusting ocularly estimated plant species biomass in different sizes and shapes of plots. The study was conducted in northeastern Colorado on shortgrass rangeland dominated by blue grama *(Bouteloua gracilis)*. No significant differences were found in clipped plant biomass in 4 quadrat sizes between 0.18 and 0.50 m^2 and 2 shapes, circular and angular quadrats. For double sampling, the scatter plots of data strongly indicated a linear relationship through the origin for estimation and clipping. There were no significant differences between the adjusted mean weights by use of regression with and without intercept. The intercept was not significantly different from zero. Interpretation of correlation coefficient and variance of regression estimate with no intercept becomes difficult because the regression is forced through zero. Therefore, it is helpful to use regression with intercept. In the present study, estimates of both green and dry weights by ratio and regression estimation were comparable. Regression estimation is a minimum variance estimation comparable to ratio estimation even when the assumption of homoscedasticity is not true.

The single factor of greatest importance to range management is an accurate appraisal of the volume of forage available. The task is difficult simply because forage varies in the weight of plant material produced by each species in a highly variable environment. Because all the forage cannot be harvested and weighed, we must obtain a reasonable estimate of the actual weight by sampling. Sampling is often not an easy task because data must be obtained within a span of 1 or 2 weeks so that growth differences are minimized. On summer ranges, the information on residual forage is obtained after the livestock have been removed and before snow makes the sampling impossible.

If estimation of biomass for individual species is needed, then available funds and personnel may pose limiting constraints. In order to overcome some of these problems, a weight estimate method was designed by the personnel of the Intermountain Forest and Range Experiment Station during the summer of 1936 (Pechanec and Pickford 1937). The data are collected in two phases. In the first phase, the desired factor (x_i) is measured by some indirect method such as ocular estimation. In the second phase, the desired factor is measured both directly and indirectly. The indirectly estimated values (x_i) in the first phase are then adjusted by developing a mathematical relationship between the direct estimates (y_i) and indirect estimates (x_i). The sample size in the first phase is usually large compared with the sample size in the second phase. The sample in the second phase is usually a random subsample from the first phase but it may be drawn independently.

The mathematical procedures used for adjusting the indirect estimates are linear regression and ratio estimation. The theory of linear regression requires the assumption that the population regression of y on x is linear, that the residual variance of y about the regression line is constant (homoscedasticity), and that the population is infinite. No assumptions are made about the line passing through the origin. In the ratio estimation procedure, assumption of homoscedasticity need not be made but the estimator works well when $\sigma^{2y_i} \propto x_i$. In many ways, the procedure is analogous to fitting a linear relationship between y and x which passes through the origin. When we are trying to decide what kind of estimate to use, a graph in which y_i is plotted against x_i is helpful. If the graph shows a straight line, relationship through the origin and variance of points y_i about the line seem to increase proportionally to x_i, then the ratio estimate is better than the least squares estimate for regression (Cochran 1963).

During the summer of 1979, data were collected by double sampling from the Central Plains Experimental Range (CPER) near Nunn, Colo. CPER is administered by the Agricultural Research Service, USDA. The objective of this study was to compare the ratio and regression estimator procedures for adjusting the ocularly estimated species biomass.

The technique of double sampling and its statistical aspects are described in most sampling technique textbooks. The technique is also described by the National Research Council (1962) and by Schumacher and Chapman (1948). Pechanec and Pickford (1937) gave a detailed outline for the training of personnel for double sampling. Burton (1944) reported the ability of different personnel to estimate the yield in plots. Double sampling determination of herbage production in different vegetation types and the results were discussed by Pickford (1940), Wilm et al. (1944), Ragsdale (1956), Hillmon (1959), Hughes (1959), Shoop and McIlvain (1963), and Tadmor et al. (1975). Double sampling was found desirable for extensive browse inventories by Carhart and Means (1941), Schawan and Swift (1941), Dasman (1948), and Blair (1959). Abstracts of the double sampling technique are given by Morris (1967). Statistical aspects of double sampling were reviewed by Francis et al. (1979).

Optimum allocation of resources to direct and indirect methods of estimation is well defined by Cochran (1963) for a single factor under study and for a given sampling procedure. The optimum allocation formulations described by Schumacher and Chapman (1948) and Wilm et al. (1944) are similar to those described by Cochran (1963). More recently, Ahmed (1980) and Ahmed and Bonham (1980) described a technique for optimum allocation in multivariate double sampling for biomass estimation.

Collection of Data

Size and Shape of Sampling Unit

The most suitable size and shape of a plot is dependent upon the type of vegetation being studied. Francis et al. (1979) recommended a circular 0.5-m^2 quadrat for determining the production of major species. A plot 30 × 60 cm was considered the smallest area that would allow for an adequate sampling of the major species. Between these 2 extremes, 4 sizes (0.18, 0.25, 0.32, and 0.50 m^2) and 2 shapes (circular and angular) were compared in our

Table 1. Regression analysis (with and without intercept) of clipped green weights (y) of plants on ocularly estimated weights (x).

Species	Regression with intercept			Regression without intercept[1]	
	Regression equation	S_{yx}	r^2	Regression equation	S_{yx}
Blue grama	y = −0.01 + 1.17 x	2.48	.942	y = 1.17 x	2.47
Sand dropseed	y = 1.49 + 1.12 x	6.40	.811	y = 1.19 x	6.44
Threeawn	y = 0.70 + 0.95 x	2.40	.843	y = 1.01 x	2.46
Scarlet globemallow	y = −0.11 + 1.05 x	1.30	.909	y = 1.04 x	1.30
Goosefoot	y = −0.18 + 1.08 x	1.70	.879	y = 1.06 x	1.70
Other species	y = 1.72 + 0.95 x	5.53	.849	y = 1.00 x	

[1]Correlation coefficient in regression without intercept is difficult to interpret. It is, therefore, not reported.

study. The 4 sizes were randomly placed in pairs of 20 randomly located, 30-m-long transects. Each pair, circular and angular quadrat, was randomly placed 2 m apart on either side of the transect. Grasses and sedges, and forbs were clipped at ground level. The time required to clip each plot was also recorded. The plants were oven dried to a constant weight and the dry weight was recorded.

Double Sampling

We selected 3 dominant grasses, *Bouteloua gracilis* (Willd. ex H.B.K.) Lag. ex Griffiths (blue grama), *Sporobolus cryptandrus* (Torr.) A. Gray (sand dropseed), and *Aristida* spp. (Threeawn), mainly *A. longiseta* (Steud.), and 2 important forbs, *Sphaeralcea coccinea* (Pursh) Rydb. (scarlet globemallow) and *Chenopodium* spp. (goosefoot), and estimated standing biomass by the double sampling technique. Species of *Aristida* and *Chenopodium* were estimated together. Any other species found in the quadrat were estimated in one category termed "other species." A ratio of 3:1 between ocularly estimated plots and clipped plots was adopted for data collection. In order to minimize differences due to growth and phenological stages, the sampling was done from the last week of July, when the plant growth is at its peak, to the second week of August, when the plants begin dispersing seed and drying. During this period, a team of 2 persons was able to clip 200 quadrats and ocularly estimate 600. Ocular estimates were not compared with the clipped weights during the day's work. The person doing ocular estimation also had no prior knowledge of which quadrat would be clipped. These precautions were taken to minimize bias. For details of this procedure, the reader is referred to Ahmed (1980). Each species was clipped at ground level, placed in a separate paper bag, and oven dried.

Results

The oven-dry weights were converted to a m² basis and a two-way analysis of variance was done for the biomass of grasses and for total biomass. In both cases, there were neither significant differences between the sizes and shapes nor a significant interaction between size and shape ($P<0.05$). There was no significant difference between the mean clipping time of circular and angular plots of the smallest size ($P<0.05$). Quadrats larger than 0.50 m² and smaller than 0.18 m² should have been compared to find an optimum size and shape, but because optimum quadrat size was not the objective of the present study, a circular plot 0.18 m² in size was selected for collection of data. A circular plot was preferred over the angular plot because there is less linear distance in a circular plot and fewer decisions need to be made about what is in or out of the plot (Cook and Bonham 1977).

Scatter plots of the data are given in Figure 1 (A-F). These plots strongly suggest a linear relationship through the origin. The variance of the points y_i about the line appears to increase as x_i increases but not proportionally to increases in x_i.

Estimation by Using Regression with and without Intercept

The results of regression of clipped green weights (CGW) and clipped dry weights (CDW) on ocularly estimated weights (OEW) are summarized in Tables 1 and 2, respectively. In all cases, the intercept was not significantly different from zero ($P<0.01$). Also, there were no significant differences between the adjusted means of the 2 methods ($P<0.01$).

Estimation with Regression and with Ratio Estimators

OEW's were adjusted by using both the regression with intercept and ratio estimation procedures. The results of analysis are given in Table 3. In the case of regression with no intercept, the interpretation of the correlation coefficient and variance of the mean became difficult because the regression is forced through the origin. Therefore, this procedure was not compared with ratio estimation. The standard error of the regression and ratio were calculated by the following equations (Cochran 1963):

$$\text{SE}(\hat{\bar{y}}_{lr}) = \frac{s_y^2 (1 - r_{yx}^2)}{n} + \frac{r_{yx}^2 \, s_y^2}{n'} \qquad (1)$$

$$\text{SE}(\hat{\bar{y}}_r) = \frac{s_y^2 - 2\hat{R} \, s_{yx} + \hat{R}^2 \, s_x^2}{n} + \frac{2\hat{R} \, s_{yx} - \hat{R}^2 \, s_x^2}{n'} \qquad (2)$$

where the subscripts lr and r denote linear regression and ratio estimation, respectively. The estimated means obtained by both procedures are identical; however, the coefficient of variation of the mean with regression estimation for both CGW's and CDW's is approximately 50% compared with the estimated means obtained with the ratio estimation procedure.
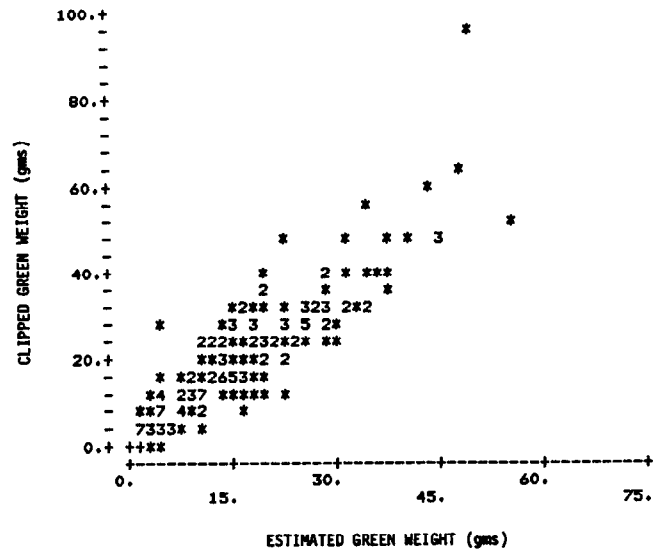
Discussion

In this study, no significant differences were found between

Table 2. Regression analysis (with and without intercept) of clipped dry weight (y) of plants on ocularly estimated weights (x).
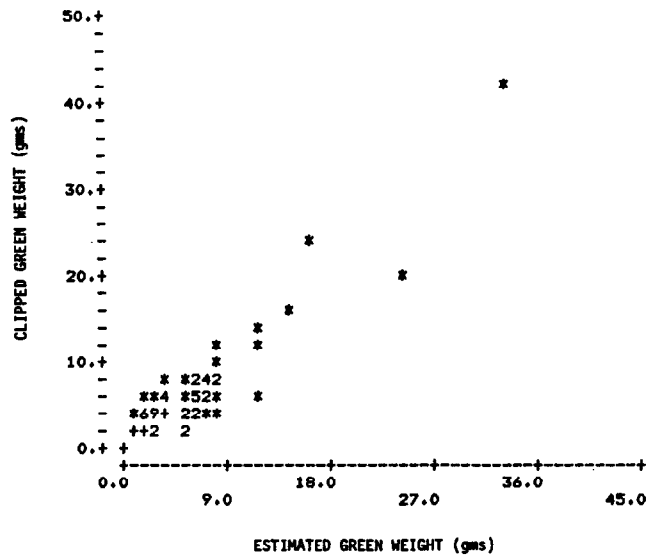
Species	Regression with intercept			Regression without intercept[1]	
	Regression equation	S_{yx}	r^2	Regression equation	S_{yx}
Blue grama	y = −0.01 + 0.69 x	1.75	.920	y = .69 x	1.75
Sand dropseed	y = 1.26 + 0.63 x	3.99	.776	y = .682 x	4.05
Threeawn	y = 0.44 + 0.64 x	1.68	.833	y = .677 x	1.71
Scarlet globemallow	y = −0.06 + 0.50 x	0.69		y = .49 x	
Goosefoot	y = 0.12 + 0.26 x	0.51	.829	y = .273 x	0.52
Other species	y = 1.94 + 3.19 x	2.98	.686	y = .382 x	3.39

[1]Correlation coefficient in regression without intercept is difficult to interpret. It is, therefore, not reported.
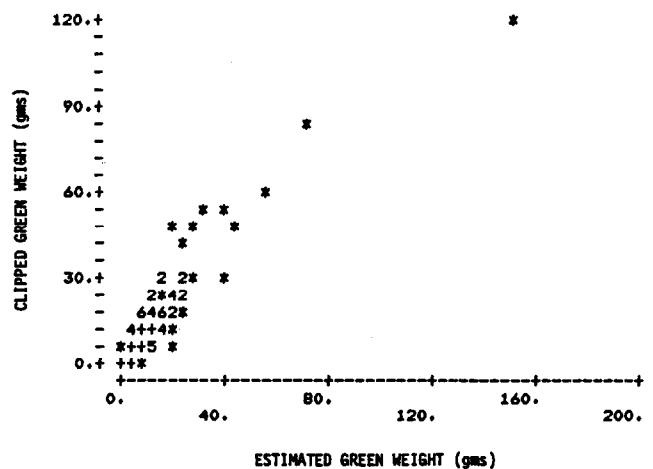
Fig. 1. *Scatter plot of clipped green weight of plants versus estimated weight of plants (g/1800 m²). Numbers indicate data point overlap.*

Table 3. Regression analysis (with and without intercept) of blue grama. Clipped weight (y) vs. estimated weight (x).

Regression equation	$\sigma y \cdot x$	r^2	Due to	D.F.	SS	MS	Fo	P(F>Fo)
			Green weight					
y = 0.01 + 1.17 x	2.48	94.2	Regression	1	20035.6	20035.6	3284.5	<.001
			Residual	198	1216.4	6.1		
			Total	199	21252.5			
y = 1.17 x	2.47		Regression	1	24453.5	24453.5	4008.8	<.001
			Residual	198	1216.5	6.1		
			Total	199	25670.0			
			Dry weight					
y = -0.01 + 0.69 x	1.75	92.0	Regression	1	6997.2	6997.2	2286.7	<.001
			Residual	198	606.7	3.1		
			Total	199	7603.9			
y = 0.69 x	1.75		Regression	1	8537.3	8537.3	2799.1	<.001
			Residual	199	606.7	3.1		
			Total	200	9144.0			

quadrat sizes ranging from 0.18 to 0.50 m² of circular and angular shapes. Quadrat sizes greater than 0.50 m² and smaller than 0.18 m² were, however, not tested because the objective of this study was not to find the optimum size of the plot.

Tadmor et al. (1975) reported a nonlinearity in their data. The scatter plots of the data obtained in this study do not indicate any nonlinearity. Francis et al. (1979) also did not find any nonlinearity in 1970 Pawnee Site data. A curvilinear plot would imply that after estimated weights (x) reach a certain value, the rate of change in clipped weights (y) with respect to change in x decreases. It does not seem logical that rate of change of ocular estimates (x_i) in the upper range of clipped weights (y_i) will consistently increase. The relationship between ocular estimates and clipped estimates, however, can vary from person to person. In the plot data of Tadmor et al. (1975), there were not enough observations in the upper range of x_i and y_i to conclusively prove nonlinearity. The results of this study and results reported by Francis et al. (1979) strongly supported a linear relationship in the double sampling data.

Blair (1959) considered the ratio estimation more valid than regression estimation for two reasons: (1) many estimates were zero or exceptionally low and use of the ratio estimator avoided the anomaly of predicting a positive or negative dry weight for a sample quadrat with no plants; and (2) the plots of data strongly suggested that the assumptions of homoscedasticity were not true. As a result of computer simulations, Francis et al. (1975) found that the regression estimate was universally the minimum variance estimation, but they suggested that for a species of relatively low biomass, either the regression or ratio technique might be employed. In any sampling procedure, the precision of individual observations is of much less importance than estimation of the population mean of the desired factor within certain specified limits of precision. Therefore, predicting a positive value for a zero or a negative value for low individual values is of little importance.

In the present study, the estimates of green and dry weight by ratio and regression estimation were comparable. For example, the estimated mean green and dry weights of scarlet globemallow were (2.36 g)·(.18 m⁻²) and (1.10 g)·(.18 m⁻²) (Table 3). The regression equation for adjusting the estimated biomass of scarlet globemallow had a negative intercept yet the estimated means obtained from both regression and ratio estimation procedures were comparable. The same was true of goosefoot. For all species, the coefficient of variation of the mean was ≤50% when regression estimation was compared with ratio estimation.

Cochran (1963) stated that in surveys in which the regression of y or x is thought to be approximately linear, it was helpful to use y_{1r} without having to assume exact linearity or constant residual variance. The plots of data we collected indicated a linear

relationship but assumptions of homoscedasticity did not appear to be true. Although the assumption of homoscedasticity was violated in almost all cases, the regression estimate was the technique that provided minimum variance estimation of the 2 employed. These results were in agreement with the simulation results of Francis et al. (1975). It is, therefore, recommended that for all practical purposes the regression estimation will give better results than ratio estimation without testing the assumption of homoscedasticity of the anomaly of predicting zero or negative values. If the predicted mean is a negative value it may be corrected to zero.
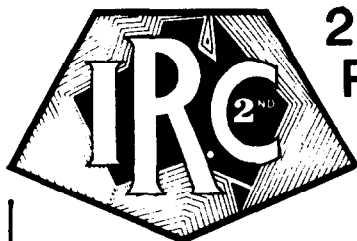
Conclusions

1. No significant differences were found in plant biomass in 4 quadrat sizes between 0.18 and 0.50 m² and 2 shapes, circular and angular quadrats. There was also no significant interaction between sizes and shapes.

2. The scatter plots of data strongly indicated a linear relationship through the origin for estimation and clipping.

3. There were no significant differences between the adjusted mean weights by using regression with intercept and by using regression with no intercept. The intercept was not significantly different from zero. Because the interpretation of correlation coefficient and variance of regression estimate with no intercept becomes difficult, it is helpful to use regression with intercept.

4. Regression estimation is a minimum variance estimation comparable to ratio estimation even when the assumption of homoscedasticity is not true.

Literature Cited

Ahmed, J. 1980. Optimum allocation in multivariate double sampling for biomass estimation. Ph.D. Diss., Colorado State Univ., Fort Collins.

Ahmed, J., and C.D. Bonham 1980. DUBSAM. Algorithm and computer program for optimum allocation in multivariate double sampling for biomass estimation. Range Sci. Dep. Range Sci. Series No. 33. Colorado State Univ., Fort Collins.

Blair, R.N. 1959. Weight techniques for sampling browse production on deer ranges. In: Techniques and methods of measuring understory vegetation, p. 26-31. Proc. Symp. at Tifton, Georgia, October, 1958.

Burton, G.W. 1944. Estimating individual forage plant yields. J. Amer. Soc. Agron. 26:709-712.

Carhart, A.H., and H. Means. 1941. Forage weight per acre method for appraisal of herbivorous animal requirements. Colo. Game and Fish Comm., Pittman-Robertson Deer-Elk Survey No. 5:29-51.

Cochran, W.G. 1963. Sampling techniques. John Wiley & Sons, Inc., New York.

Cook, C.W., and C.D. Bonham. 1977. Techniques for vegetation measurements and analysis for pre- and post-mining inventory. Range Sci. Dept. Range Sci. Series No. 28., Colorado State Univ., Fort Collins.

Dasmann, W.P. 1948. A critical review of range survey methods and their applications to deer range management. California Fish and Game 34:189-207.

Francis, R.C., G.M. Van Dyne, and B.K. Williams. 1979. An evaluation of weight estimation double sampling as a method of botanical analysis. J. Environ. Manage. 8:55-72.

Hilmon, J.B. 1959. Determination of herbage weight by double-sampling: Weight estimate and actual weight. p. 20-25. *In:* Techniques and methods of measuring understory vegetation. Proc. Symp. at Tifton, Georgia, October, 1958.

Hughes, R.H. 1959. The weight-estimate method in herbage production determinations. p. 17-19. *In:* Techniques and methods of measuring understory vegetation. Proc. Sym. at Tifton, Georgia, October, 1958.

Morris, M.J. 1967. An abstract of statistical methods in grasslands research. USDA Forest Serv. Misc. Pub. No. 1030.

National Research Council 1962. Range research: Basic problems and techniques. Nat. Acad. Sci. Res. Council Pub. No. 890. Washington, D.C.

Pechanec, J.F., and G.D. Pickford. 1937. A weight estimate method for determination of range or pasture production. J. Amer. Soc. Agron. 29:894-904.

Pickford, G.D. 1940. Range survey methods in western United States. Imp. Bur. Plant Genet. Herb. Rev. 8:1-12.

Rangsdale, D.W. 1956. Comparison of certain methods of determining utilization of range forages. M.S. Thesis, Colorado State Univ., Fort Collins.

Schumacher, F.X., and R.A. Chapman. 1948. Sampling methods in forestry and range management. Duke Univ. School of Forest Bull. No. 7.

Schwam, H.E., and L. Swift. 1941. Forage inventory methods with special references to big game ranges. Trans. Sixth North Amer. Wildlife Conf., p. 118-125.

Shoop, M.C. and E.H. McIlvain. 1963. The micro-unit forage inventory method. J. Range Manage. 16:172-179.

Tadmor, N.H., A. Brieghtt, I. Noy-Meir, R.W. Benjamin, and E. Eyel. 1975. An evaluation of the calibrated weight-estimate method for measuring production in annual vegetation. J. Range Manage. 28:65-69.

Wilm, H.G., D.F. Costello, and G.E. Klipple. 1944. Estimating forage yield by the double sampling method. Amer. Soc. Agron. J. 36:194-203.