CORE
Metadata, citation and similar papers at core.ac.uk
Provided by UNL | Libraries

University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

The Prairie Naturalist
Great Plains Natural Science Society

12-2015

Estimating Herbaceous Biomass of Grassland Vegetation Using the Reference Unit Method

Eric D. Boyda

Jack L. Bulter

Lan Xu

Follow this and additional works at: https://digitalcommons.unl.edu/tpn

 Part of the Biodiversity Commons, Botany Commons, Ecology and Evolutionary Biology Commons, Natural Resources and Conservation Commons, Systems Biology Commons, and the Weed Science Commons

This Article is brought to you for free and open access by the Great Plains Natural Science Society at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in The Prairie Naturalist by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

The Prairie Naturalist 47:73–83; 2015

Estimating Herbaceous Biomass of Grassland Vegetation Using the Reference Unit Method

ERIC D. BOYDA, JACK L. BUTLER[1], AND LAN XU

Department of Natural Resource Management, South Dakota State University, Brookings, SD 57007, USA (EDB, LX)
US Forest Service, Rocky Mountain Research Station, Forest and Grassland Research Laboratory, 8221 S. Highway 16, Rapid City, SD 57702, USA (JLB)

ABSTRACT Aboveground net primary production provides valuable information on wildlife habitat, fire fuel loads, and forage availability. Aboveground net primary production in herbaceous plant communities is typically measured by clipping aboveground biomass. However, the high costs associated with physically harvesting plant biomass may prevent collecting sufficient data to account for natural spatial and temporal variability of vegetation at a landscape scale. Various double-sampling techniques have been developed to increase sample size while reducing cost. We applied a biomass estimation technique previously developed for estimating shrub biomass using representative samples or "reference units" to estimate herbaceous grassland biomass. Our reference units consisted of major grass species and functional groups that involved combining species with similar origin (native vs. introduced) and life-form characteristics. This study was conducted in 2010 on and around prairie dog colonies on the Buffalo Gap National Grassland in southwestern South Dakota, which provided a range of plant communities for testing the method. Results of the study demonstrate that reference units can provide accurate and precise estimates of herbaceous plant biomass on grasslands, including multi-species functional groups. Twenty-five of 26 double sampling calibrations were validated by time with no changes in observer estimation trends over the sampling season. Use of the reference unit method was consistent among observers and is a viable option for estimating and monitoring herbaceous grassland aboveground biomass.

KEY WORDS ANPP, biomass, calibrating estimates, double sampling, mixed grass prairie, observer experience.

Aboveground net primary production (ANPP) is the total amount of plant biomass accumulated over a specific time period. Estimates of ANPP are an important metric in assessing wildlife habitat, fire fuel loads, forage availability, and ecological relationships and processes. A single harvest of green and current year dead plant materials at peak standing crop is the easiest and most commonly used approach in determining ANPP. Frequent harvests of standing crop through the growing season to account for plant material that may have senesced and disappeared increases the accuracy of estimating ANPP; however, frequent clipping, sorting, drying, and weighing biomass is labor intensive (Wilm et al. 1944, Haydock and Shaw 1975, Reese et al. 1980, Catchpole and Wheeler 1992). Constrained by budget and time, estimating ANPP at large spatial and temporal scales often is inadequate, especially in heterogeneous landscapes. The continuous need for economical and accurate measurements of biomass that can be repeated with precision has led to the progressive development of various techniques that have less precision at the plot level but allow for larger sample sizes that can produce large-scale ANPP estimates.

The weight-estimate method as described by Pechanec and Pickford (1937) is one of the first techniques utilized for estimating vegetation biomass and it created the foundation of double sampling. Double sampling vegetation biomass involves calculating a relationship, usually using regression analysis, between a variable that costs less to measure (indirect sample) and biomass that is accurately and precisely measured by clipping and weighing the collected biomass (direct sample; Catchpole and Wheeler 1992). With the weight-estimate method, an observer, through repeated clipping and weighing of vegetation from field plots, develops the ability to predict fresh weights (plant biomass that has not been dried) of unclipped plots using four weight units (10, 20, 50, and 100 g) as a reference. Seasonal, local, and annual variations in moisture content are periodically checked by clipping and comparing fresh weights to dry weights to determine percent moisture, which is then subtracted to produce an estimated dry weight for the plot. Pechanec and Pickford (1937) found that relative differences between estimated and actual weights were consistently below 10 percent. The weight-estimate method has been shown to accurately estimate single species biomass and requires little laboratory time outside of sampling because only about 10 to 20% of the total number of samples need to be dried and weighed to calculate the percent dry biomass for various species and sample harvest times (Pechanec and Pickford 1937, Shoop and McIlvain 1963). However, because it can take up to a week or more to develop consistency and accuracy in estimating biomass, high variability of fresh weight moisture can affect results (Pechanec and Pickford 1937, Tadmor et al. 1975). Further, observer bias can affect estimates when mental fatigue and attitude inhibit the retention of mental images of estimated units (Hutchings and Schmautz 1969).

[1] Corresponding author email address: jackbutler@fs.fed.us

Under the assumption that relative weight is easier to estimate than absolute weight (Hutchings and Schmautz 1969, Haydock and Shaw 1975, Reese et al. 1980), plots can be directly compared to a clipped sample or reference. Maintaining a set of representative clippings to compare to unclipped plots may be more efficient than using repeated experience of clipping and weighing samples in the field to develop a personal capacity for visually estimating biomass. This reference or "reference unit" method, also called the 'Adelaide' technique, was developed in Australia specifically for estimating biomass of shrubs, because the investigators determined that estimation methods developed for grasslands were unsuitable for estimating shrub forage (Andrew et al. 1979, 1981). These authors reported high correlation coefficients between estimated and actual biomass for two species of shrubs (*Atriplex vesicaria*, mean $r = 0.95$; *Maireana sedifl ora*, mean $r = 0.97$). The reference unit method also demonstrated accuracy in estimating biomass of two shrub species in Brazil (*Mimosa acutistipula* and *Auxemma oncocalyx*; Kirmse and Norton 1985), and one shrub species in Utah (*Krascheninnikovia ceratoides*; Cabral and West 1986). Similarly, Carpenter and West (1987) reported that the reference unit method was a valid method of estimating biomass of low-statured shrubs (*Artemisia tridentata* ssp. *vaseyana* and *Atriplex gardneri*) and herbs (*Polygonum aviculare* and *Salsola kali*) in Wyoming.

Several researchers (Newsome et al. 1989, Harrington and John 1990, Noble et al. 2009) have used the validation of the reference unit method with shrubs as sufficient evidence of precision and accuracy of the method for herbaceous forbs and grasses. However, none of the publications we found provided specific details on the application of the reference unit method, developed primarily for shrubs, to herbaceous vegetation or evidence of the accuracy or precision of the estimates. Although validations of the reference unit method have been reported for shrubs, further validation of the method with herbaceous vegetation is needed before the method is completely adopted for herbaceous biomass estimation. The Buffalo Gap National Grassland (BGNG) is a large, native grassland with a complex mix of mostly herbaceous vegetation that is grazed by both large and small herbivores. Hence, BGNG is an ideal laboratory for evaluating the reference unit method of estimating herbaceous biomass at the landscape level for a variety of species and plant functional groups under varying ecological conditions. The information we report here was part of a larger research project investigating vegetation heterogeneity within and among prairie dog colonies on the BGNG. Our objectives were to (1) validate the reference unit method for estimating herbaceous grassland biomass on a broad spatial scale, (2) examine multi-species reference calibrations, (3) evaluate the use of season-long calibration equations, and (4) compare differences among observers using the reference unit method. Our results provide specific information on applying and validating the reference unit method of estimating biomass to a variety of individual herbaceous species and plant functional groups.

METHODS

Study Area

Our study was conducted on the 240,000 ha Buffalo Gap National Grassland located in Custer, Fall River, Jackson, and Pennington counties in southwestern South Dakota. The climate was semiarid and continental with the majority of precipitation occurring between April and September. Precipitation from August 2009 to July 2010 approximated the long-term average (41 cm) at Edgemont (western Fall River County) but was 62 and 25% higher than long-term average at Oral (69 cm, northeastern Fall River County) and Cottonwood (55 cm, Jackson County), respectively (High Plains Regional Climate Center 2011).

The BGNG was predominantly native grassland on gently rolling sedimentary plains with scattered buttes and un-vegetated and eroded badlands. Dominant graminoids included western wheatgrass (*Pascopyrum smithii* [Rydb.] A. Löve), green needlegrass (*Nassella viridula* [Trin.] Barkworth), needle-and-thread (*Hesperostipa comata* [Trin. & Rupr.] Barkworth), threadleaf sedge (*Carex filifolia* Nutt.), blue grama (*Bouteloua gracilis* [Willd ex Kunth] Lag. ex Griffiths), and buffalograss (*B. dactyloides* [Nutt.] J.T. Columbus; Kostel 2006). Common forbs included scarlet globemallow (*Sphaeralcea coccinea* [Nutt.] Rydb.), western wallflower (*Erysimum asperum* [Nutt.] DC.), American vetch (*Vicia Americana Muhl.* ex Willd.), scurfpeas (*Pediomelum* spp.), purple coneflower (*Echinacea angustifolia* DC.), prairie coneflower (*Ratibida columnifera* [Nutt.] Woot. & Standl.), dotted gayfeather (*Liatris punctata* Hook.), and Missouri goldenrod (*Solidago missouriensis* Nutt.).

Field Sampling

We sampled nine sites in 2010 from early June through the second week of August. Each sample site consisted of an active prairie dog colony and an adjacent off-colony area. We selected sites based on size of prairie dog colonies (≥12 ha), wet-weather accessibility, and consistency of soil texture characteristics among sites. Sample sites were restricted to the Clayey (four sites) and Loamy (five sites) ecological sites of Multiple Land Resource Area (MLRA) 60A – Pierre Shale Plains (U.S. Department of Agriculture 2015). These two ecological sites represented a major component of this 2.63 million ha MLRA. All sample sites were grazed at recommended stocking rates following NRCS guidelines developed specifically for each ecological site. Collectively, sampling prairie dog colonies within two ecological sites through the growing season allowed us to evaluate the reference unit method along a considerable gradient of plant community composition and abundance.

Our selection of individual species to have their own reference unit was based on the dominant species described in state and transition models published by the NRCS for the Clayey and Loamy ecological sites. Individual species included western wheatgrass (PASSMI), green needlegrass (NASVIR), purple threeawn (*Aristida purpurea* Nutt., ARIPUR), and needle-and-thread (HESCOM). We combined less common species into functional groups based on life span (annual or perennial), basic growth form (graminoid or forb), and origin (native or introduced). The use of functional groups facilitated estimates of minor species (Mannetje and Haydock 1963, Hutchings and Schmautz 1969) as they collectively could be significant in a study of ecological relationships (Walker 1970) but difficult to estimate individually. Multiple species functional groups included shortgrass mix – buffalograss and blue grama (SHORT), perennial grass native (PGN), perennial grass introduced (PGI), annual grass native (AGN), annual grass introduced (AGI), perennial forb native (PFN), perennial forb introduced (PFI), annual forb native (AFN), and annual forb introduced (AFI).

We sampled both off- and on-colony locations at each site, then stratified on-colony locations into interior and edge, which provided additional vegetation heterogeneity. Interior on-colony locations were near the center of the colony, which was typically the oldest portion of the colony and contained a high forb component (Bonham and Lerwick 1976). Edge locations were established close to the outer perimeter of the colony and typically had a greater graminoid component than the interior locations. Off-colony locations were 100 m to 200 m from the perceived edge of the colony within the same ecological site and had an absence of prairie dog activity. We randomly placed four, 100-m transects within each location (interior, edge, and off-colony) for a total of 12 transects at each of the nine sample sites. We placed circular sample plots (0.25 m^2) at 20-m intervals along each transect for a total of five sample plots per transect. We used transects as the experimental unit because individual plot level estimations generally do not have strong relationships between direct and indirect measurements, but relationships can improve when averaged over an area, such as a transect (Ganguli et al. 2000).

We developed reference units for each site by clipping samples outside of the sample plots to ground level that best represented the current growth and phenological characteristics of each species and functional group being sampled (Kirmse and Norton 1985). Reference units typically represented approximately 5 to 10% of the biomass of the species and plant functional groups in the plots. We stored reference units in clear zip top plastic bags to reduce wilting, limit moisture loss, and reduce the potential of accidental damage to foliage during sampling. The clear bags also allowed for easy visual comparisons of reference units to plants in plots. Reference units were kept out of the sun and refrigerated overnight, which allowed them to be used for several consecutive days before being replaced. To help ensure that reference units remained a consistent proportion of the actual biomass in the plot, we collected new reference units when changes in phenological stages, height, or foliage density were noticeable or if unit size required estimates higher than 10 units or less than 0.1 units. When functional groups were composed of multiple species, we used the most visually dominant species of that functional group in the plot as the reference unit. All used reference unit material was dried at 60º C for 72 hrs and weighed to the nearest 0.01 g.

Because the SHORT functional group was often too short to develop a reference unit that could be easily compared to a plot, we used a patch of short grasses that approximated the biomass of average height of shortgrass occupying 5% of the plot area. This shortgrass reference unit was used only in plots that were similar in average height to the reference unit. If the heights were not similar, another reference unit was harvested to match the new average height.

Three designated observers estimated biomass of individual species and functional groups within each plot using reference units. Observers rotated duties of clipping and estimating biomass so that each observer would estimate one in every three transects. Observer experience was variable with Observer 1 having previous experience using the reference unit method, Observer 2 having no prior experience using the method prior to the study, and Observer 3 having received brief training on the method in an undergraduate course. The biomass of each individual species and functional group was estimated within each plot as a ratio of the reference unit: estimated unit in increments of 0.1 of a unit. Because very small plants were difficult to estimate and harvest, species and functional groups were estimated only if their foliar cover was estimated to be above 1%. Total biomass was estimated by summing all estimates on a plot. After each plot was estimated, we clipped biomass by species and functional group to ground level, oven dried samples for 72 hrs at 60° C, and then weighed each sampled to the nearest 0.01g.

Statistical Analysis

Comparisons of actual (clipped) to estimated biomass of each species and functional group for each plot was averaged for each transect. Because our goal was to evaluate the reference unit approach under a gradient of ecological conditions, we pooled the 108 transects across ecological sites and prairie dog colonies (interior and edge). We evaluated the relationship between estimated biomass, using the reference unit, and clipped biomass with least squares regression for each observer within each species/functional group in a calibration regression (PROC REG with INFLUENCE option, SAS Institute Inc., Cary, North Carolina, USA). For each regression, we split the data by sampling time to create two datasets, one to construct the model and another to validate the model. We used the first 65% of transects sampled during

the field season (included a mix of the two ecological sites) as the construction dataset while the remaining transects formed the validation dataset (Picard and Berk 1990). We weighted the procedure because data did not meet the assumption of constant variance (homoscedasticity) required for ordinary least squares procedures (Ganguli et al. 2000, Zar 2010). Data points with relatively large studentized residuals (>2), Hat matrix leverage ($2p/n$), or DFFITs ($> \sqrt{p/n}$), where p = number of parameters and n = number of transects, were investigated as outliers and possibly removed (SAS Institute 2008); Tables 1 and 4 show the number of transects that were removed as outliers. Results from q-q plots indicated that data were normally distributed. Because it is unlikely that ocular estimations will change consistently as plot biomass increases (Ahmed et al. 1983), we ignored statistically significant curvilinear relationships ($\alpha = 0.05$) unless there was a substantial number of points conclusively demonstrating a curvilinear relationship (Andrew et al. 1979). Regressions with a substantial number of points demonstrating curvilinear relationships were transformed to establish the best linear fit.

We calibrated the validation data set using the previously created regressions. We compared calibrated reference weights to the actual clipped weights of the validation transect for each observer, species, and plant functional group using paired t-tests. We considered the models valid if the mean difference between calibrated reference weight and actual weight was not different from zero ($\alpha = 0.05$). Creating regressions was not possible for species and functional groups that were uncommon among transects (ARIPUR, HESCOM, PGI, and AGI). For these species and functional groups, we used non-calibrated estimation values that were directly compared to actual biomass values using a paired t-test in SAS to determine if estimations were accurate enough without calibrations to estimate minor functional groups or species.

Double sampling methods that are consistent over a variety of factors and locations are ideal because data can then be pooled to decrease sampling costs and increase precision of measurements (Laca et al. 1989). Testing for coincident lines between observers and key species or functional groups can determine if data can be pooled across certain species or observers. We conducted a 2 × 2 between groups analysis of covariance (ANCOVA) using a weighted general linear model (IBM SPSS Statistics 19, SPSS Inc., Chicago, Illinois, USA). The dependent variable was clipped biomass, with two independent variables of observer and key species or functional groups. We used estimated biomass as the covariate and considered the regression lines coincidental if there was no significant difference among groupings ($\alpha \leq 0.05$).

RESULTS

For each observer, all species and functional groups total biomass regressions had a significant relationship ($P \leq 0.05$) between estimated biomass (x) and actual biomass (y; Table 1). Regression model intercepts were not different from zero ($P > 0.05$) except PASSMI (Observer 2) and total biomass (all three observers). The mean coefficient of determination for all regressions was 0.86 (range = 0.61 to 0.97). Observers 1 and 3 had similar coefficient of determination for regressions, with Observer 1 having a range of 0.67 to 0.95 ($\overline{\times} = 0.88$) and Observer 3 having a range of 0.81 to 0.95 ($\overline{\times} = 0.88$). In contrast, Observer 2 had slightly lower coefficient of determination values compared with Observers 1 and 3, with a range of 0.61 to 0.97 ($\overline{\times} = 0.80$). PASSMI and SHORT shared the highest mean coefficient of determination ($\overline{\times} = 0.91$) while PGN had the lowest mean coefficient of determination ($\overline{\times} = 0.73$). All three observers appeared to have estimates of AGI that produced slopes near 1 with relatively high levels of precision (R^2 range 0.77 to 0.90). All regression models were validated (no difference between estimated and actual biomass, $P > 0.05$) except for the AGI group for Observer 2 ($P = 0.03$; Table 2). Actual biomass and non-calibrated biomass estimates for uncommon species/functional groups did not statistically differ from one another ($P > 0.05$; Table 3).

The ANCOVA to test for coincident lines showed no significant interaction between observer and functional groups ($F_{16, 446} = 0.95$, $P = 0.51$), indicating that all of the slopes for regressions were similar. After adjusting for the covariate of estimated biomass, there was no significant difference in regression intercepts between observers ($F_{2, 446} = 0.83$, $P = 0.44$) while there was a significant difference between functional groups ($F_{8, 446} = 2.93$, $P = 0.003$). This indicates that observers tended to estimate species and functional groups similarly, but significant differences were found among species and functional groups. Clipped transect biomass differed among species and functional groups with PASSMI as the only species that was different than the other species and functional groups with mean biomass 44% higher than the next closest functional group, AGI.

Because slopes and intercepts were similar among observers but different among functional groups, observer regressions could be compiled by functional groups. All 9 functional groups along with a 10th total biomass group had a significant relationship ($P < 0.05$) between estimated and actual biomass (Fig. 1, 2). Estimated biomass explained an average of 82% of the variation in actual biomass (range = 64 to 90%) suggesting that the regressions have high predictive value. Collectively, observers tended to estimate biomass close to unity (y = 1x) as the average absolute difference between regression slopes and a slope of one was only 0.09 (Fig. 1). However, the scatterplot shows slight tendencies for observers to variously under- and over-estimate several groups. When estimates of individual groups were summed at the plot level to estimate total biomass, we found that the cumulative effects of over- and under-estimating were more pronounced compared to the group level (Fig. 2). The net effect was that small quantities of total biomass tended to be

Table 1. Species and functional group regression equations for three different observers [x = estimated biomass, y = actual biomass (g/0.25m^2)]. Functional groups include AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = shortgrass mix, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN = perennial forb introduced, PGN = perennial grass native, TOTAL = total biomass (*n* = number of transects).

Species/ Functional Group	Outliers	Observer	*n*	Equation	*P*		
					INT[1]	*X*[3]	R^2
AFI	1	1	25	y = −0.01+1.50x	0.88	< 0.001	0.88
	0	2	25	y = −0.24+0.87x	0.27	< 0.001	0.61
	0	3	26	y = −0.15+1.12x	0.39	< 0.001	0.87
AFN	0	1	26	y = −0.05+1.43x	0.62	< 0.001	0.93
	1	2	25	y = 0.15+0.91x	0.37	< 0.001	0.78
	0	3	26	y = −0.02+1.08x	0.91	< 0.001	0.95
AGI	0	1	25	y = 0.16+0.95x	0.60	< 0.001	0.90
	0	2	25	y = 0.56+1.03x	0.23	< 0.001	0.77
	0	3	26	y = 0.67+0.91x	0.23	< 0.001	0.88
AGN	1	1	25	y = 0.11+1.55x	0.77	= 0.003	0.85
	0	2	25	y = −0.08+1.03x	0.61	< 0.001	0.97
	2	3	25	y = 0.03+0.73x	0.77	< 0.001	0.83
SHORT	1	1	25	y = 0.25+0.79x	0.33	< 0.001	0.95
	1	2	25	y = 0.39+0.94x	0.16	< 0.001	0.91
	0	3	26	y = 0.36+0.77x	0.37	< 0.001	0.87
PASSMI	1	1	25	y = 0.44+1.16x	0.14	< 0.001	0.95
	1	2*	25	$y = -3.74+5.06x^{1/2}$	< 0.001	< 0.001	0.90
	0	3	26	y = 0.80+0.95x	0.11	< 0.001	0.88
PFI	0	1	25	y = −0.10+2.06x	0.24	< 0.001	0.91
	2	2	24	y = −0.04+0.81x	0.24	= 0.005	0.64
PFN	0	1	25	y = 0.08+1.35x	0.70	< 0.001	0.87
	1	2	24	y = −0.01+0.97x	0.94	< 0.001	0.91
	1	3	26	y = 0.14+0.93x	0.31	< 0.001	0.92
PGN	1	1	24	y = 0.21+1.08x	0.29	< 0.001	0.67
	1	2	25	y = 0.35+0.72x	0.16	< 0.001	0.71
	1	3	26	y = 0.18+1.03x	0.21	< 0.001	0.81
TOTAL	1	1	25	y = 7.87+0.89x	0.01	< 0.001	0.87
	1	2	25	y = 4.11+0.80x	0.02	<0.001	0.92
	1	3	26	y = 4.84+0.84x	0.03	<0.001	0.92

* = curvilinear relationship transformed using $\sqrt{x}$. [1]Probability the intercept (INT) is different from zero. [3]Probability the slope (*X*) is different from zero.

Table 2. Validation paired t-test for functional groups by observer between calibrated biomass estimations and actual biomass. Species and functional groups include AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = shortgrass mix, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN = perennial forb introduced, PGN = perennial grass native. Mean difference between calibrated and actual biomass in g/0.25m^2 (*n* = number of transects).

Functional group	Observer	*n*	Mean Difference	SD	*P*[1]
AFI	1	13	0.40	1.83	0.44
	2	15	0.09	0.36	0.43
	3	14	0.05	0.11	0.12
AFN	1	14	−0.11	1.09	0.72
	2	13	0.43	0.99	0.14
	3	15	1.38	3.32	0.13
AGI	1	13	0.36	1.28	0.34
	2	13	4.89	6.96	0.03
	3	15	1.59	4.23	0.17
AGN	1	13	0.26	0.68	0.20
	2	14	0.23	0.71	0.25
	3	14	−0.02	0.35	0.83
SHORT	1	14	−0.13	1.77	0.79
	2	13	1.57	3.45	0.13
	3	14	−1.34	2.56	0.08
PASSMI	1	14	−0.63	2.41	0.35
	2	13	−0.20	6.70	0.90
	3	14	1.44	3.24	0.12
PFI	1	14	−0.02	0.04	0.07
	2	13	0.05	0.09	0.08
PFN	1	14	0.82	1.65	0.09
	2	13	−0.26	0.73	0.23
	3	14	0.21	1.28	0.55
PGN	1	14	-0.02	0.34	0.83
	2	13	0.21	0.48	0.14
	3	14	0.28	0.61	0.12
TOTAL	1	14	0.80	7.09	0.68
	2	13	−11.08	18.89	0.06
	3	14	−1.60	6.36	0.36

[1]Probability the mean difference between calibrated biomass and actual biomass was different from zero.

Table 3. Non-regression paired t-test for uncommon species and functional groups by observer. Species and functional groups include ARIPUR = three-awn, HESCOM = needle-and-thread, NASVIR = green needlegrass, PFI = perennial forb introduced, PGI = perennial grass introduced. Results in g/0.25m^2 (*n* = number of transects).

Group	Observer	*n*	Mean[1]	SD	*P*[2]
ARIPUR	1	40	0.03	0.50	0.73
	2	39	0.03	0.20	0.32
	3	41	−0.29	1.50	0.21
HESCOM	1	40	0.04	0.20	0.32
	2	39	0.01	0.10	0.38
	3	41	0.01	0.10	0.32
NASVIR	1	40	0.01	0.10	0.22
	2	39	−0.34	2.10	0.31
	3	41	−0.260	1.60	0.32
PFI	3	41	0.02	0.20	0.46
PGI	3	41	0.02	0.10	0.32

[1]Mean difference between non-calibrated estimated biomass and actual biomass; [2]Probability that the mean difference between estimated and actual biomass is different from zero.

Table 4. Results of validation paired t-test for compiled functional groups between calibrated biomass estimations and actual biomass in the validation set for species (PASSMI) and functional groups compiled from three different observers. Functional groups include A = annual, P = perennial, N = native, I = introduced, G = graminoid, F = forb, and SHORT = shortgrass mix. Results in g/0.25m^2 (*n* = number of transects).

Group	*n*	Mean[1]	SD	*P*[2]
AFI	42	−0.05	0.87	0.74
AFN	42	1.33	4.73	0.08
AGI	42	2.01	4.65	0.01
AGN	41	0.29	0.70	0.01
SHORT	42	−0.23	2.23	0.50
PASSMI	41	2.97	10.87	0.09
PFI	42	0.01	0.27	0.77
PFN	42	0.22	1.28	0.28
PGN	42	0.17	0.66	0.10
TOTAL	41	4.18	15.34	0.09

[1] Mean average difference between calibrated and actual biomass; [2]Probability mean difference between calibrated and actual biomass is different from zero.

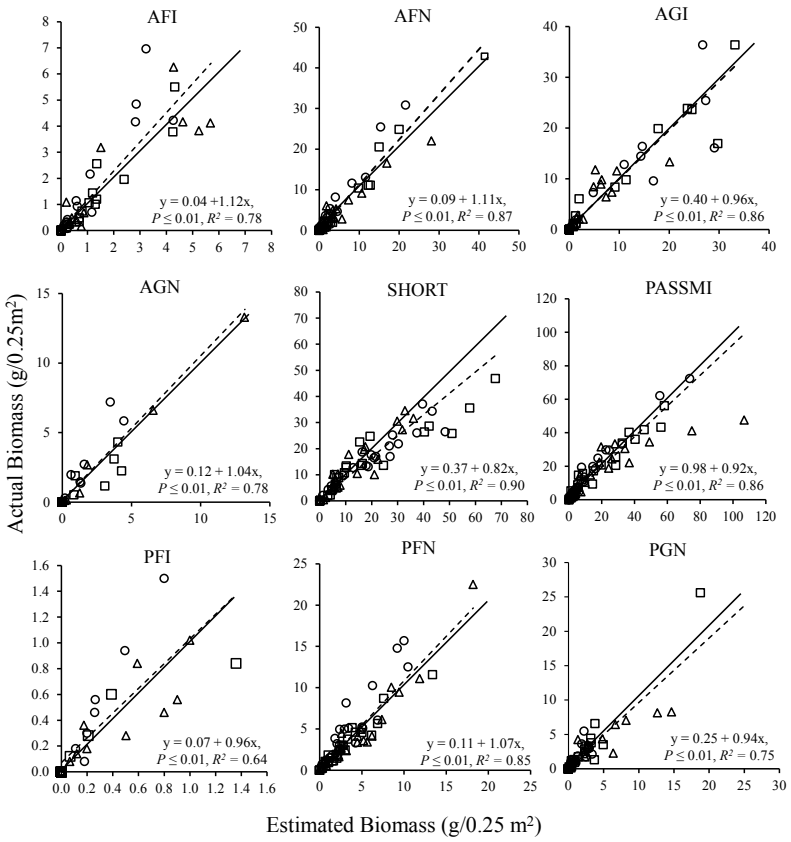

Figure 1. Regression lines by functional groups compiled across observer (dashed line). Solid line is a reference to y = 1x. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Note different scales of y- and x-axis. P = perennial, A = Annual, F= Forb, G=Grass, N = Native, I = Introduced, Short = buffalograss and blue grama, PASSMI = western wheatgrass. Circles, squares, and triangles represent Observer 1, 2, and 3, respectively.

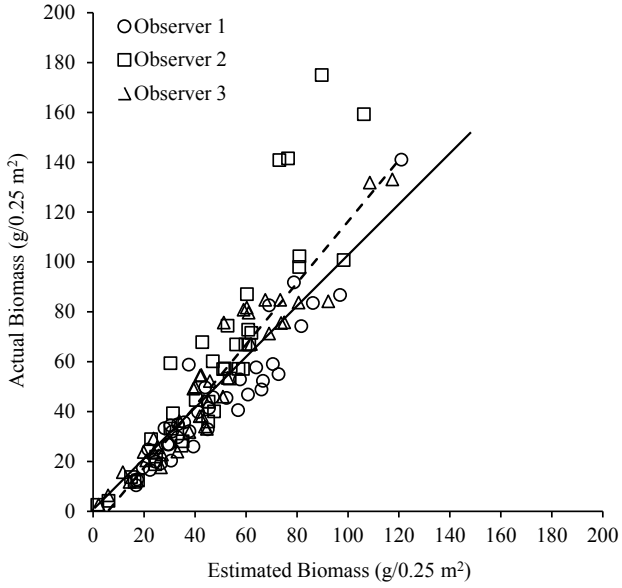

Figure 2. Regression line for total biomass compiled across observer (dashed line). Solid line is a reference to y = 1x. Data were collected in 2010 on the Buffalo Gap National Grasslands, South Dakota, USA. Circles, squares, and triangles represent Observer 1, 2, and 3, respectively.

overestimated while large quantities tended to be underestimated.

After we combined estimates among the three observers for each functional group, mean differences between calibrated and actual biomass estimations ranged, in g/0.25 m^2, from 0 (AFI and PFI, $P = 0.74$ and 0.77, respectively) to 4.2 (TOTAL, $P = 0.09$; Table 4). Despite the high precision and accuracy of calibrations for AGI and AGN ($r^2 = 0.86$ and 0.78, and slope = 0.96 and 1.04, respectively, see Fig. 1), we found significant differences ($P \leq 0.05$) between calibrated and actual biomass. All other paired comparisons were similar ($P > 0.05$) and calibrations appeared to closely approximate the actual mean with an average mean difference of 0.76 g/0.25m^2 (SE = 0.37).

DISCUSSION

Ideally, biomass estimation methods would express relationships between actual and estimated biomass with high precision (Tadmor et al. 1975) and accuracy (Hutchings and Schmautz 1969) with intercepts that pass through or close to the origin (Carpenter and West 1987). Additionally, estimates should be reliable in a variety of environmental conditions and species compositions (Reese et al. 1980) and remain consistent among observers through the sampling period (Friedel et al. 1988). Overall, the reference unit approach described here is broadly applicable in short- and mixed-grass prairies. In terms of accuracy and precision, our results for a variety of herbaceous species and functional groups are comparable to those reported for several species of shrubs and two forb species (Andrew et al. 1979, 1981, Kirmse and Norton 1985, Carbral and West 1986, Carpenter and West 1987). All three observers in our study demonstrated the ability to consistently make accurate estimates of biomass for individual species and functional groups using the reference unit method. With a few exceptions, comparisons of most slopes among observers revealed that calibrations were generally close to a 1:1 relationship of estimated to actual biomass with most intercepts equal to zero ($P > 0.05$). We found one nonlinear relationship between actual and estimated biomass (PASSMI, Observer 2). We suspect this may have been the result of inadequate visual separation of living PASSMI stems from standing dead while estimating biomass in tall dense stands of vegetation before physically separating live and dead material after clipping. Similar issues were reported by Shoop and McIlvain (1963) who found that it was difficult to estimate overly-mature, high production areas using the micro-unit method. Hutchings and Schmautz (1969) also had difficulties in visually sorting live biomass from standing dead while using the relative-weight method. Additionally, Reese et al. (1980) found that the amount of standing dead plant material present in grasslands impacted the ability to use the weight-estimate method. Because sampling duties rotated from transect to transect in our study, observers often did not harvest the plots they referenced.

Observer experience can impact biomass estimation, especially when estimating absolute biomass (Gillen and Smith 1986). Several authors suggested intensive training periods before implementing an estimation method (Pechanec and Pickford 1937, Mannetje and Haydock 1963, Shoop and McIlvain 1963). However, we found that relatively inexperienced observers could be as consistent in their estimations as their experienced counterparts. Researchers using methods dependent upon relative weight comparisons reported no significant difference between experienced and inexperienced observers, similar to our findings (Hutchings and Schmautz 1969, Hutchinson et al. 1972, Haydock and Shaw 1975, Reese et al. 1980, Kirmse and Norton 1985). Collectively, the evidence suggests that individuals may have a greater natural aptitude for relative estimates than direct estimates (but see Tadmor et al. 1975). Regardless, it seems that individuals can become capable estimators of relative weight in a brief amount of time compared to direct estimation, likely because individuals make many more relative than direct comparisons during their everyday life.

Despite high precision of regressions with slopes approaching unity, observers tend to underestimate biomass (slopes > 1) for some species and functional groups while overestimating others (slopes < 1). Haydock and Shaw (1975) reported similar results and argued that over- and under-estimation is expected if a method is unbiased. With one exception (SHORT), we found that overestimations by one observer were somewhat balanced by underestimations of another. Consequently, when regressions among observers were pooled by functional groups, slopes approached 1 with intercepts approaching zero (Fig. 1). Such a pattern implies the need for a balanced sample design when observer data are pooled. If individual observers sample a greater quantity of experimental units than others, or if observers sample stratified areas alone, there could be a potential bias in estimates. Shoop and McIlvain (1963) similarly noted that observers varied in their ability to estimate biomass and recommended that each observer sample a proportionate number of plots in the pasture to avoid bias of results.

All observers tended to overestimate biomass of SHORT (buffalograss and blue grama). Blue grama is a bunch grass while buffalograss is stoloniferious. Stolons are aboveground stems that have a relatively greater mass density than the leaf blades; consequently, reference units that contained more stolons than were actually present in a plot could result in an overestimation of biomass. Overestimations of biomass of SHORT may be further compounded by occasional changes in the proportions of buffalograss to blue grama and *Carex* spp. that were intermixed within the plot. *Carex* spp. often were difficult to see while estimating, but were sorted out into perennial native grasses (PGN) during clipping. This issue was reported by Shoop and McIlvain (1963), who noted that observers using the micro-unit method tended to overlook small plants. In contrast, observers underestimated intro-

duced and native annual forbs (AFI and AFN) probably because of the difficulties of finding all of the annual forb species in high diversity plots. Hutchings and Schmautz (1969) encountered a similar issue using the relative-weight method.

Small plants may be especially susceptible to over- and under-estimates of biomass. Trends in overlooking small plants and having difficulties in estimating high levels of biomass become apparent when individual estimates are added together to estimate total biomass (Fig. 2). Pechanec and Pickford (1937) and Tadmor et al. (1975) found that weight-estimate methods tended to have less dispersion around the mean than actual values, which meant low quantities of biomass were overestimated and high quantities were underestimated. Conversely, our study suggests that estimated data with the reference unit method may have greater dispersion than actual biomass (i.e. small quantities were overestimated and large quantities were underestimated). Such trends often are not noticed on individual species and functional groups; however, the cumulative impacts of these small errors are noticeable when individual estimates are summed to estimate total biomass.

Although previous research has not thoroughly indicated the potential for using multiple species calibration equations for reference units, the reference unit method is capable of accurate and precise collective estimates of multiple species. While using the dry-weight-rank method, Walker (1970) reported difficulties in estimating weight contributions because individually insignificant species had combined weights that may be significant. Mannetje and Haydock (1963) and Hutchings and Schmautz (1969) suggested collectively measuring uncommon or minor species to improve estimates. In our study, functional groups contained assemblages of similar species that were collectively estimated. Individually creating calibrations for all these species with enough data to stabilize regression coefficients would have been time consuming, especially given the patchy or uncommon nature of many of these species. However, almost all of the multiple species groups used in our study were validated (22 of 23 calibrations).

Composition and structure of functional groups and reference units may influence the ability to accurately estimate biomass. Large differences in plant structure and weight densities can make estimation more difficult as seen with PGN, which had the lowest mean coefficient of determination. This particular functional group contained both short *Carex* (grass-like) species and taller grass species. Most of the remaining functional groups had species with less disparity between growth characteristics that allowed for easier estimation of the species collectively. Reese et al. (1980) had similar difficulties with the relative-weight method when dissimilar species exhibited different weight densities. Kirmse and Norton (1985) reported that regressions tended to have higher coefficients of determination when the appearance of reference units more closely resembled the shrub foliage they were estimating. Collectively, this evidence suggests that more precise estimates will occur for single species or simple functional groups because reference units used for these groups will more closely resemble the biomass being estimated.

MANAGEMENT IMPLICATIONS

The reference unit method described here is capable of precise and accurate estimates of biomass for herbaceous forbs and grasses with multiple observers. Double sampling can be used to create a calibration equation for specific species and functional groups. Multi-species functional groups have limited impacts on the precision and accuracy of the method as long as species within the group have similar growth habits and weight densities. However, reference units created for these groups should only contain the species of the functional group that is most dominant within the plot. A single calibration can be made for each functional group using data from estimations made with numerous reference units. Conscientious observers can maintain consistency in estimations throughout a season and within a variety of ecological conditions. Experience using the method has little impact on estimations and training can be completed in one hour. Our results demonstrate that the reference unit method maintains consistency and is unaffected by vegetation heterogeneity and observer experience.

ACKNOWLEDGMENTS

We thank E. Barnett and H. Bergeleen for their hard work in collecting the field data. J. Ott, M. Greer, and two anonymous reviewers provided valuable comments that greatly improved early drafts of the manuscript. We thank C. Ren for assistance on the statistical analysis. We greatly appreciate the assistance and hospitality of the staff of the US Forest Service, Buffalo Gap National Grasslands. This research was supported in part by funds provided by the US Forest Service Rocky Mountain Research Station and the Nebraska National Forests and Grasslands.

LITERATURE CITED

Ahmed, J., C. D. Bonham, and W. A. Laycock. 1983. Comparison of techniques used for adjusting biomass estimates by double sampling. Journal of Range Management 36:217−221.

Andrew, M. H., I. R. Noble, and R. T. Lange. 1979. A non-destructive method for estimating the weight of forage on shrubs. Australian Rangeland Journal 1:225−231.

Andrew, M. H., I. R. Noble, R. T. Lange, and A. W. Johnson. 1981. The measurement of shrub forage weight, three methods compared. Australian Rangeland Journal 3:74−82.

Bonham, C. D., and A. Lerwick. 1976. Vegetation changes induced by prairie dogs on shortgrass range. Journal of Range Management 29:221−225.

Cabral, D. R., and N. E. West. 1986. Reference unit-based estimates of winterfat browse weights. Journal of Range Management 39:187−189.

Carpenter, A. T., and N. E. West. 1987. Validating the reference unit method of aboveground phytomass estimation on shrubs and herbs. Vegetatio 72:75−79.

Catchpole, W. R., and C. J. Wheeler. 1992. Estimating plant biomass; a review of techniques. Australian Journal of Ecology 17:121−131.

Friedel, M. H., V. H. Chewings, and G. N. Bastin. 1988. The use of comparative yield and dry-weight-rank techniques for monitoring arid rangeland. Journal of Range Management 41:430−435.

Ganguli, A. C., L. T. Vermeire, R. B. Mitchell, and M. C. Wallace. 2000. Comparison of four nondestructive techniques for estimating standing crop in shortgrass plains. Agronomy Journal 92:1211−1215.

Gillen, R. L., and E. L. Smith. 1986. Evaluation of the dry-weight-rank method for determining species composition in tallgrass prairie. Journal of Range Management 39:283−285.

Harrington, G. N., and G. G. John. 1990. Herbaceous biomass in Eucalyptus savanna woodland after removing trees and/or shrubs. Journal of Applied Ecology 27:775−787.

Haydock, K. P., and N. H. Shaw. 1975. The comparative yield method for estimating dry matter yield of pasture. Australian Journal of Experimental Agriculture and Animal Husbandry 15:663−670.

High Plains Regional Climate Center. 2011. Historic climate data summaries. <http://www. hprcc.unl.edu/index. php>. Accessed 6 June 2011.

Hutchings, S. S., and J. E. Schmautz. 1969. A field test of the relative-weight-estimate method for determining herbage production. Journal of Range Management 22:408−411.

Hutchinson, K. J., R. W. McLean, and B. A. Hamilton. 1972. The visual estimation of pasture availability using standard pasture cores. Journal of the British Grassland Society 27:29−34.

Kirmse, R. D., and B. E. Norton. 1985. Comparison of the reference unit method and dimensional analysis methods for two large shrubby species in the Castinga woodlands. Journal of Range Management 38:425−428.

Kostel, G. M. 2006. A floristic inventory of the Buffalo Gap National Grassland (South Dakota) and Oglala National Grassland (Nebraska). Thesis. University of Wyoming, Laramie, USA.

Laca, E. A., M. W. Demment, J. Winckel, and J. G. Kie. 1989. Comparison of weight estimate and rising-plate meter methods to measure herbage mass of a mountain meadow. Journal of Range Management 42:71−75.

Mannetje, L. T., and K. P. Haydock. 1963. The dry-weight-rank method for the botanical analysis of pasture. Journal of the British Grassland Society 18:268−275.

Newsome, A. E., I. Parer, and P. C. Catling. 1989. Prolonged prey suppression by carnivores: predation removal experiments. Oecologia 78:458−467.

Noble, J. C., W. J. Müller, W. G. Whitford, and G. H. Pfitzner. 2009. The significance of termites as decomposers in contrasting grassland communities of semi-arid eastern Australia. Journal of Arid Environments 73:113−119.

Pechanec, J. F., and G. D. Pickford. 1937. A weight estimate method for the determination of range or pasture production. Journal of the American Society of Agronomy 29:894−904.

Picard, R. R., and K. N. Berk. 1990. Data splitting. The American Statistician 2:140−147.

Reese, G. A., R. L. Bayn, and N. E. West. 1980. Evaluation of double-sampling estimators of subalpine herbage production. Journal of Range Management 33:300−306.

SAS Institute. 2008. The SAS system for Windows. Release 9.2. SAS Institute, Inc., Cary, North Carolina, USA.

Shoop, M. C., and E. H. McIlvain. 1963. The micro-unit forage inventory method. Journal of Range Management 16:172−179.

Tadmor, N. H., A. Brieghet, I. Noy-Meir, R. W. Benjamin, and E. Eyal. 1975. An evaluation of the calibrated weight-estimate method for measuring production in annual vegetation. Journal of Range Management 28:65−69.

U. S. Department of Agriculture. 2015. Field Office Technical Guide. <http:efotg.sc.egov.usda.gov/treemenuFS.aspx> Accessed 23 April 2015.

Walker, B. H. 1970. An evaluation of eight methods of botanical analysis on grasslands in Rhodesia. Journal of Applied Ecology 7:403−416.

Wilm, H. G., D. F. Costello, and G. E. Klipple. 1944. Estimating forage yield by the double-sampling method. Journal of the American Society of Agronomy 36:194−203.

Zar, J. H. 2010. Biostatistical analysis. Fifth Edition. Prentice Hall, Upper Saddle River, New Jersey, USA.

Submitted 27 March 2015. Accepted 20 September 2015. Associate Editor was G. Larson.