EVALUATION OF THE REFERENCE UNIT METHOD FOR HERBACEOUS BIOMASS ESTIMATION IN NATIVE GRASSLANDS OF SOUTHWESTERN SOUTH DAKOTA

By

Eric D. Boyda

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Biological Science

South Dakota State University

2013

ACKNOWLEDGEMENTS

This research has been supported through a joint venture agreement between South Dakota State University and the U.S. Forest Service – Rocky Mountain Research Station. A thank you is owed to my advisors, Dr. Jack Butler and Dr. Lan Xu, for your support and patience. A special thanks to the Rocky Mountain Research Station and the Fall River and Wall Ranger Districts of the Buffalo Gap National Grasslands for assistance during the project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

EVALUATION OF THE REFERENCE UNIT METHOD FOR HERBACEOUS BIOMASS ESTIMATION IN NATIVE GRASSLANDS IN SOUTHWESTERN SOUTH DAKOTA

Eric D. Boyda

2013

The high costs associated with physically harvesting plant biomass may prevent sufficient data collection, which is necessary to account for the natural variability of vegetation at a landscape scale. A biomass estimation technique was previously developed using representative samples or "reference units", which eliminated the need to harvest biomass from all sample plots. Use of reference units increases the number of plots that can be sampled, allowing monitoring to capture the variability of rangeland. This technique was developed for use on shrub foliage with limited validation of the method in herbaceous grassland vegetation. The objectives of this study were to: 1) validate reference unit method for herbaceous grassland biomass estimation, 2) examine multi-species reference calibrations, 3) validate the use of season-long calibration equations, and 4) compare differences among observers using the reference unit method. The study was conducted in 2010 on prairie dog colonies of the Buffalo Gap National Grasslands in southwestern South Dakota. These prairie dog colonies and the surrounding areas provided a range of vegetation height and plant communities to use for testing the method. Research revealed that reference units provided accurate and precise estimates of the mean for herbaceous plants on grasslands, including multi-species functional groups.

Twenty-five of the 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. Additionally, a technique to create shortgrass reference units was developed. The reference unit method is recommended for grassland evaluation and monitoring projects.

INTRODUCTION

Accurate estimates of aboveground production of a site that can be repeated with precision are an important metric in assessing wildlife habitat, fire fuel loads, forage availability, and ecological relationships and processes. Harvesting biomass through clipping can give an accurate and precise estimate of production at the plot level, however, the approach has high costs regarding time spent clipping, sorting, drying, and weighing biomass (Catchpole and Wheeler 1992; Haydock and Shaw 1975; Reese et al. 1980; Wilm et al. 1944). For example, Reese et al. (1980) estimated that clipping takes anywhere from 16-70 times longer than visual or metering estimation methods. The number of samples taken at higher spatial scales in heterogeneous landscapes is often constrained by budget or time. Consequently, estimation methods that have less precision at the plot level but allow for larger sample sizes could provide more accurate results than those provided by precisely clipped plots with a small sample size (Reese et al. 1980; Wilm et al. 1944). However, the statistical parameters obtained through estimation must be unbiased (Green 1949; Haydock and Shaw 1975), reliable in the range of environments and vegetation encountered in an area (Reese et al. 1980), and consistent among various observers (Friedel et al. 1988).

A continuous need for economical and accurate measurements of biomass that can be repeated with precision has led to the progressive development of the double sampling method. In using this method, a relationship is calculated between a variable that costs less to measure (indirect sample) and biomass that is accurately and precisely measured through destructive means (direct sample) (Catchpole and Wheeler 1992). Many samples can then be taken through indirect measurement and converted to an

estimated biomass using the calibration regression, thus potentially reducing the sampling burden (Catchpole and Wheeler 1992). Various techniques have been developed over time to take indirect measurements of biomass that range from modifying previous methods, developing different measurements, or creating new tools for measurements. It is difficult to explicitly declare a method superior to another, because each was developed for a specific need; consequently, each method has strengths and weaknesses in certain situations

Although it is highly unlikely an estimation technique will meet every desirable trait, they should include a combination of accuracy, speed, precision, limited interference from environmental conditions and topography, and if instruments are required, they should be inexpensive and easy to operate (Tucker 1980). Because the study of ecological relationships and processes could involve the estimation of biomass of individual species or functional groups, techniques that are ineffective or unable to estimate single species biomass may not be appropriate. Additionally, many techniques that sample total biomass require tools (such as drop discs, robel poles, capacitance probes, and leaf area index probes) that are susceptible to differing calibrations from season to season or within a single season (Tucker 1980) and from site to site. Estimation techniques that require tools may be best suited for grasslands with limited variability such as tame pastures. With the addition of the dry-weight rank method (a method that estimates the percent contribution of a species to total biomass), many of the estimation methods that can only estimate total biomass could serve this purpose. However, the accuracy of the dry-weight rank method depends on observer capabilities and requires initial training and subsequent training with changes in species compositions (Mannetje

and Haydock 1963). Monitoring and research activities are often dependent on seasonal or temporary employees with high turnover rates; therefore, training periods are usually necessary every year. Clearly, training observers to use a single effective method can save time and energy compared to the time and energy needed to train individuals in multiple methods that lead to similar results.

Visual estimation methods could allow for an observer to overcome the complications and bias that height, density, mass distribution, standing dead, terrain, and over-mature vegetation have on measurement tools (Ogura et al. 2005). In addition, the methods can estimate single species or functional groups, which makes them ideal estimation methods in many situations. The weight estimate method, a form of direct visual estimation, has been demonstrated to accurately estimate single species biomass (Hughes 1959; Pechanec and Pickford 1937; Shoop and McIlvain 1963); however, high variability of fresh weight moisture can affect results (Tadmor et al. 1975). The weight estimate method involves clipping and weighing samples to develop a capacity for estimating plant biomass. It can take several days, to a week, to develop consistency and accuracy in estimations with reoccurring retraining as seasonal changes in fresh weight moisture are observed (Pechanec and Pickford 1937). Often, estimates are impacted by mental fatigue, attitude, and the retention of mental images of units for too long (Hutchings and Schmautz 1969). Moreover, the weight estimate method may not be ideal for all situations because of the long training periods, variability in moisture content of fresh weight estimates, and bias caused by observers.

Rather than building a personal capacity for direct biomass estimation through clipping samples and weighing them in the field, it may be more efficient to carry the

clippings around and make comparisons relative to them. Under the assumption that it is easier to estimate relative weight than it is to estimate absolute weight (Haydock and Shaw 1975; Hutchings and Schmautz 1969; Reese et al. 1980), plots can be directly compared to a clipped sample or reference. This reference or "reference unit" method has demonstrated accuracy in biomass estimation of shrubs (Andrew et al. 1979; Andrew et al. 1981; Cabral and West 1986; Kirmse and Norton 1985) and a few shrub-like herbaceous forbs (Carpenter and West 1987); furthermore, training can be completed in as little as a few hours. Although validations of the reference unit method have only been reported for single species groups, Kirmse and Norton (1985) felt that regressions could collectively be developed for groups of species with similar foliage appearances.

Several researchers (Harrington and John 1990; Newsome et al. 1989; Noble et al. 2009; Stoltenberg 2004; Woodland 2004) have used the validation of the reference unit method with shrubs as sufficient evidence of the precision and accuracy of the method with herbaceous forbs and grasses. Although this is not necessarily inappropriate, further validation of the method with herbaceous vegetation, especially grasses should be undertaken before the method is completely adopted for herbaceous biomass estimation. This is particularly important for rhizomatous grasses because accurately estimating biomass for this growth form using the weight estimate method is difficult (Pechanec and Pickford 1937).

The reference unit method depends on calibration models to correct for personal bias in estimations. Most applications of reference units create a calibration model from five to ten plots external to the study using simple linear regression models. However, statistically significant regressions with high coefficients of determination is not enough

to determine capabilities of a method; models must encompass the many conditions or changing variables observed in the field to achieve a degree of stability in coefficients (Reppert et al. 1962). Purposely selecting a small scale validation set as recommended by Haydock and Shaw (1975) and Tadmor et al. (1975) can also bias results when the validation set does not truly reflect the variability of the larger estimated set (Reese et al. 1980). Additionally, most five to ten plot calibration models were created daily and require plot selection that encompasses the range of biomass values expected to be encountered during that period of sampling. Research or monitoring that requires estimation of multiple species or functional groups would notice a decrease in efficiency of the method since a large portion of a sampling day would be dedicated to creating calibrations only to have them be recreated the next day.

Consequently, the objectives of this study are to validate the reference unit method for use with herbaceous forbs and grasses, examine the potential for multiple species reference calibrations, confirm the use of season long calibration models with referenced data, and test the differences among observers using the reference unit method. I hypothesize that reference units will be validated for use with herbaceous forbs and grasses; be capable of multi-species calibrations; remain consistent throughout a sampling season, allowing for single season calibrations; and that they are used similarly among observers.

LITERATURE REVIEW

Biomass clippings can give an accurate and precise estimate of production but have high costs regarding time spent harvesting, sorting, drying, and weighing. Too often, research and monitoring efforts are constrained by time, which can limit their ability to examine an adequate number of experimental units needed to overcome the variability of grasslands (Catchpole and Wheeler 1992; Haydock and Shaw 1975; Reese et al. 1980; Wilm et al. 1944). Overtime, various estimation methods have been developed to increase the efficiency of sampling in various habitats for distinct purposes that include: determining grazing capacity and evaluating ecological relationships and processes.

Weight Estimates for Determining Grazing Capacity

Numerous methods were created to address estimates of total production of tame pasture and native rangeland. These methods are, in some cases, still used to determine grazing capacity but they were not established to evaluate complex ecological relationships and processes.

Direct Estimates of Biomass. The weight-estimate method (Pechanec and Pickford 1937) is one of the first methods described for biomass estimation and it created the foundation of the double sampling method. With the weight-estimate method, an observer develops the ability to accurately predict fresh weights (plant biomass that has not been dried) based on four weight units (10, 20, 50, and 100 g) by repeated clipping and weighing of vegetation from field plots. After an initial training period of several days to a week, the observer would then visually estimate the biomass of species in plots to the nearest 10 g using any combination of multiples of the four units (Pechanec and

Pickford 1937). These authors suggested that observers clip 10-20% of all the estimated plots to develop a linear regression model to correct for personal bias in estimations. Seasonal, local, and annual variations in moisture content were periodically checked by clipping and comparing fresh weights to dry weights to determine percent moisture. Moisture content of the estimations could then be subtracted, leaving an estimated dry weight for the plot. Mean estimates of weights were found to closely approach mean actual weights and the percent relative differences of estimated to actual weights were consistently below 10 percent, especially for forbs. However, differences were occasionally higher for species with small average plot biomass and rhizomatous grasses that lacked the easily definable units of forbs or bunch grasses. To remedy the larger relative weight differences in species with small average plot biomass, the authors recommended estimating to the nearest 1 or 5 grams where possible because 10 gram units were not precise enough for these specific species. No suggestions were given to help with the estimation of rhizomatous species.

The weight-estimate method requires little laboratory time outside of sampling because only a few samples need to be dried and weighed to calculate the percent dry biomass for various species and sample harvest times (Pechanec and Pickford 1937). However, frequent verification the ability of an observer to estimate the fresh weights of plants is required to maintain accuracy during sampling.

In their assessment of the weight estimate method, Wilm et al. (1944) stressed the importance of using the double sampling method as a quantitative control measure. Double sampling involves measuring a large sample indirectly (estimation), and a small subset of that sample directly (clipping). A linear regression is then used to convert or

calibrate the larger sample to that of the smaller sub set, thus creating a data set capable of analysis. Without these calibrations, estimators cannot provide data on productivity directly.

In a study conducted by Wilm et al. (1944), the weight-estimate method used in double sampling was 14% more efficient compared to direct harvest alone because fewer plots needed to be clipped. Furthermore, in a cost to benefit analysis, Wilm et al. (1944) found that 6 plots, rather than 4 plots, could have been indirectly sampled for each plot double sampled in their study. Had there not been this inefficiency in the study design, fewer plots would have been directly sampled (harvested) and the net gain in information would have drastically increased.

Double sampling using the weight-estimate method was highly adaptable and by 1959 was recommended for use in determining grazing capacity, calculating range condition and trends, taking inventory of flammable fuels, estimating the biomass of understory vegetation, and evaluating vegetation responses to management (Hughes 1959). Although the original instructions for the weight-estimate method were derived for estimating grasses and forbs, Hughes (1959) suggested the weight-estimate method for estimating current seasonal growth of woody-stemmed plants. Although no data were provided, the author, through personal experience, reported that estimates of the mean production can be obtained within 10% of the actual mean with an intensive training period.

Over time, modifications and suggestions for using the weight-estimate method have appeared in the literature. Shoop and McIlvain (1963) suggested using the weight-estimate method with only one easily estimated visual unit or “micro-unit”. Their fresh

weight micro-unit was around 10 grams, and plots were estimated in terms of the number and fractions of the unit. Because the micro unit was only 10 grams, observers rapidly refined their estimation abilities during training by clipping and weighing the small units. Shoop and McIlvain (1963) felt this method was easier because only the number of micro-units in a plot were estimated, rather than the total biomass of the plot. This approach lead to estimations based upon specific observations of plants rather than overly generalized views of plot mass (Shoop and McIlvain 1963). In ten comparisons, the estimated weight averaged 8% less than the actual clipped weights (Shoop and McIlvain 1963). Correlation coefficients between estimated weights and clipped weights averaged 0.87 in the ten comparisons and ranged from 0.74 to 0.96. Furthermore, repeated estimates of plots on successive days demonstrated that observers could consistently estimate plot biomass. However, observers tended to underestimate the mean, predominately because underestimation of high biomass plots was greater than the overestimation of low biomass plots. In addition many smaller species were overlooked during sampling. Shoop and McIlvain (1963) noted that individual observers varied slightly in their ability to estimate biomass. To overcome this, the authors recommended each observer sample a proportionate number of plots in the pasture or experiment. To maintain reliability in micro-units, the authors recommended clipping either the micro-unit or entire plot about every tenth plot. Clipped biomass was then weighed in the field and an observer adjusted his or her estimation technique, bypassing the double sample calibration regression. Reasons given by Shoop and McIlvain (1963) for not using the double sample method included: 1) the increase in time for sampling, 2) the need for a second person to clip and weigh the plot biomass to avoid the introduction of unintended

bias from the learning process, and 3) the need to sample enough plots for valid regression constants.

Tadmor et al. (1975) evaluated the weight-estimate method in semi-arid, annual grasslands and reported methodological issues that had not been previously discussed. The first of these methodological issues was the direct estimation of dry weight rather than the estimation of fresh weight. Occasionally, high daily and seasonal fluctuations in moisture content of vegetation produced large variations in estimates of fresh weight. For example, data collected in February, when all the vegetation was green and moisture content was more consistent, had a coefficient of determination of 0.94, whereas in March and April, when moisture variability was high, the coefficient of determination was 0.38 (Tadmor et al. 1975). The authors suggested that dry weights be estimated directly to overcome the inconsistency of fresh weight estimations. However, estimating dry weights required an observer to have a strong aptitude for consistency in their estimates.

The second issue reported by Tadmor et al. (1975) was nonlinearity noted in calibrating fresh weight to dry weight. Nonlinear regressions were transformed using logarithmic transformations, which improved mean coefficient of determination by 0.20 (Tadmor et al. 1975). However, Ahmed et al. (1983) was critical of the conclusions made by Tadmor et al. (1975) regarding nonlinearity of calibrations and contested that there was an insufficient number of points to conclusively demonstrate a curvilinear relationship and that it was illogical for ocular estimations to consistently change as plot biomass increases.

Tadmor et al. (1975) reported that regressions with direct estimations of dry weight with logarithmic transformations were more consistent than fresh weight estimates with coefficients of determinations that ranged from 0.72 to 0.91. The authors described the method as highly successful at estimating biomass of an area despite high variability in plot biomass, cover, and vegetation height, as long as species composition and phenological stage were consistent. The third issue reported by Tadmor et al. (1975) was that local differences in species composition, growth forms, and phenological stages can decrease the consistency of estimates in a given time period. The authors expressed that experienced observers should be able to overcome these differences, but if not, the authors suggested stratifying calibrations by compositional or phenological stages.

Reese et al. (1980) reported that the three most significant variables affecting double sampling regressions for dry weight-estimation are variations in foliar cover, species richness at the plot level, and average litter weight. The most accurate dry weight-estimations required high variation of foliage cover and species composition with minimal amounts of litter. This is in contrast to the conclusion made by Tadmor et al. (1975) in which low variability in species composition was required for successful dry weight-estimations. Reese et al. (1980) concluded that high variation among plots can compensate for inaccuracies of estimations by untrained observers using the dry weight-estimation method. The importance of the impact of litter on dry weight-estimations may reflect the difficulty of estimating grasses and tall forbs in a grassland, which typically have high litter levels, compared to that of broad leaf forbs in forest understory that have low litter levels (Reese et al. 1980).

Reese et al. (1980) reported that observer experience did not significantly impact double sampling correlations for dry weight-estimates in their study. This generally supports conclusions made by Tadmor et al. (1975) who implied estimation consistency among observers is attributed to natural abilities to estimate rather than experience or training. Reese et al. (1980) also found that the dry weight-estimation method caused mental fatigue.

Ahmed et al. (1983) explored the statistics of double sampling and examined ratio and linear regression techniques for the calibration of indirect estimates using the weight-estimate method. Linear regression requires that data meet the assumptions of homoscedacity (constant variance), have a linear relationship between [x,y] values, and have no assumptions about the intercept passing through the origin. Ratio estimations are similar except that they do not require homoscedacity and have no intercept, so the relationship must pass through the origin (Ahmed et al. 1983). In a comparison between the two methods, the authors found that biomass estimates were similar. Furthermore, the intercepts for linear regression were not significantly different from zero. Despite violations of homoscedasticity in all cases, linear regression had the lower variance estimation compared to ratio estimates, therefore linear regression was the recommended method of calibration (Ahmed et al. 1983). Additionally, interpretation of correlation coefficients and variance is difficult with ratio estimation. However, linear regression can return a positive value for a species that is absent from a plot or a negative value for a species with small plot biomass. This is unimportant though, because the mean of the population is the more desirable value than individual values at a plot level (Ahmed et al. 1983).

Relative Weight Estimates. Hutchings and Schmautz (1969) noted several inherent difficulties in utilizing the weight-estimate methods for biomass estimation, that included 1) extensive training periods required to gain precision and accuracy in estimations with frequent periods of retraining throughout the sampling season, 2) biased estimates that required calibration through double sampling, 3) retained mental images of weight units that do not correspond to current phenological conditions, and 4) the effects of mental exhaustion, mood, and changes in light conditions on estimates. Under the assumption that it is easier to estimate relative weight than it is to estimate absolute weights, several methods were developed that compared sample plot biomass to reference plots or clipped samples.

Reference Plots. The relative weight-estimate method was used to estimate production as a percentage of a nearby base plot (Hutchings and Schmautz 1969). In this method, five 1ft × 2ft plots (30.48 cm × 60.96 cm) are arranged in a cluster similar to the pips of the five-side on a six-sided die with plot centers approximately 3 to 4 ft (0.91 to 1.21 m) apart. Total biomass, life-form biomass (grasses, forbs, and shrubs), and species biomass of the corner plots were all estimated as a percentage of total biomass of the center base plot. After the base plot was harvested, it was used to estimate corner plots as a percentage of the total weight of the base plot. Almost all estimates of total biomass were within 20% of the actual mean with correlation coefficients that ranged from 0.70 to 0.99. Total biomass was the most accurate and consistent estimate in this study. However, estimates of individual plots were highly variable and in two of the three sites sampled, individual plot estimated and actual biomasses were significantly different (Hutchings and Schmautz 1969). In this instance, high variability was not considered of

practical importance by the authors, which reinforces the general conclusion that estimations cannot accurately predict individual plot biomass, but only the mean of an area.

Hutchings and Schmautz (1969) reported that estimating biomass of life-forms of grasses, forbs, and shrubs was less reliable than total biomass, with correlations between actual and estimated biomass that ranged from 0.42 to 0.93. Furthermore, individual species estimations were erratic and unreliable, with the exception of *Festuca idahoensis* and *Erigeron* spp. (genus grouped because of difficulties in species identification) with correlations between actual and estimated biomass ranging from 0.83 to 0.90 and 0.92 to 0.99, respectively. The authors suggested that this was the result of an inability to recognize or see all the species (as many as 15) in a sample plot of the vegetation. Often smaller plants were hidden under taller vegetation and distinction between old and new growth was often difficult. Hutchings and Schmautz (1969) recommended that the base plot be harvested by life-form, rather than total biomass. They also suggested the base plot be selected from the five plots that best represent the others and contain the same life-forms of vegetation as the remaining plots in the cluster. Minor species should also be clumped together in a single estimate to improve biomass estimates. The application of a double sampling procedure, where the occasional estimated plot is randomly harvested, would also correct much of the personal bias in the estimates (Hutchings and Schmautz 1969).

Double sampling regressions used to correct personal bias of the relative weight-estimation method were optimized when pastures had high variability in foliage cover, high species richness, and minimal inter-plot variations in composition (Reese et al.

1980). Plots in heterogeneous, species rich, grasslands were difficult to compare because dissimilar species had different fresh to dry weight ratios (Reese et al. 1980). Despite these difficulties, the relative weight-estimate method appeared to: be reliable and unbiased across all subalpine ecosystems, cause less mental fatigue than dry weight estimation, and not be impacted by observer experience (Reese et al. 1980).

Building off of the success of the relative weight method, the standard pasture core method was created for estimating biomass in closely grazed and short grass pasture that are difficult to harvest (Hutchinson et al. 1972). In this study, eight soil cores were ranked relative to each other and arranged on a circular tray numbered 1 through 8 with a representing the lowest biomass and 8 representing the highest. The soil cores (10.8 cm diameter and approximately "60 cm" [sic] deep) were taken in areas that best represented equal-interval yields found in the pasture. A ninth class of zero was used to represent bare ground. The center of the circular tray contained a hole of equal diameter to the soil core. The soil core tray was carried to and placed over the sample plot so that the vegetation could be viewed through the center and scored according to the closest matching soil core. After the plots were scored, all aboveground biomass was collected from the eight ranked soil cores, dried, and weighed. Mean biomass was calculated by direct conversion of the ranks to weights and double sampling to correct for personal bias in estimations. While using a direct conversion of the pasture cores to weight, Hutchinson et al. (1972) found that: estimates were significantly different among observers, different sets of pasture cores yielded significantly different estimates, and experienced observers estimated no better than completely inexperienced observers. Although the use of double sampling reduced discrepancies between different observers and sets of pasture cores,

and appeared to be superior to direct conversion, Hutchinson et al. (1972) reported difficulties in using the standard pasture core method in stands of vegetation more than 20 cm in height.

In an attempt to overcome the issues associated with using the standard pasture core method to estimate biomass over 20 cm in height, Haydock and Shaw (1975) developed the comparative yield method. In this method, sample plots are compared to a 5 or 9 equal-interval scale of reference plots in the sampling area. This scale should encompass the expected range of vegetation yields for an area and function best if centralized within the sampling area with the quadrats in close proximity to each other (Haydock and Shaw 1975). Total biomass in sample plots was ranked using the reference scale to a degree of 0.25 increments. In a non-double sampling procedure the authors found a significant difference between scales (5 plot vs. 9 plot scales) and observers (Haydock and Shaw 1975). However, with a double sampling procedure differences between observers were reduced and greater accuracy was achieved (Haydock and Shaw 1975). The authors also noted that completely inexperienced observers were able to achieve accurate estimates of mean biomass. Scale plots at either end of the ranked scale (1 and 5 or 1 and 9) were rarely used if they were representative of extreme low or high biomass that was not typically found in the pasture. As a solution to this issue, Haydock and Shaw (1975) recommended selecting ranked scale plot number 1 and 5 or 9 to better represent the plots more likely to be encountered during sampling and to extrapolate the scale to fit plot biomass values at the extremities of low or high biomass. If comparisons between plants with different growth habits are necessary, they recommended that separate calibration lines be created (Haydock and Shaw 1975). The use of photographs

of each ranked scale was also suggested to save time and possibly reduce observer bias during sampling (Haydock and Shaw 1975).

Friedel et al. (1988) evaluated the comparative yield method in arid rangelands on the basis of maintaining consistency in estimates among observers and time requirements. In this study, slopes and intercepts of estimation regressions among observers were significantly different and, because of this, only large differences in biomass could be comfortably detected (Friedel et al. 1988). Friedel and Bastin (1988) suggested that differences among observers in the study conducted by Friedel et al. (1988) were much larger relative to results from Haydock and Shaw (1975) because of the increased heterogeneity of vegetation attributes in arid rangelands. Friedel et al. (1988) reported that the selection of the five-plot reference scale along with collecting data for double sampling calibrations was too time consuming to be effective for monitoring, and suggested the use of photographic reference scales, as described by Friedel and Bastin (1988), instead of selecting reference scale plots in the field. Photographic references of $1m^2$ plots were developed by taking an oblique photo of the plot from a normal standing position, and a vertical photo from the top of a vehicle (Friedel and Bastin 1988). Photos were taken to represent the wide variety of species composition, biomass, height, and structure encountered in the field (Friedel and Bastin 1988). After photos were taken, all standing vegetation in plots were harvested by species, dried, and weighed, and used to calculate relative proportions of species and total biomass.

Photographs, combined with dry yield and species compositional data, can be assembled in a portfolio to use as the five-plot reference scale in the comparative yield method (Friedel and Bastin 1988). Friedel and Bastin (1988) selected one to two

photographs to represent 0.25 scale increments on the reference scale from 0.25-1.5, and in 0.5 scale increments thereafter to a maximum of 5. In this study, observers ranked plots relative to the scale in 0.25 increments despite pictures not being available for all rankings. Differences in regression slopes among observers were not significantly different in grasslands with composition and structure similar to those in the photographic scale (Friedel and Bastin 1988). However, when photographs were not representative of the grassland, differences in slopes were significant among observers. If slopes are consistent among observers, significant differences in regression intercepts, resulting from consistent over or underestimations of yields, could be corrected with a calibration created through double sampling (Friedel et al. 1988).

Using photographic references with the comparative yield method took approximately two hours to sample 100 1 m^2 plots in an arid pasture (Friedel and Bastin 1988), compared to the four hours it took to create a five-plot scale in the field and sample 100 plots in the same area (Friedel et al. 1988). A library of photographs that represented the various conditions throughout a season and the different vegetation types encountered over a region could further increase the efficiency and accuracy of the method. Friedel and Bastin (1988) recommended that observers with limited experience go through an initial training period, and that occasional plots are clipped for calibration purposes. Additionally, the authors suggest using a single observer for vegetation monitoring programs if possible, to maintain the most consistent results.

*Clipped Samples as References***.** Although techniques for estimating aboveground biomass were primarily developed for grasslands, their accuracy and precision with shrubs remained unreported, or were unsatisfactory because of an inability to separate the

biomass of leaves from the stems (Andrew et al. 1979). To address this issue, a double sampling application was developed for estimating shrub foliage biomass, called the 'Adelaide' or reference unit method, which entailed matching clipped samples with biomass in plots (Andrew et al. 1979). In the reference unit method, representative branches, or "reference units", were selected at approximately 10-20% of the total size of a shrub to be sampled. Observers then estimated the number of reference units that comprised the total sampled shrub. A hand-held reference unit was used for a day and then the foliage was stripped off the reference unit, dried, and weighed. Estimated biomass was the product of the reference unit foliage weight and the number of references in the sampled shrub. A regression to correct for personal bias in estimations, using a single reference unit, was created from double sampling ten plots that contained the spectrum of biomass values encountered during the sampling period. The reference unit method accurately estimated saltbush (*Atriplex vescaria* Heward ex Benth) and bluebush (*Maireana sedifolia* (F. Muell.) P. G. Wilson) shrubs with an average r^2 of 0.96 (min $r^2 = 0.84$, $n = 86$) and was used to estimate both grazed and ungrazed shrubs (Andrew et al. 1979). The time required to estimate a single shrub was only 20 seconds excluding travel time between plots, while the time taken to harvest and strip the foliage was up to 25 minutes. Furthermore, the authors found inexperienced observers became competent estimators in as little as two days.

Andrew et al. (1981) re-evaluated the reference unit method and again demonstrated the ability to accurately estimate saltbush (*Atriplex vescaria*) and bluebush (*Maireana sedifolia*) shrubs with an average r^2 of 0.94. However, despite the precise calibrations, individual shrub forage estimates were ± 10% of the actual biomass more

than 60% of the time (Andrew et al. 1981). Consequently, the authors supported the general recommendation that the estimated data be averaged over an area and not to use this technique to estimate the biomass of individual shrubs or plots.

Further expansion of the reference unit method to Northeastern Brazil lead to the validation of the technique to estimate biomass of two widely distributed shrub species: jurema (*Mimosa acutistipula* Benth.) and pau branco (*Auxemma oncocalyx* [Fr. Alem.] Taub.) (Kirmse and Norton 1985). In this study, three observers, one with previous experience using reference units and the other two inexperienced, estimated the biomass of jurema with three reference units to determine how reference size and foliage densities influence estimations. One of the three reference units was approximately 7% of the average shrub size, while the other two were approximately 19%. The two larger references were used to test differences in foliage density of reference unit construction, therefore, one was densely foliated and the other had more disperse foliage. Due to the compact and un-branched growth of pau branco, biomass was estimated for this species using a whole shrub of average size in close proximity ($\leq$19 m). Initial training took only two hours before estimations began. Kirmse and Norton (1985) reported no significant difference in the slopes and intercepts of the estimation equations among observers; however, there was a significant difference among the three reference units. Because of these results, regression equations were combined for all observers rather than fitting one equation to each observer. Although reference units of any density of foliage accurately estimated shrub foliage (r^2 ranged from 0.890 to 0.985), the authors found that larger reference units that matched the foliage density of the measured shrub increased the precision of estimates, which suggests that reference units are best selected to closely

reflect the foliage of shrubs to be estimated. In addition, Kirmse and Norton (1985) suggested that regression equations could be applied to groups of species with similar foliage characteristics.

Due to a lack of validation of the reference unit method for smaller, half shrubs typically found in the United States, Cabral and West (1986) sought to apply the technique to estimating the biomass of winterfat (*Ceratoides lanata* [Moq.] J. T. Howell). Shrub branches that represented approximately 30-100% of the average aboveground winterfat weight were selected to serve as reference units and were used to estimate foliage to the nearest 0.01 unit. No significant difference was found between estimated and actual mean forage biomass, and the coefficients of determinations were high, ranging from 0.839 to 0.993. Although Cabral and West (1986) recommended that the intercept of calibrations be forced through the origin to avoid predicting negative biomasses for small plants, they found no significant difference in average biomass if the intercept was included. Estimating plots using the reference unit method took only 0.6 minutes per plant to estimate, while biomass clipping and weighing took 3.1 minutes per plant (Cabral and West 1986). Therefore, if the reference unit method was applied to a double sampling scheme with a greater number of indirect samples, a substantial amount of time could have been saved. Cabral and West (1986) noted that the weight of smaller plants tended to be slightly overestimated by the reference unit method and that the technique tended to cause mental fatigue after several hours of use. Despite this, the authors reported that the reference unit was an acceptable method for biomass estimation in their study. As an afterthought to the study, Cabral and West (1986) found wrapping

the base of a reference with a wet paper towel surround by polyethylene reduced the potential bias caused by the wilting of reference units.

Carpenter and West (1987) continued exploring the reference unit method by attempting to validate the approach using two shrubs, mountain big sagebrush (*Artemisia tridentate* spp. *vaseyana* [Rydb.] Beetle) and Gardner's saltbrush (*Atriplex gardneri* (Moq.) Dietr.), and for the first time with herbaceous plants, two forbs, *Polygonum aviculare* L. and *Salsola kali* L.. Reference units were selected to be between 10-20% of the average weight for each species. After an hour long training period, the three observers estimated the biomass of 40 individual plants of each species to the nearest 0.1 unit. Based on sampling order, the 40 individual plants were divided into two groups of 20 plants. The first group was used to create a calibration regression, and the second group was used to validate the regression. Precision of the calibration equations was highly variable among the observers (r^2 = 0.39-0.93, $\bar{x}$ = 0.75) and was generally lower than reported by previous validations of reference units (Carpenter and West 1987). Ten of the twelve calibration equations had intercepts not significantly different than the origin and the slopes of all regressions were highly significant. The average percent relative error of individual plants estimated using the calibration ranged from 17.1 to 47.3% error but when all plants were averaged, the percent relative error dropped to 0.2 to 8.0%. Consequently, Carpenter and West (1987) suggested reference units should be applied only to situations where the mean or total biomass of many plants is required, which agrees with previous suggestions by Andrew et al. (1981). The use of reference units requires consistency in estimations over the time period a reference unit is used (Carpenter and West 1987). To determine this, the authors compared regression equations

created from the first group with regressions created using the first and second groups collectively. In 11 of the 12 equations the slopes and intercepts of the two equations did not significantly differ (Carpenter and West 1987). Carpenter and West (1987) suggested further investigations to determine how and under what conditions estimation relationships change over time.

A number of studies have used the reference unit method to estimate herbaceous biomass; however, some authors have provided no details on how calibrations were created (e.g., Newsome et al. 1989; Stoltenberg 2004) or created calibrations for each reference unit using only five plots (e.g., Woodland 2004) and ten plots (e.g., Harrington and John 1990; Noble et al. 2009). Only one study was found that actually validated the use of references units to estimate herbaceous biomass (i.e., Carpenter and West 1987).

Indirect Estimates of Biomass. There have been several efforts to correlate biomass with indirect factors (e.g., vegetation height or canopy cover) as an approach to estimating production yields and total biomass. These efforts included integrated height and cover estimates, disc meters, and visual obstruction readings.

Integrated height and cover estimates. By the late 1950s and early 1960s height and cover was explored as an estimator of production. Pasto et al. (1957) demonstrated that multiple correlations of height, ground cover, and the product of height and cover explained much of the variation in production of tame grasslands. They reported an average correlation coefficient of 0.91 for a bluegrass pasture (assumed *Poa* spp.) and 0.88 for an orchardgrass and Ladino clover pasture (assumed *Dactylis glomerata* L. and *Trifolium repens* L.) using the product of estimated cover and vegetation height (Pasto et al. 1957). The authors demonstrated that regardless of whether cover was estimated

before or after clipping, similar results were found. Despite strong correlations, Pasto et al. (1957) concluded that this method may be impractical because areas of pasture would need to be reserved from grazing in order to measure stand height.

Evans and Jones (1958) reported correlation coefficients ranging from 0.54 to 0.99, which were lower than those reported by Pasto et al. (1957). However, Evans and Jones (1958) used the product of average maximum height and cover rather than multiple correlations with height, cover, and the product of height and cover. In addition, Evans and Jones (1958) compared three different sizes of clipped sub-plots. The lower of the correlation coefficient values were the result of an inadequate proportion of clipped sub-plots relative to the overall plot size (i.e., a single $1ft^2$ circular sub-plot [929 cm^2] for one pasture [size not provided]). With such a small sample size, the maximum height × cover only accounted for an average of 45% of the variability in production among annual grasslands of Northern California. However, plots that contained three clipped subplots ($1ft^2$ circular sub-plot [929 cm^2]) or a 33 ft^2 (3.07 m^2) belt transect increased the average coefficient of determination to 74.1% (Evans and Jones 1958). In a comparison of fertilizer treatments, the combined measurement of height and ground cover appeared to be an adequate index of yield on nearly all stages of growth, as long as plots had a large enough clipped sub-plot. However, relationships between production and average maximum height × cover were not constant throughout the growing season and required several re-calibrations (Evans and Jones 1958). Additionally, as stands became overly mature, relationships became erratic and could not produce reliable estimates of production.

Reppert et al. (1962) later suggested converting height × ground cover measurements into the same terms as production (e.g. kg/ha) by developing a correction term using double sampling procedures. In that study, 84% of variation in yield was accounted for by a multiple regression model that included using height, cover, and the product of height and cover. However, the authors suggested that using combined height and cover methods had little value in estimating absolute yield and instead should be maintained only as an index of relative yields. They argued that statistically significant regressions with high correlations alone do not adequately evaluate the accuracy and precision of estimations, and concluded that tests must also incorporate the many conditions or changing variables observed in the field (i.e. height and cover in this instance) to achieve a degree of stability in coefficients (Reppert et al. 1962).

Although strong correlations between the average maximum height × cover and biomass have been reported (Evans and Jones 1958; Pasto et al. 1957; Reppert et al. 1962), methods using maximum height may not provide reliable estimates of biomass because the volume of a plant is not evenly distributed throughout the height of plant growth (Crafts 1938). A method described by Williamson et al. (1987) to calculate percent basal cover and mean total leaf blade length (a measurement of height) addresses this issue. In this method, vegetation was sampled within $0.7m^2$ square plots containing a 10cm × 10cm string grid. The basal area of grass that intersected the string was measured along the total length of the string grid (13.3m). At alternating intersections of the grid, the nearest tiller was selected and the length of each blade was measured. Biomass was clipped to ground level, oven dried, and weighed. Three double sampling calibration equations were created by correlating biomass to basal cover, mean total blade length per

tiller, and standing crop index calculated from basal cover × mean total blade length. Calibration data were obtained from five dates throughout the 1979 and 1980 growing seasons from areas containing high, intermediate, and low biomass. The standing crop index (a volume measurement) demonstrated a greater coefficient of determination (r^2 = 0.86) over both basal cover (an area measurement, r^2 = 0.62) or mean total blade length (a height measurement, r^2 = 0.74) (Williamson et al. 1987). Data from 1980 demonstrated different year to year regression coefficients and, within 1980, the regression coefficients changed over the season from early summer to mid-summer (Williamson et al. 1987). Because of seasonal changes in dry weight to volume ratios, Williamson et al. (1987) suggested creating new regression coefficients during each sampling period to increase biomass estimation accuracy. They also suggested using this technique for experiments requiring non-destructive estimates of biomass over time. However, with measurements taking up to 20 minutes per plot with a three person team, the technique may not be a suitable estimation replacement for clipping in studies were re-sampling permanent plots is not necessary (Williamson et al. 1987).

Disc Meters. The inherent difficulties of measuring both height and cover lead to the eventual development of a method that could simultaneously estimate both measurements in the field using a single expression (Alexander et al. 1962). In this method, a 2ft (0.61m) square of cardboard was dropped horizontally from waist height. The height of the midpoint of each side of the board was averaged and used as an overall estimate of the board height in the center. The weight of the cardboard is a sensitive test of the combined height and percent cover of the foliage while the resting height measures the forage bulk of the plot (Alexander et al. 1962). Correlations between resting height

and dried biomass were high (mean r = 0.86, n = 14). Similar results were observed with ¼ inch plywood in place of the cardboard (mean r = 0.87, n = 17), suggesting measurements of forage bulk have high predictive ability for the estimation of biomass (Alexander et al. 1962).

Eventually, a tool was created in an effort to simplify the measurement of forage bulk (Holmes 1974). The Massey grass meter (Holmes 1974), or more generally called the disc meter, is constructed by drilling a hole in the center of either a circular or rectangular plate. A tubular handle is attached at the base of the hole and holds the disc perpendicular to a measuring rod placed through the center hole. The disc is allowed to freely move up and down the rod. Forage bulk is measured using a disc meter by dropping the disc from a predetermined height above the surface of soil in a sample plot. Upon settling, a measurement is taken off the center rod. Because disc size and shape match that of the sample plot, only a single height measurement needs to be taken from the soil surface.

Height measurements of the dropped disc and total biomass can be used to create a calibration regression to use during double sampling. This method is very efficient (Douglas and Crawford 1994; Michalk and Herbert 1977) and can lead to fast sample speeds with up to 50 plots sampled in 10-15 minutes (Bransby et al. 1977; Castle 1976). However, with changes in phenology, species composition, and vegetation moisture content, it is recommended that a new regression be created when temporal and spatial variations are noticeable (Bransby et al. 1977; Castle 1976; Holmes 1974; Martin et al. 2005; Rayburn and Rayburn 1998; Vartha and Matches 1977) or a new regression should be created every time an area is sampled (Karl and Nicholson 1987; Santillan et al. 1979).

Discs can be made from wood, plastic, metal, or cardboard, but the use of acrylic plastics was encouraged because it is inexpensive and readily available and does not absorb water, which could change the regression relationship (Rayburn and Rayburn 1998). Changes in disc size and weights did not appear to affect the precision of calibrations in tall fescue stands (Bransby et al. 1977). In another study, increased disc size lead to reductions in errors and increased precision probably because of increased area of contact with biomass and the decrease in sampling errors associated with the determination of what biomass was in and out of a plot (Santillan et al. 1979).

In an attempt to eliminate potential bias caused by gravitational acceleration from dropping the disc among vegetation of varying heights, a method was developed that gently lowered the disc until it smoothly settled on the vegetation (Santillan et al. 1979). The slide down technique resulted in smaller intercepts and decreases in the standard deviation of the double sampling regressions (Santillan et al. 1979).

The rising plate method also decreased variability caused by the falling disc. As the measuring rod is inserted into the plot, the disc is pushed up by the foliage (Harmoney et al. 1997) rather than falling down to the foliage like the drop-disc method. As a result, these reading can be more consistent because the force to resist the plate does not change (Harmoney et al. 1997). Rising plates had a smaller regression intercept and standard deviation than comparably sized drop-disc measurement (Santillan et al. 1979). The rising plate method was easy to use, eliminated the need to bend over during monitoring to lift the disc to the top of the sampling pole (Harmoney et al. 1997), and remained consistent between observers (Earle and McGowan 1979).

Despite the many models of disc meters, the approach often produced strong relationships between yields and measured bulk forage (Santillan et al. 1979; Sharrow 1984) with many authors achieving correlation coefficients greater than 0.90 in tame grassland (Alexander et al. 1962; Baker et al. 1981; Bransby et al. 1977; Earle and McGowan 1979; Michalk and Herbert 1977; Michell and Large 1983; O'Donovan et al. 2002; Santillan et al. 1979; Stockdale and Kelly 1984). However, high variability in regressions were a consequence of management practices that resulted in: high levels of accumulated litter (Karl and Nicholson 1987; Michell and Large 1983; Vartha and Matches 1977), livestock trampling of vegetation (Stockdale and Kelly 1984; Vartha and Matches 1977) and soil (Santillan et al. 1979), and uneven soil surfaces from seed bed preparation (Santillan et al. 1979). High levels of variability can also be produced by site characteristics including soil microrelief (Karl and Nicholson 1987; Sharrow 1984), dominant vegetation (as seen in shortgrass (Sharrow 1984) and bunch grass pastures (Santillan et al. 1979), and overly mature stands (Douglas and Crawford 1994).

Disc methods appear to be less suitable for use in heterogeneous pasture and native grasslands (Tucker 1980). As ratios of forbs to grasses exceeds 10%, undesirable impacts arise for measuring of forage bulk, which include increased sampling error (Karl and Nicholson 1987). In native grasslands, combined height and cover estimates may also encounter increased variation in regression calibrations because diversity in growth forms between species produce variable species-specific relationships between yield and forage bulk (Laca et al. 1989). Accuracy of the method is also dependent on species, as species have different characteristics that make them more or less predictable (e.g. non-jointing grasses tending to be easier to predict than jointed grasses that can bend or break during

sampling) (Harmoney et al. 1997). Still, some native communities may be simple enough for disc meters since reliable relationships ($r^2 = 0.83$) were found on mixed grass prairie in Texas that were reportedly dominated by only a few short to mid-height grasses (Ganguli et al. 2000).

Visual Obstruction Readings. Integrated measures of height and vertical cover, or visual obstruction/ screening effect (Emlen 1956), appear to be a promising indirect measure of biomass (Robel et al. 1970). Visual obstruction can be measured using a Robel pole, which was originally graduated in decimeters. Each decimeter is delineated by a colored band (alternating light brown and white) with halfway marks to measure 0.5 decimeter increments. The visual obstruction reading (VOR) is the lowest decimeter or half decimeter band not completely obstructed from view by vegetation. Robel et al. (1970) demonstrated that a measurement taken at a height of one meter, at a distance of four meters from the pole, will explain most of the variability between biomass and visual obstruction in a tall grass prairie. The addition of a four meter rope or string connecting the observation pole to the measurement pole helps maintain a consistent measurement point (Vermeire et al. 2002).

Placement of the measurement pole depends on plot shape. If a plot is rectangular, the pole should be placed immediately behind the plot on the short side, farthest from the observer, to capture the full plot area (Ganguli et al. 2000; Robel et al. 1970; Vermeire et al. 2002; Vermeire and Gillen 2001). If a plot is circular, the pole is required to be in the center with four readings taken at perpendicular angles and averaged to adequately capture the most vegetation within the plot (Benkobi et al. 2000; Uresk and Benzon 2007; Uresk and Juntti 2008; Uresk et al. 2009). In a general literature review of the VOR

method, Jackson and Paine (2006) suggested that placement on the far side of a rectangular plot will yield the highest coefficient of determination values. Numerous studies have altered bandwidths on measurement poles to 2.5 cm (Higgins 1986; Vermeire and Gillen 2001), 2.54 cm (Benkobi et al. 2000; Volesky et al. 1999), 2 cm (Ganguli et al. 2000; Vermeire et al. 2002), and 1.27cm (Uresk and Benzon 2007; Uresk and Juntti 2008; Uresk et al. 2009). Because of the decreased ability to delineate shorter vegetation heights using measurement increments of 0.5 decimeters, the band width is decreased to maintain precision (Uresk and Juntti 2008).

Double sampling procedures using VOR can be used to create regressions to estimate total plot biomass. The VOR method has proven to be precise, accurate, fast, and easy to learn (Benkobi et al. 2000; Vermeire and Gillen 2001). As a result, the VOR method has been suggested for use on large grassland projects (Benkobi et al. 2000). However, with spatial changes in vegetation structure, separate regression calibrations should be produced for each location of interest (Benkobi et al. 2000; Vermeire et al. 2002), as a general VOR regression calibration may not be possible (Uresk and Juntti 2008). Since the VOR method lacks the ability to estimate single species or functional groups, only total biomass can be estimated. In addition, the distinction between live and standing dead vegetation is not possible, which could reduce the predictive ability of the method. Visual obstruction readings are also strongly influenced by large levels of live or dead biomass that have accumulated beyond the point that the pole is visually obstructed (e.g., dense coastal marshland grasses or high density stands of vegetation) (Whitbeck and Grace 2006). Greater heterogeneity of vegetation, especially that created from

irregularly grazed tall grass prairie, is likely to cause an increase in sample variability, which makes it more difficult to obtain a strong correlation (Jackson and Paine 2006).

The use of transects as the experimental unit with the VOR method has generally created stronger correlations [mean $r^2 = 0.82$ from (Ackerman et al. 1999; Benkobi et al. 2000; Vermeire et al. 2002)] than that of single plots [mean $r^2 = 0.52$ from (Harmoney et al. 1997; Volesky et al. 1999)] (Jackson and Paine 2006). This suggests that VOR and biomass may not always correlate well with single readings on plots, but that multiple readings averaged over an area (e.g. a transect) may result in stronger correlations between VOR and biomass (Jackson and Paine 2006).

Electronic Measurements of Biomass. Some measurements of vegetation attributes can be difficult to visually estimate (e.g., moisture content and canopy structure). As electronic devices were being developed to overcome some of the difficulties associated with visual estimates, their use in double sampling for the estimation of biomass was explored. These efforts included measurements of capacitance and leaf area index.

Measuring Capacitance. Capacitance methods were originally tested for their ability to estimate moisture in grains, bales of cotton, and soils; however, they were eventually recognized as a potential method to estimate forage (Fletcher and Robinson 1956). In using a capacitance meter, the dielectric constant of above-ground vegetation is measured and regressed with the weight of clipped vegetation to create a calibration equation (Fletcher and Robinson 1956). Because a general calibration regression with acceptable accuracy could not be found (Back 1968), Back et al. (1969) suggested using the capacitance method as a double sampling procedure. Although various changes were

made to the capacitance meter (Neal and Neal 1973), they remained relatively insensitive to dry vegetation and overly sensitive to succulent plants. In order to resolve these issues, Currie et al. (1973) suggested stratifying regressions by vegetation characteristics. Despite stratification, the double sampling procedure paired with the capacitance meter produced inconsistent regressions in native annual range vegetation. In vegetated areas with various states of phenology and dryness, capacitance probes explained 48% of the variability, compared to the 84% explained in wetter, more uniform green sites (Neal et al. 1976). Differences in chemical makeup of a species were also noted to influence readings as changes in species proportions are encountered (Neal et al. 1976). This evidence suggests that capacitance meters may be better suited for use in mostly homogenous vegetation (Terry et al. 1981), such as tame pastures, where species composition, vegetation densities, and heights have limited variability (Jones et al. 1977).

Early models of capacitance meters mostly detected differences in plant moisture content. They had either multiple plates or probes that used a combination of vegetation and air acting as the dielectric. Together, the vegetation and air would produce variable dielectric constants since vegetation can have inconsistent moisture content. Because of this, later models consisted of a single probe capacitance meter that detected differences in leaf surface area rather than moisture content (Vickery et al. 1980). The single probe model used a probe as one plate in the circuit, and the surrounding grass as the other plate with the air between acting as the dielectric. While using the probe, the variable component being measured is the leaf and stem area rather than the moisture in the vegetation (Vickery et al. 1980). Meter calibrations could be corrected by taking a measure of the air capacitance above each plot and subtracting that from the plot reading.

This calibration increased the coefficient of determination from 0.75 to 0.85 (Vickery et al. 1980). Although the modified meter was less sensitive to changes in moisture, it would not function in vegetation that was so dry that it no longer acted as an electrical conduit (Vickery et al. 1980). Vickery et al. (1980) suggested that general calibrations may be suitable for monitoring, but recommended recalibrations monthly or when large changes in vegetation are noticed during grazing research. Currie et al. (1987) took recalibration a step further and suggested creating a new calibration through double sampling at each sampling date and location.

Serrano et al. (2011) evaluated the single probe capacitance meter in Mediterranean pastures and reported that calibration regressions were variable among homogeneous pastures, heterogeneous pastures, and legume-planted pastures ($r^2 = 0.90$, $r^2 = 0.87$, and $r^2 = 0.48$, respectively). Generally, poor calibration equations where also produced in pastures with heterogeneous species composition, unevenness in ground, and trampling and lodging of vegetation (Sanderson et al. 2001; Stockdale and Kelly 1984). Due to the degree of uncertainty in calibrations, Serrano et al. (2011) suggested that the single probe capacitance meter be approached with some reservation. Reese et al. (1980) reviewed the capacitance probe and other popular estimation techniques used in double sampling and found that if [x, y] values are clumped together, the relationships between [x, y] values becomes less defined; consequently, the validation set does not truly reflect the variability of the larger estimated set. This is counter to recommendations made by Haydock and Shaw (1975) and Tadmor et al. (1975) to purposely select a small scale validation set during double sampling. Species diversity and large variance in foliage cover between quadrats can lead to a greater scattering of [x, y] values in the capacitance

probe calibration regression, which tends to increase the coefficient of determination, as long as there is a relationship between [x, y] values (Reese et al. 1980).

*Canopy Analysis***.** The position, orientation, shapes, and area of the stems, leaves, and flowers is typically referred to as canopy structure, which plays a valuable role in the description of vegetation and environmental interactions (Welles and Norman 1991). The leaf area index (LAI) is an expression of leaf area per unit of ground cover and is frequently used for the measurement of canopy structure (Welles and Norman 1991). In stands of orchardgrass (*Dactylis glomerata* L.), Pearce et al. (1965) observed a linear relationships between net photosynthesis and LAI. Linear relationships between biomass and LAI have also been found on stands of smooth bromegrass (*Bromus inermis* Leyss.) and tall fescue (*Schedonorus phoenix* (Scop.) Holub) (Engel et al. 1987; Trott et al. 1988). Collectively, this lead Harmoney et al. (1997) to believe that LAI could be a useful tool for the indirect measurement of biomass in double sampling. The LAI method uses a fish-eye optical sensor to discern canopy gaps by averaging measurements of light interception below the canopy and compares it to one measurement above the canopy. The number of measurements taken in a plot is determined by the degree of heterogeneity of the canopy, but is typically 10 or less (Welles and Norman 1991). After measurements are taken, the built-in computer attached to the sensor calculates the LAI value (Welles and Norman 1991). These measurements of LAI include any plant part or object that is opaque, such as stems, fruits, flowers, and leaves; however, the sensor cannot distinguish between live plant material from dead (Welles and Norman 1991). Harmoney et al. (1997) found that double sampling regressions between LAI, using eight below canopy readings, and biomass only explained 32% of the variability in biomass of tame pasture

biomass in Iowa. In a similar study, Miller-Goodman et al. (1999) found LAI, using five below canopy readings, explained only 43% of the variability of biomass in rangeland of the Nebraska Sandhills. In another study in the Nebraska Sandhills, Volesky et al. (1999) found that LAI accounted for 33% of biomass variability when averaging three below canopy readings; however, when below canopy measurements increased to eight, LAI accounted for 59% of the biomass variability. In the shortgrass plains of Texas, LAI could only explain 34% of biomass variability at a plot level, and improved to 67% when an entire pasture was averaged (Ganguli et al. 2000). Ganguli et al. (2000) concluded that VOR was a superior method for estimating biomass on shortgrass prairie compared to the poorly preforming LAI method.

Because LAI uses various canopy properties to estimate the leaf area, Whitbeck and Grace (2006) suggested that light penetration (LP), which is simply the percentage of ambient light that reaches the ground, could provide a better estimate of biomass compared to LAI. Light penetration measurements explained 65% of the variance in clipped plot biomass in coastal marshlands (Whitbeck and Grace 2006). Multiple regression increased the coefficient of determination to 0.75 when LP measurements and percent canopy cover of the dominant marshland species, marshhay cordgrass (*Spartina patens* (Ait.) Muhl.), were included (Whitbeck and Grace 2006). However, the steep learning curve associated with the technical requirements for programming and retrieving data (Ganguli et al. 2000) and the general poor performance of the LAI and LP methods (Ganguli et al. 2000; Harmoney et al. 1997; Miller-Goodman et al. 1999; Volesky et al. 1999) may limit their widespread use.

Individual Biomass Contributions for Determining Ecological Relationships

Many of the techniques described above can only estimate total biomass, which limits their utility. The dry-weight rank method can estimate individual species contributions to total biomass (Mannetje and Haydock 1963). When the dry-weight rank method is used along with a method for estimating total biomass, individual species biomass can be estimated. In the dry-weight rank method, the top three species or species groups contributing to dry-weight production are ranked first, second, or third with the remaining vegetation unranked. The proportion of plots sampled in which each species was given either a rank of first, second, or third are calculated and multiplied by factors of 0.7019, 0.2108, and 0.0873 respectively (Mannetje and Haydock 1963). These factors were determined using least squares from 18 data sets of clipped plots hand sorted into 4-5 groups in which the exact ranking within each plot could be obtained and the number of plots ranked first, second, and third could be calculated (Mannetje and Haydock 1963). These three products are added to give a percent of dry-weight contribution of the species. For example, if species A was ranked first in 50/100 plots (0.50), second in 30/100 plots (0.30), and third in 20/100 plots (0.20) the contribution to total production of species A would be 0.50*0.7019+0.30*0.2108+0.20*0.0873 = 0.432 or 43.2%.

A minimum of three species are necessary for the dry-weight rank method to work and minor species may never have a high enough presence to obtain a ranking, especially in stands with high species diversity. However, minor species could be collectively grouped to obtain sufficient levels for ranking (Mannetje and Haydock 1963). Walker (1970) applied the approach in Rhodesian grasslands and reported that few species received ranks and that weight contributions could not be accurately estimated because the unranked species, although individually insignificant, had combined weights

that were “undoubtedly significant”. In three tests of the dry-weight-rank method using the top four species, only three of the 12 cases had significant differences between dry-weight-rank and hand sorted ranks ranging from 3.6% to 1.9% relative difference, which was considered too small to be of practical concern (Mannetje and Haydock 1963).

The advantage of estimating ranks rather than weight is that an observer only determines whether a species has a greater weight than another species. Despite this advantage, estimating ranks can be difficult because of large differences between the weight to volume ratios of differing species with variable growth habits (Mannetje and Haydock 1963). However, it is easier and faster to train observers to rank species than to directly estimate actual weight or the percent of weight directly (Mannetje and Haydock 1963).

Jones and Hargreaves (1979) noted that an underlying assumption made by Mannetje and Haydock (1963) is that there is no consistent relationship between plot yields and species ranks. That is, a plot dominated by species A, should not have a consistently higher yield than a plot dominated by species B. However, in some grasslands, consistent relationships between yield and composition were evident (Jones and Hargreaves 1979). In this study, dry-weight-rank was modified using a weighting-factor calculated by multiplying the plot yield, either clipped or estimated, by the multipliers for each component (Jones and Hargreaves 1979). Typically, if data on species composition are collected, then total production is also calculated; therefore, including a weighting-factor would increase costs negligibly. In the dry-weight rank method as described by Mannetje and Haydock (1963), a single species could only receive a rank of first, second, or third and the highest any single species could contribute

to total yield was 70%. However, in certain conditions (e.g., simple polyculture pasture) total yield of a single species can exceed 70%. To address this, the use of "cumulative ranking" was recommended by Jones and Hargreaves (1979) , in which a dominate species could be ranked as a combination of first, second, and third in a plot. However, this only addresses irregular problems and if frequent use is required in an experiment it is best to estimate the percentages directly (Jones and Hargreaves 1979). Independently, Jones and Hargreaves (1979) calculated a new set of multipliers for their study, arriving at 0.714, 0.247, and 0.039 for first, second, and third, respectively. Because this new set of multipliers and those presented by Mannetje and Haydock (1963) were similar, the two data sets were combined to further refine the multipliers to 0.705, 0.238, 0.057; thus expanding the data set to include 12 locations, rather than the original four locations (Jones and Hargreaves 1979).

Gillen and Smith (1986) evaluated the dry-weight-rank method and found that the original multipliers described by Mannetje and Haydock (1963) were adequate for three of the four tallgrass prairie sites they surveyed. In one of the sites, a single species received the first rank in 90% of plots; however, with limited variability in ranking between plots, the sensitivity of the method was reduced and the percent composition could not be accurately determined (Gillen and Smith 1986). Gillen and Smith (1986) noted that the amount of error depended on observer experience, predominately because of the difficulties associated with comparing grouped uncommon or rare species to that of a single dominate species. Bbservations by Gillen and Smith (1986) support Mannetje and Haydock (1963) conclusions that estimating dry-weight ranks requires training. Time spent ranking plots was approximately 1.6 minutes, compared to the 28 minutes required

for hand clipping and sorting (Gillen and Smith 1986). Despite Mannetje and Haydock (1963) multipliers being derived from different ecosystems, the multipliers gave acceptable estimates of percent composition in a tallgrass prairie; however, the use of weighted multipliers did not improve estimates (Gillen and Smith 1986). Similar conclusions for the use of Mannetje and Haydock (1963) multipliers, which were not weighted, were reached for both arid rangeland (Friedel et al. 1988) and desert scrub brush (Mazaika and Krausman 1991).

Friedel et al. (1988) evaluated the dry-weight rank method in arid rangeland and reported that increasing the number of plots sampled, increased the consistency of estimates among observers. Slight increases in similarity among observers were also noticed with the use of fixed plots rather than plots selected by each individual observer. However, by incorporating fixed plots into monitoring projects, the time benefits of the method may be lost to the extra time spent finding plot locations (Friedel et al. 1988).

Cost Benefit Analysis

During a double sampling procedure, the ratio of indirect samples to direct samples strongly influences the efficiency of sampling. If too many plots are sampled directly, the costs become unnecessarily high. If too few plots are sampled directly, the regression becomes unreliable. To properly balance the ratio of direct to indirect samples, an optimum ratio equation for single factors was developed by Wilm et al. (1944) and modified by Cochran (1963).

$$C = nc_n + n'c_{n'}$$

Where

$n' = number\ of\ indirect\ samples$

$n = number\ of\ direct\ samples$
$c_n = cost\ of\ obtaining\ one\ direct\ sample$
$c_{n'} = cost\ of\ obtaining\ one\ indirect\ sample$
$C = total\ cost\ of\ sampling$

For a fixed cost C, an optimum ratio of indirect to direct samples is obtained by

$$\frac{n'}{n} = \sqrt{\left(\frac{c_n \rho^2}{c_{n'}(1-\rho^2)}\right)}$$

Where

$\rho = coefficient\ of\ correlation$

Multivariate double sampling designs are more complicated. A technique was described that can find the optimum ratio of indirect to direct samples for a multivariate double sampling that can either minimize costs or variance (Ahmed and Bonham 1982; Ahmed et al. 1983). An optimum ratio problem with k variates will:

$$\text{Minimize } C = nc_n + n'c_{n'}$$

$$\text{Constrained by } \frac{V_{n_j}}{n} + \frac{V_{n'_j}}{n'} \leq V^{\circ}_j$$

$$(j = a, \ldots \ldots, k)$$

$$n' > n > 0$$

Where

$V_{n_j} = Variance\ of\ direct\ sample\ j$
$V_{n'_j} = Variance\ of\ indirect\ sample\ j$
$V^{\circ}_j =$ Specified variance tolerance for the mean of the j^{th} $variate$

The variance tolerance can be specified from past experience or literature, the coefficients of variation for $\bar{y}$, or setting a bound of 2*standard error of estimate.

METHODS

Study Area

The study was conducted on the Buffalo Gap National Grassland (BGNG) located in Custer, Fall River, Jackson, and Pennington counties in southwestern South Dakota (Figure 1). Its 240,000 hectares are intermixed with Badlands National Park, Pine Ridge Indian Reservation, and state and privately owned land (USDA U.S. Forest Service 2006). The BGNG is administered for multiple use that includes, but is not limited to, sustainable livestock production, habitat preservation, and recreation.

The regional climate is continental and semiarid with extreme seasonal temperature variation and low precipitation. Climatic data for three sites, Oral, Cottonwood, and Edgemont (Figure 1), encompassing the range of the BGNG were compiled. A majority of precipitation in the area falls between April and September (Table 1). During the study period of August 2009 to July 2010, annual precipitation was 62 and 25% higher than long-term averages at Oral and Cottonwood respectively (Figure 2). The majority of higher than normal precipitation recorded at Oral during the study period occurred during April, May, and June. This is a sharp contrast to the Cottonwood and Edgemont stations, which had near normal precipitation during that timeframe. Annual precipitation recorded at Edgemont for the study period more closely approximated the long-term averages.

The landscape of the BGNG is predominantly native grassland on gently rolling sedimentary plains with scattered buttes and un-vegetated and eroded badlands. Dominant graminoid species include western wheatgrass (*Pascopyrum smithii* (Rydb.) A. Löve), green needlegrass (*Nassella viridula* (Trin.) Barkworth), blue grama (*Bouteloua*

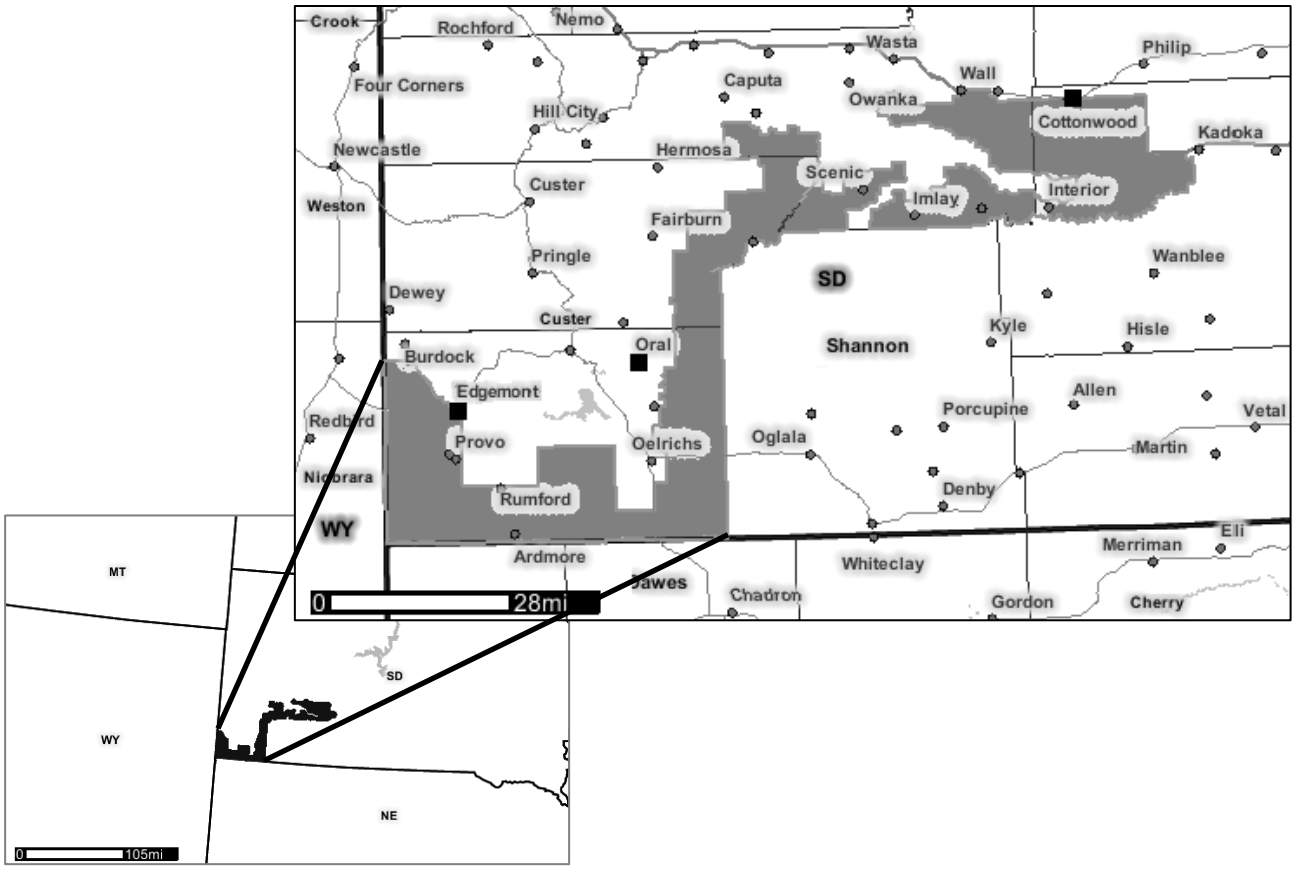

Figure 1. Map of study site in South Dakota with location of Buffalo Gap National Grasslands (BGNG) in gray and locations of weather stations at Oral, Cottonwood, and Edgemont represented by black squares (USDA NRCS 2010).

Table 1. Temperature and precipitation. Thirty-year averages (1971 to 2000) and 2009 to 2010 annual temperature (°C) and precipitation (cm) for Cottonwood, Edgemont, and Oral Automated Weather Data Network (AWDN) weather stations in South Dakota (High Plains Regional Climate Center 2011) (http://www.hprcc.unl.edu/index.php). Precipitation values based on August through July.

	30-year Average 1971-2000				2009-2010	
	Temperature		Precipitation		Precipitation	
	max	min	Total Annual	% Apr-Sept	Total Annual	% Apr-Sept
Cottonwood	22.8	-7.4	43.6	74%	54.5	82%
Edgemont	22.9	-6.5	41.3	74%	41.4	68%
Oral	22.7	-4.6	42.5	77%	68.8	76%

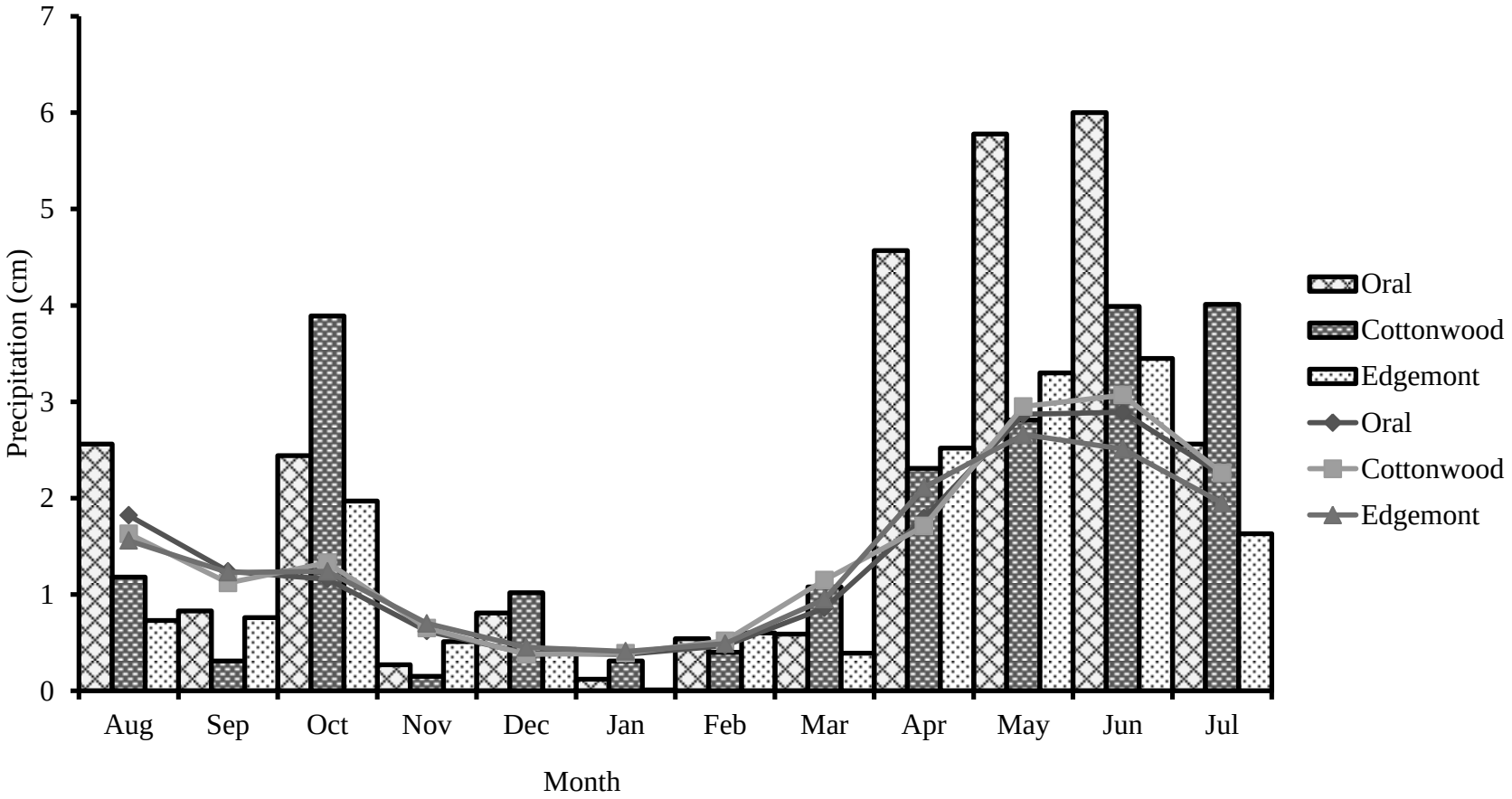

Figure 2. Thirty-year average (lines) (1971-2000) and August 2009- July 2010 bar graph of precipitation (cm) for Cottonwood, Edgemont, and Oral Automated Weather Data Network (AWDN) weather stations in South Dakota (High Plains Regional Climate Center 2011) (http://www.hprcc.unl.edu/index.php).

gracilis (Willd. ex Kunth) Lag. ex Griffiths), needle and thread (*Hesperostipa comata* (Trin. & Rupr.) Barkworth), threadleaf sedge (*Carex filifolia* Nutt.), needleleaf sedge (*Carex duriuscula* C.A. Mey.), little bluestem (*Schizachyrium scoparium* (Michx.) Nash), and buffalograss (*Bouteloua dactyloides* (Nutt.) J.T. Columbus) (Johnson and Larson 1999). Common forbs include 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.), Missouri goldenrod (*Solidago missouriensis* Nutt.), prickly pear (*Opuntia polyacantha* Haw.), while shrubs included western snowberry (*Symphoricarpos occidentalis* Hook.), leadplant (*Amorpha canescens* Pursh), and fringed sagewort (*Artemisia frigida* Willd.). A recent plant species inventory reported a total of 530 species within the BGNG (Kostel 2006).

Study Site

Nine sites were sampled 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. Prairie dog colonies were included because the foraging, clipping, and burrowing activities of the prairie dogs create vegetation heterogeneity at a variety of spatial scales. Specific colonies were selected based on size and accessibility, consistency of soil characteristics both on and off the colony, and an objective evaluation of a gradient of ecological condition and plant species diversity. The sites were restricted to the Clayey (four sites) and Loamy (five sites) ecological sites of the Western Great Plains Range and Irrigated Region-Pierre Shale (060A) Multiple Land Resource Area (MLRA)

as defined by the Natural Resource Conservation Service (NRCS) because these two ecological sites represent a major component of this 2.63 million ha MLRA. All sample sites were grazed at recommended stocking levels using NRCS guidelines for each ecological site.

Sampling

Sampled sites were stratified into off colony locations and interiorand edge for the on colony locations. The interior was toward the center of the colony, which was typically the oldest portion of the colony and contained a high forb component (Bonham and Lerwick 1976). The edge was still within the colony but around the outer portions and typically had a greater graminoid component than the interior, but less than off colony. Off colony locations were within the same ecological site as the colony and had an absence of prairie dog activity and were generally between 100 m to 200 m from the perceived edge of the colony.

Four, 100 m transects (each transect serving as the experimental unit) were randomly placed within each location (interior, edge, and off-colony) for a total of 12 transects at each sample site (Figure 3). A total of 108 transects were sampled in 2010 on the 9 colonies. Circular sample plots (0.25 m^2) were placed at 20 m intervals along each transect. Transects were chosen as the experimental unit because individual plot level estimations tend not have strong relationships between direct and indirect measurements, but when averaged over an area, such as a transect, relationships can improve (Ganguli et al. 2000; Jackson and Paine 2006).

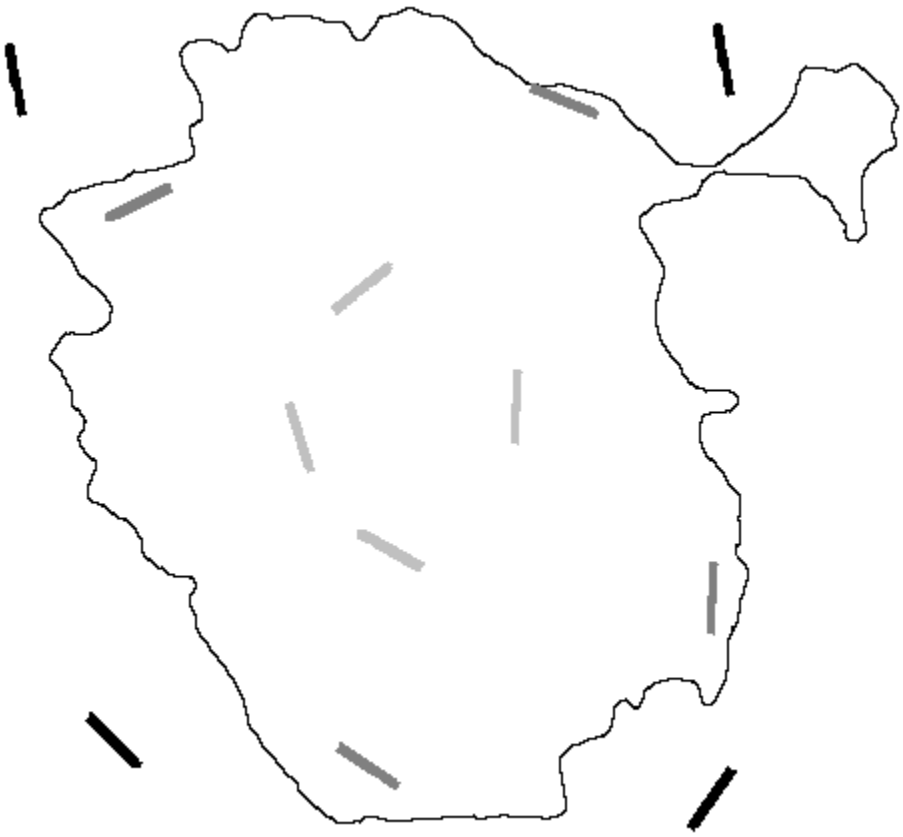

Figure 3. Transect layout. Colonies were stratified into 3 locations with 4 transects placed in each location: Interior (light gray), edge (dark gray), and off (black).

The biomass of individual species and functional groups was estimated using reference units within each plot. Individual species include western wheatgrass (PASSMI), green needlegrass (NASVIR), purple threeawn (*Aristida purpurea* Nutt., ARIPUR), needle and thread (HESCOM). Multiple species functional groups include 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). Functional groups were based on life form (forb or graminoid), origin (native or introduced), and life span (annual or perennial). The use of functional groups was necessary to improve the estimates of minor species (Hutchings and Schmautz 1969; Mannetje and Haydock 1963), as they collectively could be significant in a study of ecological relationships (Walker 1970) but difficult to estimate individually. Individual key species were selected based upon the dominate species in descriptions of the plant communities in the state and transition models published by the NRCS for the Clayey and Loamy ecological sites. Species and functional groups were referenced only if their cover was above 1% foliar cover using Daubenmire cover classes (Daubenmire 1959). One percent was used as a threshold for inclusion because very small plants were difficult to both estimate and harvest. Total biomass was not estimated independently, but was calculated by summing all estimates on a plot and averaged by transect.

Reference units were selected to best represent the current growth and phenological characteristics of the functional groups being sampled (Kirmse and Norton 1985). Biomass of individual species and functional groups within each plot was

estimated as a ratio of the reference unit in increments of 0.1 of a unit. New reference units were collected when changes in phenological stages, height, or foliage density were noticeable and if unit size required estimates higher than 10 units or less than 0.1 units. This helped ensure that reference units remained a consistent proportion of the actual biomass in the plot. In other words, as actual biomass increased or decreased in a series of plots so did the reference units. Such adjustments may also reduce the potential for "mental fatigue" (Cabral and West 1986) that is often associated with either estimating large amounts of biomass with small reference units or small amounts of biomass with large reference units. When functional groups were composed of multiple species, the most dominate species of that functional group in the plot was selected as the reference unit. This made it easier to evaluate different species collectively compared to having a mixed reference unit that contained many species. The difference in foliage density, size, or shape and the densities of stem biomasses were easier to compare when only working with a single species reference unit.

A modification of the reference unit technique was required for use with the shortgrass group because the vegetation was often too short to develop a reference unit that held together and could be easily compared to a plot through sight and touch. Building upon the success of Williamson et al. (1987), who used average height multiplied by percent cover to estimate biomass of blue grama, a modified height × cover method using reference units was developed for the shortgrass functional group. To create a shortgrass reference unit, a patch of shortgrasses that was approximately 5% of the total area of the plot with foliage density that approximated a lawn was located, clipped, and used as the reference unit for that site. Stolons were harvested if they were

not rooted at the nodes. The goal was to use a reference unit that approximated the biomass of average height shortgrass in a 5% area of the plot. To use the reference unit required a visual confirmation that the area was of similar average height to the reference unit. If the heights were not similar, another reference unit would be harvested to match the new average height. To determine the number of reference units in a plot, clumps or patches of shortgrasses were visually compiled together in an effort to match the density and size of the reference unit. Using a cupped hand as a standard size for approximating 5% of plot helped maintain consistency during reference unit creation and use.

To reduce wilting, reference units were stored in clear zip top bags to limit moisture loss and accidental damage to foliage during sampling. The use of clear bags also allowed for easy visual comparisons to the reference unit. If kept out of the sun and refrigerated overnight, the references typically lasted for two to three days. However, the longer a reference unit was used, the greater the potential for data loss if the unit was misplaced or destroyed. A plastic, expandable file-folder helped organize references, kept them out of direct sunlight, and prevented their misplacement.

After plots were referenced, biomass within plots was clipped to ground level, collected by species or functional groups, oven dried for 72 hours at 60°C, and weighed to the nearest 0.01g. Three observers would rotate 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 experience using the method prior to training, and Observer 3 had received brief training on the method in an undergraduate course. To look at temporal variations in prairie dog sites for a larger study, a site was sampled twice

during the single season. Plots on the sites sampled twice were marked with colored flags placed in the center of the plot to allow easier transect placement during a second sampling. Since the initial plots were previously harvested, new plots were moved five meters further along transects during the second sampling period.

Data Analysis

Regression analysis is a widely used data analysis method for developing predictive or calibration models (Snee 1977). Although the fit of the regression to the data is important, a high coefficient of determination is not enough to validate a regression (Snee 1977). One approach that can be used to validate a model is to use a portion of the available data in an independent assessment of the accuracy of predictions from the regression (Snee 1977). Proper data splitting requires using a sufficient portion of the data to create the regression while maintaining enough points to assess the fitted regression using a validation set (Picard and Berk 1990). Most optimal data splits reserve approximately 25 to 33% of the data for the validation set (Picard and Berk 1990). Data sequentially collected over a period of time can be easy split by picking a point in time to separate the estimation set used to create the regression and the independent validation set (Snee 1977).

To create the calibration regression in this study, transects were separated by observer and by functional groups within each observer. Referenced and harvested weights on plots were averaged at the transect level. Regression data were examined using ordinary least squares with the REG procedure in SAS, SAS Institute Inc., Cary, NC, with the influence option. 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 samples, were investigated as outliers and possibly removed (SAS Institute 2008). Results from q-q plots found data had normal distributions. Transects for each of the observers were divided by time into groups with approximately the first 65% of transects (24-26 transects per observer for each key species and functional group and 74-77 transects per key species and functional group pooled by observer) used in a calibration regression building set and the last 35% used in a validation set (13 to 15 transects per observer for each functional group and 41-41 transects per functional group pooled by observer). With the data split based on time, validation of the model confirms consistency of estimations over a season.

Linear regressions for the calibration sets were created using weighted ordinary least squares with the REG procedure in SAS. Linear regression was chosen over ratio estimation because, although linear regression can return a positive value for a species that is absent from a plot or a negative value for species with low plot biomass, the mean of the population is the most desirable value (Ahmed et al. 1983). Additionally, ratio estimation is difficult to interpret regarding the correlation coefficient and variance (Ahmed et al. 1983). Weighting the procedure was necessary because the data did not meet the assumption of constant variance (homoscedasticity) required for ordinary least squares procedures (Ganguli et al. 2000; Zar 2010). Because it is unlikely that ocular estimations will consistently change as plot biomass increases (Ahmed et al. 1983), statistically significant curvilinear relationships ($\alpha = 0.05$) were ignored unless there was a substantial number of points conclusively demonstrating a curvilinear relationship (Andrew et al. 1979). Regressions with substantial number of points demonstrating curvilinear relationships were transformed to establish be best linear fit.

The validation data set was then calibrated using the created regressions. A paired t-test in SAS was used to compare calibrated reference weights to the actual clipped weights of the transects for each observer and plant functional group. Models were considered valid if the mean difference between reference weight and actual weight was not significantly different from zero ($\alpha = 0.05$). Creating regressions was not possible for species and functional groups that were uncommon among transects. Within these species and functional groups, non-calibrated estimation values 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 be pooled to decrease sampling costs and increase precision of measurements (Laca et al. 1989). Testing for coincident lines between observers and key species and functional groups can determine if data can be pooled across certain species or observers. A two by two between groups analysis of covariance (ANCOVA) was conducted using a weighted general linear model in IBM SPSS Statistics 19, SPSS Inc., Chicago, IL. The dependent variable was clipped biomass, with two independent variables of observer and key species and functional groups. Estimated biomass was used as the covariate. Regression lines were coincidental if there was no significant difference among groupings.

A multivariate cost benefit analysis using methods described by Ahmed and Bonham (1982) was calculated. This cost analysis was constrained by the coefficients of variation for functional groups compiled across observers with the time invested in clipping taking five times longer than estimation using references.

RESULTS

All 26 functional group and 3 total biomass regressions had a significant relationship ($P<0.05$) between estimated biomass (x) and actual biomass (y) (Table 2). A scatterplot of estimated and actual biomass suggests a curvilinear relationship for PASSMI regression for Observer 2 (Figure 4a), which was addressed by a square root transformation to approximate a linear relationship (Figure 4b). Intercepts of regressions were significantly different from zero ($P<0.05$) for all three observers for Total biomass and Observer 2 for PASSMI (Table 2).Intercepts of all remaining regressions were not significantly different than zero ($P>0.05$). The coefficient of determination for all regressions ranged from 0.61 to 0.97 with a grand mean of 0.86. Observers 1 and 3 had similar coefficient of determination for regressions, with Observer 1 having a range of 0.67 to 0.95 (mean = 0.88) and Observer 3 having a range of 0.81 to 0.95 (mean = 0.88). In contrast, Observer 2 had slightly lower coefficient of determination values compared with Observer 1 and 3 with a range of 0.61 to 0.97 (mean = 0.80). Two groups of species shared the highest average coefficient of determination, PASSMI and SHORT with averages of 0.91. PGN had the lowest average coefficient of determination with 0.73.

Accuracy of direct estimates using references were evaluated by visually comparing equation slopes relative to unity ($y = 1x$) (Figure 5). For the most part, predicted biomass of each observer tended to be scattered around unity, with some overestimating (regression below unity) and underestimating (regression above unity). All three observers appeared to have estimates of AGI which produced slopes near 1 with relatively high levels of precision (r^2 range 0.77 to 0.90). Additionally, all three observers

Table 2. Regression equations [x = estimated biomass, y = actual biomass; grams/0.25m^2] for functional groups by three observers. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Functional groups are as follows: AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss and blue grama, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN =perennial forb introduced, PGN = perennial grass native. Total = total biomass. *n* = number of transects. Outliers = number of transects excluded from data analysis. P values for intercept (int) and slope (X), MSE = Mean square error, r^2 = coefficient of determination. * curvilinear relationship transformed using $\sqrt{x}$.

					P-value			
Group	Tech	*n*	Outlier	Equation	int	X	MSE	r^2
AFI	1	25	1	y=-0.01+1.50x	0.880	≤0.001	0.31	0.88
	2	25	0	y=-0.24+0.87x	0.273	≤0.001	0.66	0.61
	3	26	0	y=-0.15+1.12x	0.388	≤0.001	0.18	0.87
AFN	1	26	0	y=-0.05+1.43x	0.621	≤0.001	0.55	0.93
	2	25	1	y= 0.15+0.91x	0.374	≤0.001	1.03	0.78
	3	26	0	y=-0.02+1.08x	0.907	≤0.001	0.37	0.95
AGI	1	25	0	y=0.16+0.95x	0.600	≤0.001	1.19	0.90
	2	25	0	y=0.56+1.03x	0.235	≤0.001	1.14	0.77
	3	26	0	y=0.67+0.91x	0.212	≤0.001	1.32	0.88
AGN	1	25	1	y=0.11+1.55x	0.770	0.003	0.65	0.85
	2	25	0	y=-0.08+1.03x	0.610	≤0.001	0.18	0.97
	3	25	2	y=0.03+0.73x	0.766	≤0.001	0.23	0.83
SHORT	1	25	1	y=0.25+0.79x	0.330	≤0.001	0.62	0.95
	2	25	1	y=0.39+0.94x	0.161	≤0.001	1.21	0.91
	3	26	0	y=0.36+0.77x	0.375	≤0.001	1.60	0.87
PASSMI	1	25	1	y=0.44+1.16x	0.144	≤0.001	1.05	0.95
	2*	25	1	y=-3.74+5.06x^1/2	≤0.001	≤0.001	17.64	0.90
	3	26	0	y=0.80+0.95x	0.100	≤0.001	1.84	0.88
PFI	1	25	0	y=-0.10+2.06x	0.238	≤0.001	0.06	0.91
	2	24	2	y=-0.04+0.81x	0.242	0.005	0.23	0.64
PFN	1	25	0	y=0.08+1.35x	0.705	≤0.001	0.64	0.87
	2	24	1	y=-0.01+0.97x	0.940	≤0.001	0.30	0.91
	3	26	1	y=0.14+0.93x	0.305	≤0.001	0.16	0.92
PGN	1	24	1	y=0.21+1.08x	0.294	≤0.001	0.49	0.67
	2	25	1	y=0.35+0.72x	0.158	≤0.001	0.76	0.71
	3	26	1	y=0.18+1.03x	0.212	≤0.001	0.40	0.81
TOTAL	1	25	1	y=7.87+0.89x	0.010	≤0.001	1.80	0.87
	2	25	1	y=4.11+0.80x	0.017	≤0.001	2.05	0.92
	3	26	1	y=4.84+0.84x	0.033	≤0.001	1.40	0.92

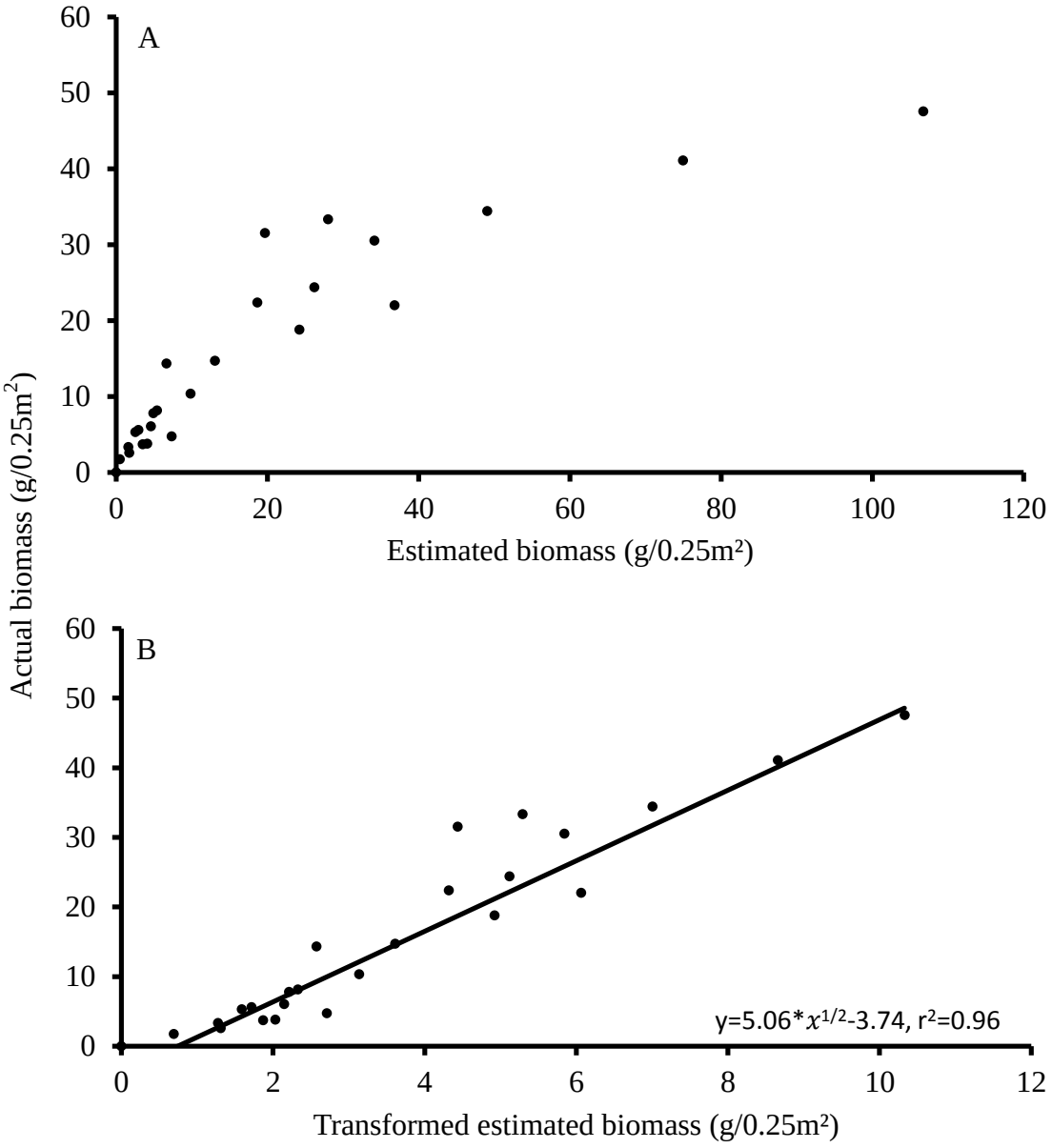

Figure 4. PASSMI regression for Observer 2 with transformation. Note the different x-axis scales. Curvilear relationship is evident on top graph (A). Regression equation with square root transformation approximates a linear relationship (B). Data were collected in 2010 on the Buffalo Gap National Grasslands, SD.

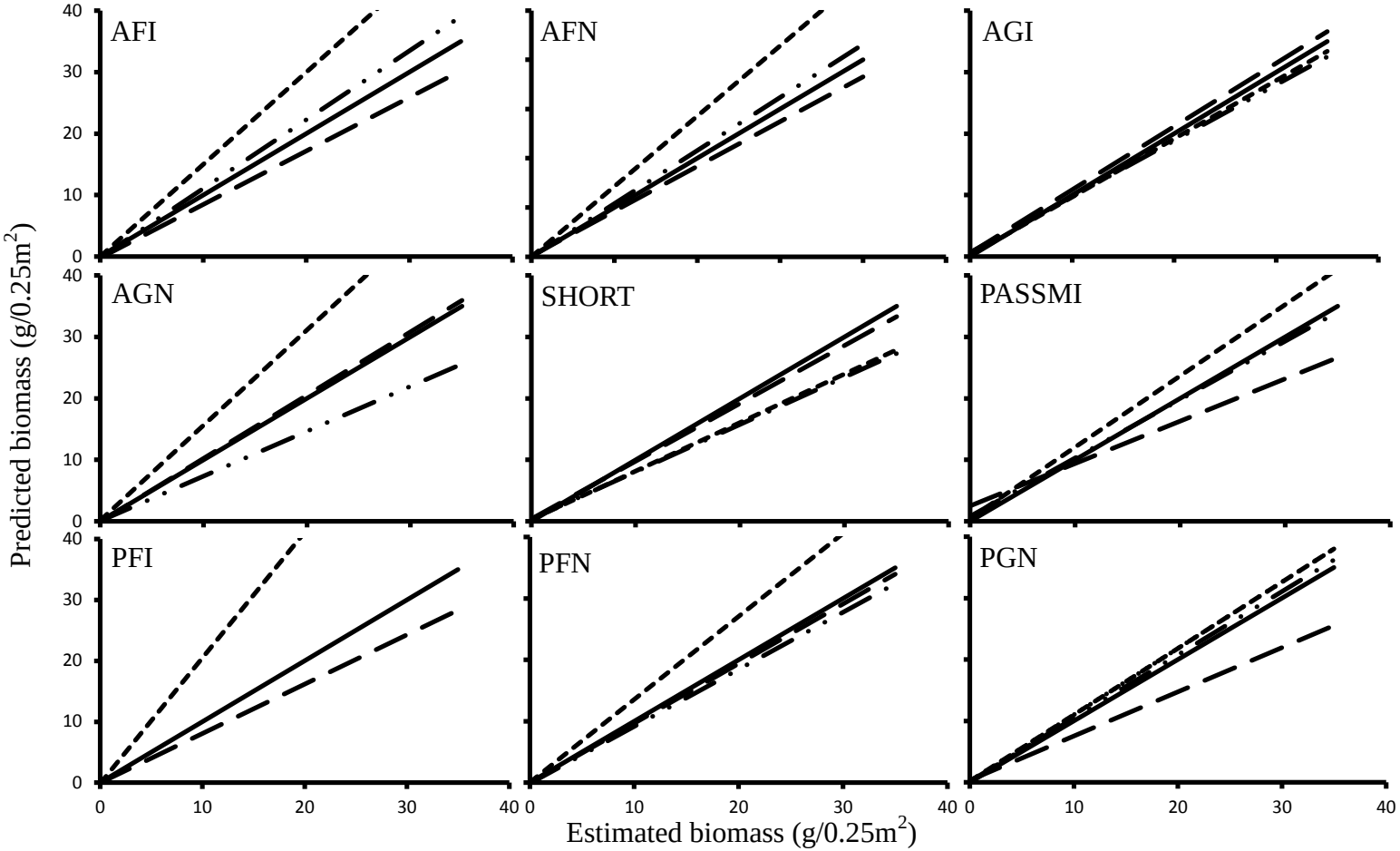

Figure 5. Observer regression lines (Observer 1 = short dash line, Observer 2 = long dash line, and Observer 3 = dotted and dashed line) by functional groups. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = shortgrass mix, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN =perennial forb introduced, PGN = perennial grass native. Solid line is a reference to y = 1x. Predicted biomass for PASSMI by Observer 2 is untransformed data.

tended to overestimate SHORT (regressions under unity), and underestimate AFN (regressions over unity). When individual estimates were collectively added to estimate total biomass, small quantities of biomass tended to be underestimated and large quantities were overestimated with Observer 1 underestimating small biomass to a larger extent and Observers 2 and 3 overestimating to a greater extent (Figure 6). Observer 1 drastically underestimated five functional groups (PFI, AGN, AFI, AFN, PFN). Observer 2 drastically underestimated AFI and moderately overestimated three functional groups (PGN, AFI, PFI). Observer 3 appeared the most consistent and only distinctly overestimated two functional groups (SHORT and AGN; Figure 7). Average slopes were Observer 1 = 1.32 (underestimated an average of 32%), Observer 2 = 0.91 (overestimated an average of 9%) and Observer 3 = 0.94 (overestimated an average of 6%) (Table 2).

Paired t-test between calibrated biomass estimations and actual biomass in the validation set found only the AGI set for Observer 2 to be significantly different ($P = 0.03$) (Table 3). All other sets had no significant differences ($P>0.05$), validating all models except for AGI for Observer 2. When plotting the model building and validation set of data for Observer 2 for AGI, it becomes evident there are two distinct linear relationships based upon time (Figure 8). There was no significant difference ($P>0.05$) between actual biomass and non-calibrated biomass estimates for functional groups with an inadequate number of data points for the creation of a calibration equation (Table 4).

The analyses of covariance (ANCOVA) to test for coincident lines resulted in no significant interaction effect between observer and functional groups (Table 5). This suggests homogeneity of regression slopes, meaning that none of slopes for regressions

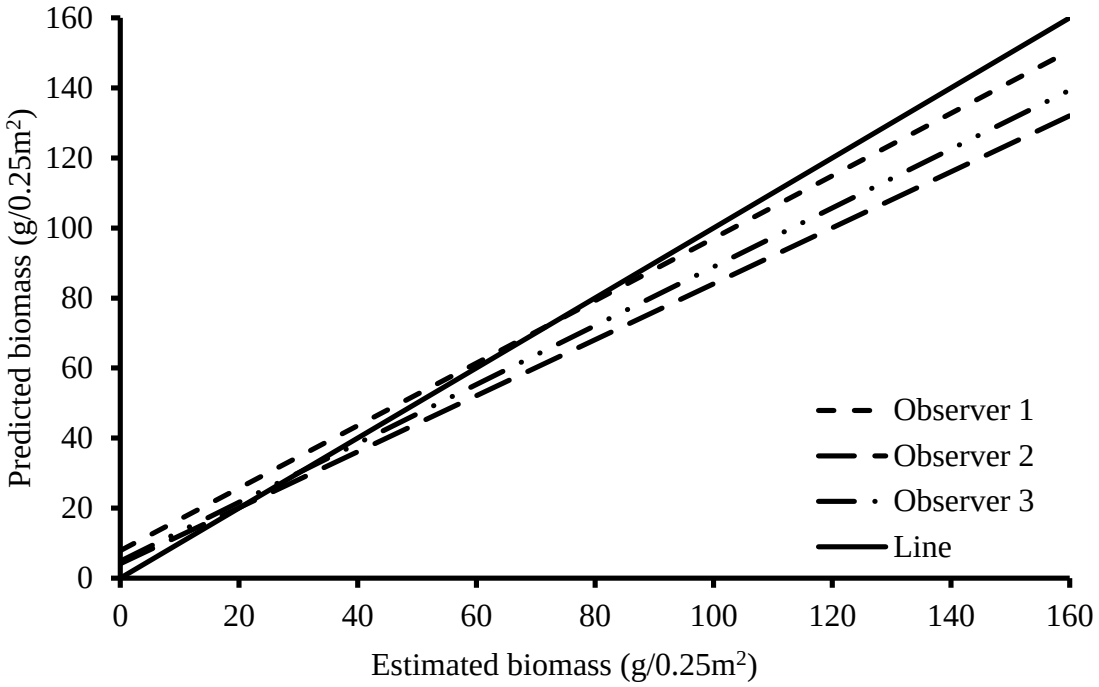

Figure 6. Observer regression lines for total biomass. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Solid line is a calibration of y = 1x.

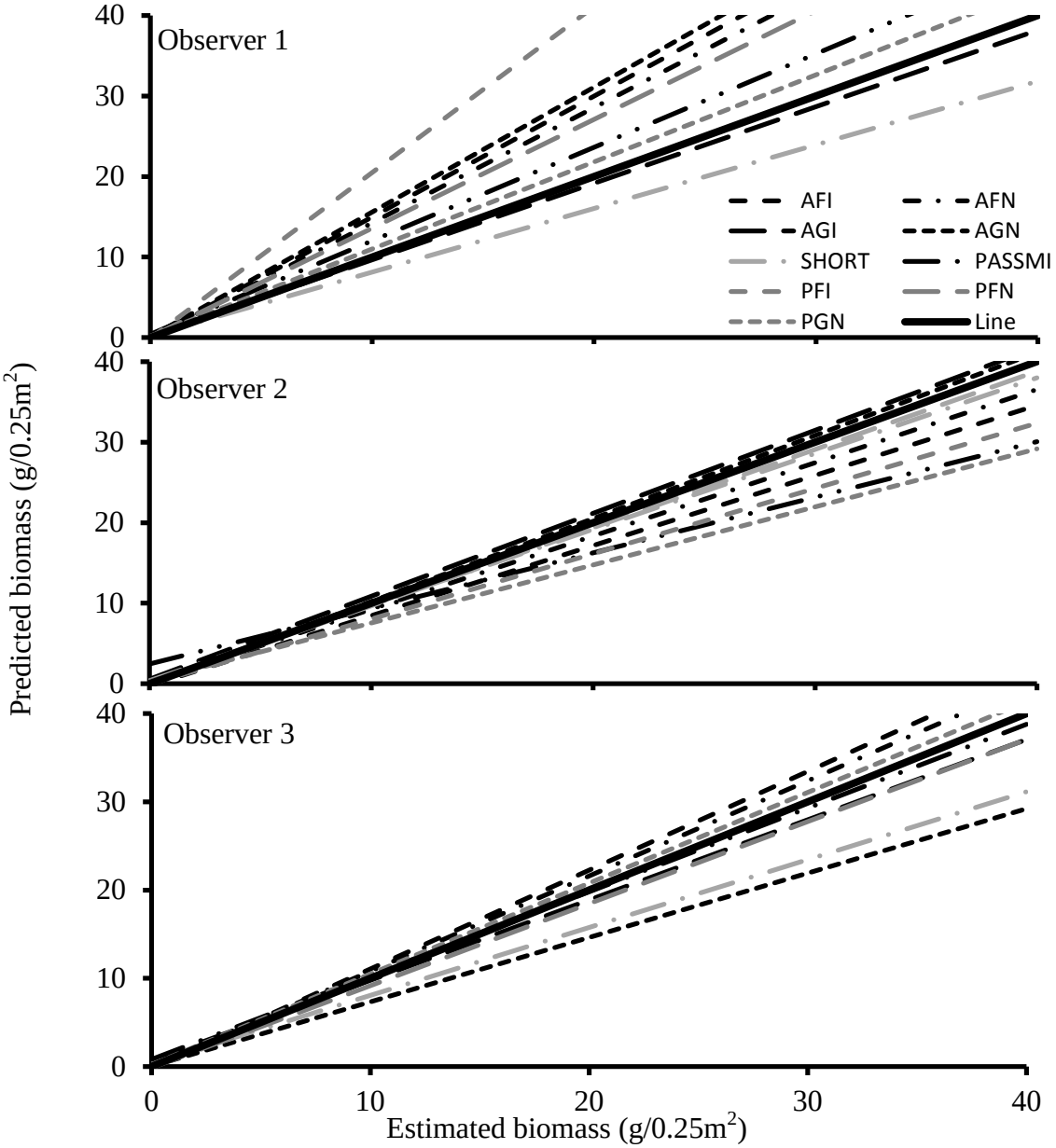

Figure 7. Functional group regression lines by observer. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss and blue grama, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN = perennial forb introduced, PGN = perennial grass native. Solid line is a calibration of y = 1x (unity). Regressions above unity underestimate actual biomass, while regressions below unity overestimate actual biomass.

Table 3. Validation paired t-test between calibrated biomass estimations and actual biomass for functional groups by three observers. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Functional groups: AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss and blue grama, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN = perennial forb introduced, PGN = perennial grass native. Mean = mean difference between calibrated and actual biomass (g/0.25 m^2, *n* = number of transects).

Group	Observer	*n*	Mean	Standard deviation	Minimum detectable difference	P-value
AFI	1	13	0.4	1.8	1.8	0.444
	2	15	0.1	0.5	0.4	0.435
	3	14	0.1	0.1	0.1	0.109
AFN	1	14	-0.1	1.1	1.0	0.723
	2	13	0.4	1.0	1.0	0.141
	3	15	1.4	3.3	3.0	0.129
AGI	1	13	0.4	1.3	1.3	0.337
	2	13	4.9	7.0	6.8	0.026
	3	15	1.6	4.2	3.8	0.167
AGN	1	13	0.3	-0.7	0.7	0.200
	2	14	0.2	0.7	0.7	0.251
	3	14	0.0	0.4	0.3	0.834
SHORT	1	14	-0.1	1.8	1.7	0.793
	2	13	1.6	3.5	3.4	0.128
	3	14	-1.3	2.6	2.4	0.072
PASSMI	1	14	-0.6	2.4	2.3	0.345
	2	13	-0.2	6.7	6.6	0.907
	3	14	1.4	3.2	3.0	0.120
PFI	1	14	0.0	≤0.1	0.0	0.072
	2	13	0.1	0.1	0.1	0.082
PFN	1	14	0.8	1.7	1.5	0.086
	2	13	-0.3	0.7	0.7	0.227
	3	14	0.2	1.3	1.2	0.549
PGN	1	14	0.0	0.3	0.3	0.833
	2	13	0.2	0.5	0.5	0.138
	3	14	0.3	0.6	0.6	0.120
TOTAL	1	14	0.8	7.1	6.6	0.680
	2	13	-11.1	18.9	18.5	0.056
	3	14	-1.6	6.4	6.0	0.364

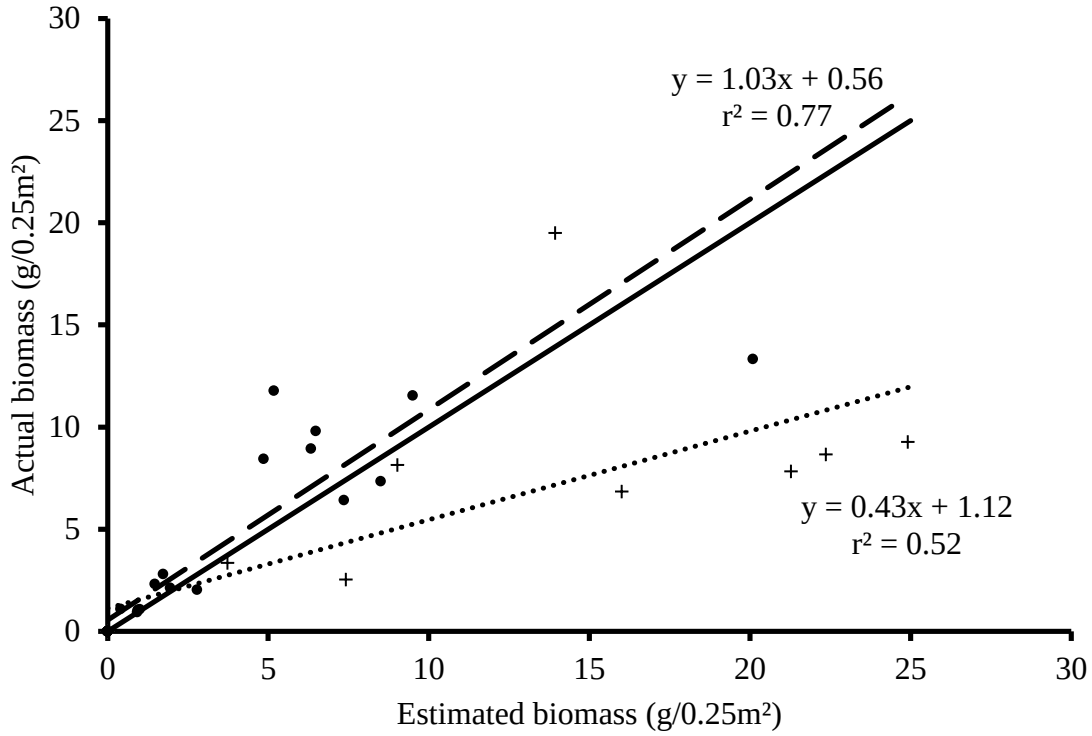

Figure 8. AGI regression for Observer 2. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Two distinct relationships were found for AGI from early to late season. Early season regression (black circles, dashed line) equation and late season regression (black crosses, dotted line) equation are presented together. Solid line is a calibration of y = 1x.

Table 4. Non-regression paired t-test between direct biomass estimations and actual biomass for functional groups that where not found on enough transects to create a regression by three observers. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Functional groups: ARIPUR = purple threeawn, HESCOM = needle and thread, NASVIR = green needlegrass, PFI = perennial forb introduced, PGI = perennial grass introduced. Results in g/0.25m^2, mean = average difference between direct estimates and actual biomass with *n* = number of transects.

Group	Observer	*n*	Mean	Standard deviation	Minimum detectable difference	P-value
ARIPUR	1	40	0.0	0.5	0.3	0.729
	2	39	0.0	0.2	0.1	0.320
	3	41	-0.3	1.5	0.8	0.208
HESCOM	1	40	0.0	0.2	0.1	0.323
	2	39	0.0	0.1	0.1	0.380
	3	41	0.0	0.1	0.0	0.323
NASVIR	1	40	0.0	0.1	0.0	0.216
	2	39	-0.3	2.1	1.1	0.306
	3	41	-0.3	1.6	0.9	0.323
PFI	3	41	0.0	0.2	0.1	0.457
PGI	1	0				
	2	0				
	3	41	0.0	0.1	0.1	0.323

Table 5. Two-way ANCOVA results with a dependent variable of clipped transect biomass (g/0.25m^2), with two independent variables of observer and functional group (group) and estimated biomass as a covariate. Data were collected in 2010 from the Buffalo Gap National Grasslands, SD.

Source	Type III Sum of Squares	Degrees of Freedom	Mean Square	F-value	P-value
Corrected Model	2937.885	27	108.811	101.985	≤0.001
Intercept	34.682	1	34.682	32.507	≤0.001
Estimated	2545.316	1	2545.316	2385.659	≤0.001
Group	25.036	8	3.129	2.933	0.003
Observer	1.762	2	.881	.826	0.439
Group * Observer	16.201	16	1.013	.949	0.513
Error	447.041	419	1.067		
Total	3896.844	447			
Corrected Total	3384.926	446			

were significantly different. After adjusting for the covariate of estimated biomass, there was no significant difference of regression intercepts between observers (Figure 9) while there was a significant difference between functional groups (Figure 10). This indicates that observers tended to estimate functional groups similarly, but significant differences were noticed between functional groups. PASSMI was the only functional group that was significantly different than the other groups and had a mean biomass 44% higher than the next closest group, AGI.

Observer regressions could be compiled by functional groups because there was no significant difference between the slopes and intercepts of observers but a significant difference between 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 (Table 6). Four of the regressions, PASSMI, SHORT, PGN, and Total, had intercepts significantly different than zero. Estimated biomass explained 82% of the variation in actual biomass (range $r^2 = 0.64$ to 0.90) suggesting the regressions have a high predictive value. Collectively, observers tended to estimated biomass close to unity ($y = 1x$) as the average absolute difference between regression slopes and a slope of one was only 0.09 (Figure 11). Slight tendencies for the underestimation of AFI and AFN and an overestimation of SHORT are noticeable. Regressions for Total biomass demonstrated a tendency to underestimate small amounts of biomass and overestimate large amounts of biomass (Figure 12). Paired t-test between calibrated biomass estimations and actual biomass in the validation set found only AGI and AGN to be significantly different (Both $P = 0.01$, Table 7) despite the high precision and accuracy of these calibrations ($r^2 = 0.86$ and 0.78, and slope = 0.96 and 1.04 respectively). All other sets had no significant

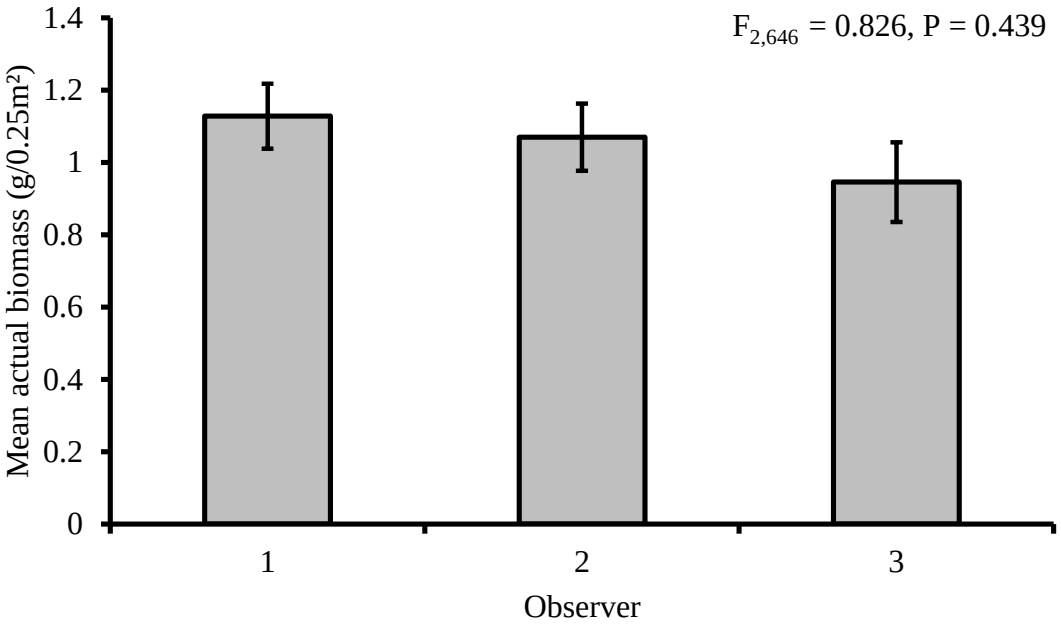

Figure 9. ANCOVA adjusted actual biomass means with standard errors of three observers taken at estimated = 0.76 g/0.25m^2. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD.

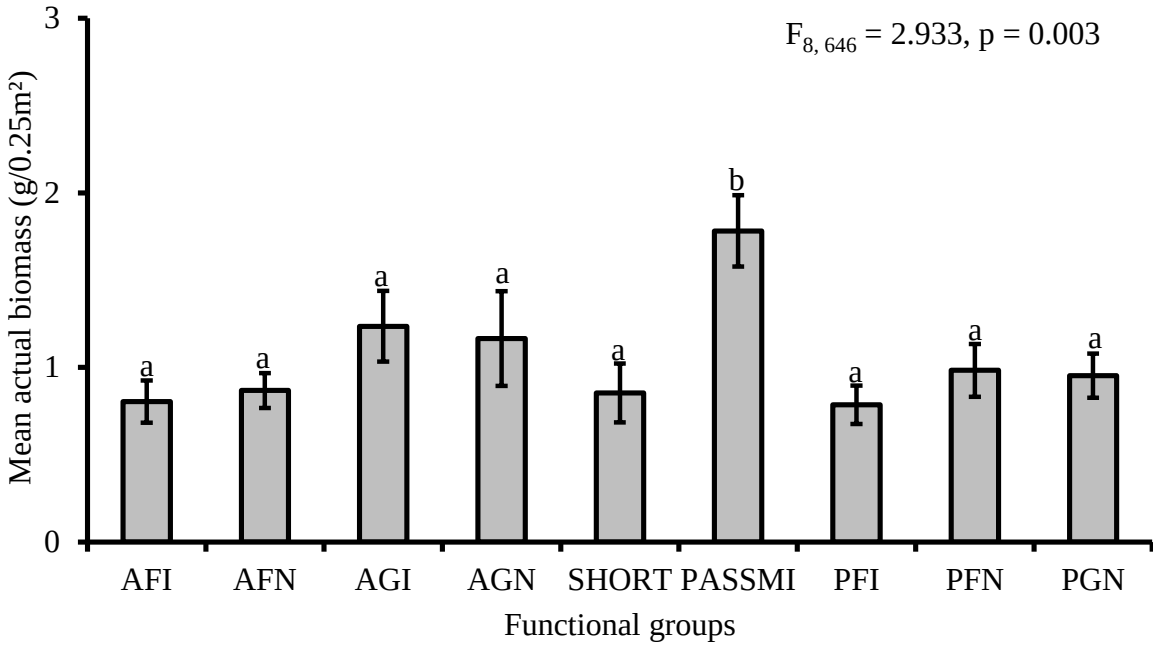

Figure 10. ANCOVA adjusted actual mean biomass with standard errors of functional groups taken at estimated = 0.76 g/0.25m^2. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Functional groups with the same similar letter were not significantly different. Functional groups: AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss and blue grama, PASSMI = western wheat grass, PFI = perennial forb introduced, PFN = perennial forb introduced, PGN = perennial grass native.

Table 6. Regression equations [x = estimated biomass, y = actual biomass; grams/0.25m2] for compiled functional groups across three different observers. Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Functional groups are as follows: AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss and blue grama, PASSMI western wheatgrass, PFI = perennial forb introduced, PFN = perennial forb introduced, PGN= perennial grass native. Total = total biomass. *n* = number of transects. Outliers = number of transects excluded from data analysis. P values for intercept (int) and slope (X), MSE = Mean square error, r^2 = coefficient of determination.

				P-value			
Group	*n*	Outlier	Equation	int	X	MSE	r^2
AFI	75	3	y=0.04+1.12x	0.551	≤0.001	0.43	0.78
AFN	77	1	y=0.09+1.11x	0.270	≤0.001	0.79	0.87
AGI	77	1	y=0.40+0.96x	0.072	≤0.001	1.18	0.86
AGN	74	4	y=0.12+1.04x	0.564	≤0.001	0.69	0.78
SHORT	76	2	y=0.37+0.82x	0.044	≤0.001	1.18	0.90
PASSMI	76	2	y=0.07+0.96x	0.001	≤0.001	2.16	0.86
PFI	74	4	y=0.98+0.92x	0.177	≤0.001	0.13	0.64
PFN	75	3	y=0.11+1.07x	0.405	≤0.001	0.49	0.85
PGN	74	4	y=0.25+0.94x	0.030	≤0.001	0.64	0.75
TOTAL	76	2	y=5.09+0.85x	≤0.001	≤0.001	1.94	0.90

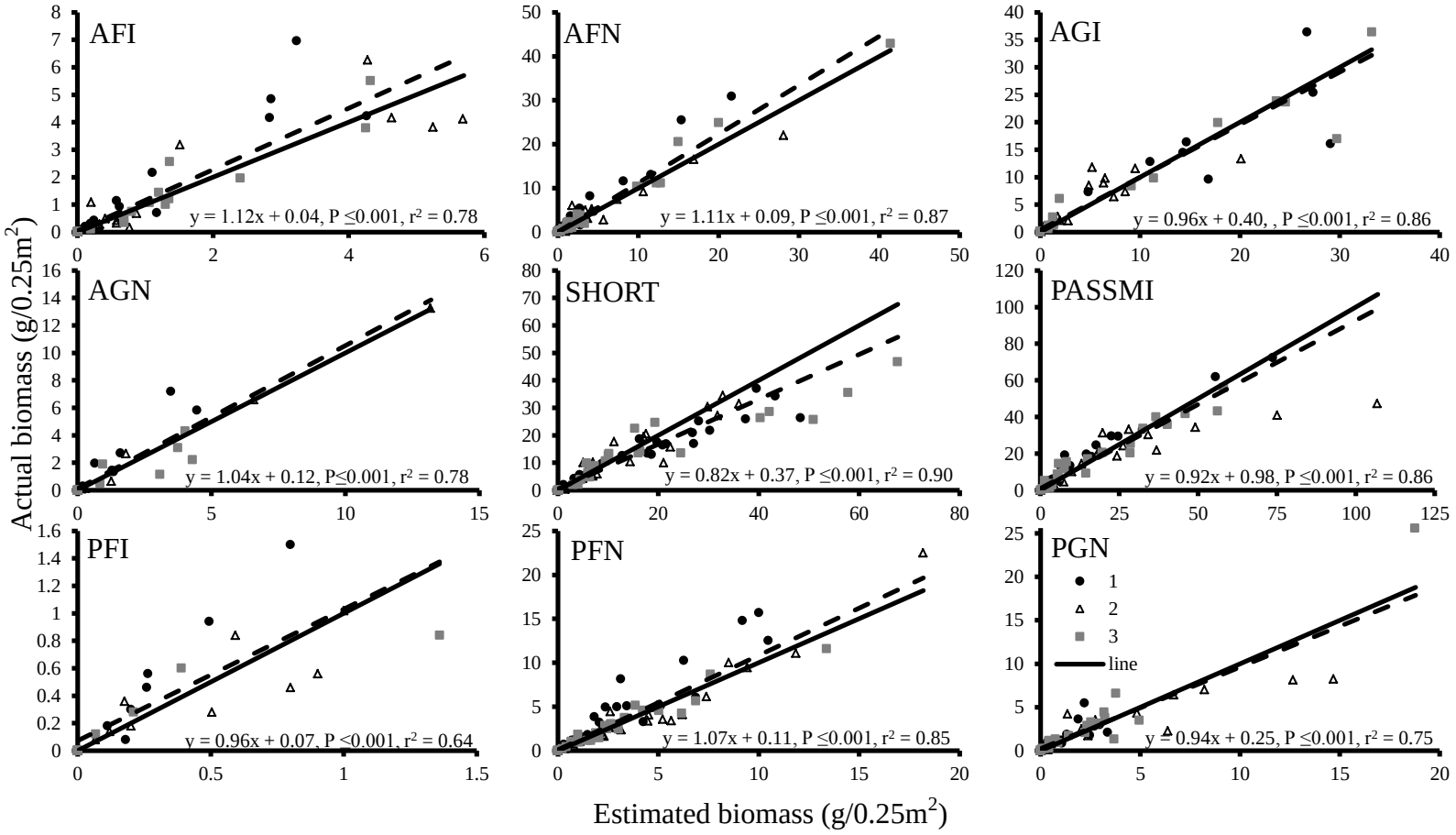

Figure 11. Regression lines by functional groups compiled across observer (dashed line). Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Note different scales of y- and x-axis. Solid line is a reference to y = 1x. AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss blue grama, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN =perennial forb introduced, PGN = perennial grass native. 1 (black circles), 2 (triangles), and 3 (grey squares) represent Observer 1, 2, and 3 respectively.

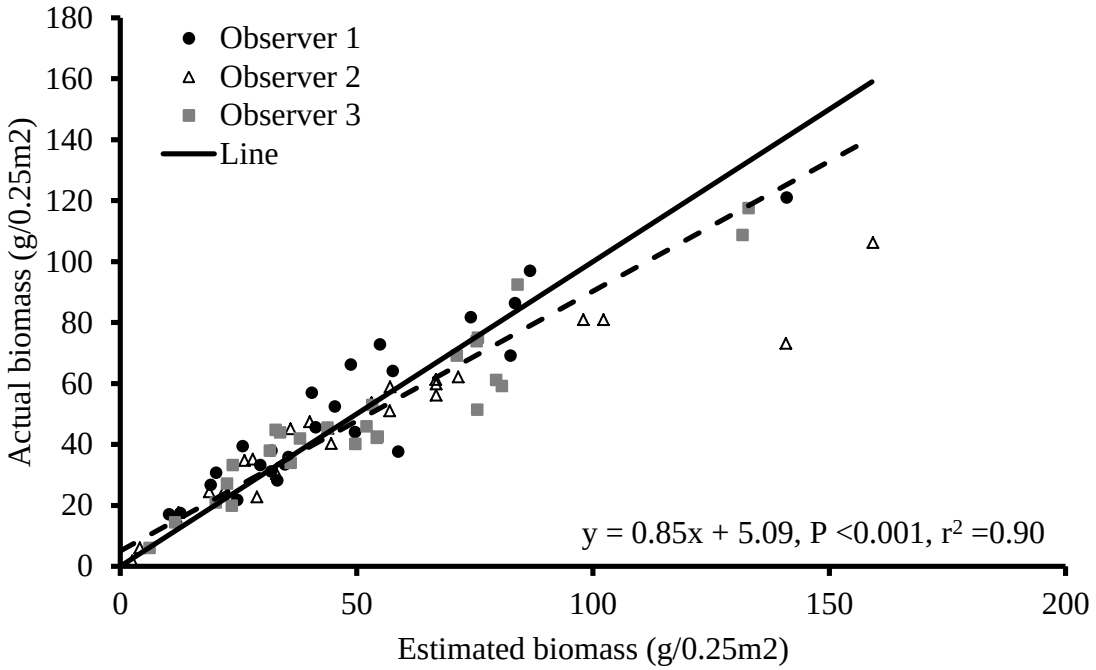

Figure 12. Regression line for total biomass compiled across observer (dashed line). Data were collected in 2010 on the Buffalo Gap National Grasslands, SD. Line is a reference to y = 1x. Observer 1 = black circles, Observer 2 = triangles, and Observer 3 = grey squares.

Table 7. Validation paired t-test between calibrated biomass estimations and actual biomass for compiled functional groups from three different observers. Data were collected in 2010 on the Buffalo Gap National Grassland, SD. Functional groups are as follows: AFI = annual forb introduced, AFN = annual forb native, AGI = annual grass introduced, AGN = annual grass native, SHORT = buffalograss and blue grama, PASSMI = western wheatgrass, PFI = perennial forb introduced, PFN =perennial forb introduced, PGN = perennial grass native. Results in g/0.25m^2, mean = average difference between calibrated and actual biomass with *n*= number of transects. * = significant difference (α = 0.05). CV = SD/mean actual biomass.

Group	*n*	Mean	Standard Deviation	Minimum Detectable Difference	P-value
AFI	42	0.0	0.9	0.4	0.737
AFN	42	1.3	4.7	2.4	0.076
AGI	42	2.0	4.6	2.4	0.008*
AGN	41	0.3	0.7	0.4	0.011*
SHORT	42	-0.2	2.2	1.1	0.499
PASSMI	41	3.0	10.9	5.6	0.088
PFI	42	0.0	0.3	0.1	0.774
PFN	42	0.2	1.3	0.7	0.280
PGN	42	0.2	0.7	0.3	0.096
TOTAL	41	4.2	15.3	8.0	0.089

differences ($P>0.05$). Despite high variability, calibration appeared to closely approximate the mean, with an average mean difference of approximately $0.76 \pm 0.37 g/0.25 m^2$. An additional multivariate cost benefit analysis found that ¼ of all transects sampled should be double sampled (data not shown).

DISCUSSION

Biomass estimation is an important metric in assessing ecological relationships and processes, especially site productivity. Ideally, biomass estimation methods would express relationships between actual and estimated biomass that have high precision (Tadmor et al. 1975) and accuracy (Hutchings and Schmautz 1969), with an intercept passing through or closely 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), remain consistent among observers (Friedel et al. 1988), and optimize sample costs (Wilm et al. 1944). However, in my exploration of the reference unit method, I encountered similar complications as previous researchers that include 1) nonlinear relationships between actual and estimated biomass, 2) influence of observer experience and bias, 3) underestimation of biomass in small quantities and overestimation in large quantities, 4) difficulties estimating infrequent species, 5) difficulties estimating multi-species groups, and 6) problems maintaining consistency in a variety of ecological conditions.

Ahmed et al. (1983) were critical of conclusions made by Tadmor et al. (1975) that calibration regressions could be nonlinear. Despite data presented by Tadmor et al. (1975) that demonstrated an average increase in the coefficient of determination by 0.20 with a nonlinear calibration, Ahmed et al. (1983) maintained that it was illogical for ocular estimation relationships to change (nonlinearity) as plot biomass increased. However, in this study, Observer 2 had a significant curvilinear relationship with a substantial number of points conclusively demonstrating nonlinearity for biomass estimations of PASSMI (Figure 4). The nonlinear relationship may be the result of

inadequate separation of living stems from standing dead while estimating biomass in tall dense stands of PASSMI. Furthermore, lodging exacerbated the difficulty in separating live from dead biomass. 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, and Hutchings and Schmautz (1969), who 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 litter present in grasslands impacted the ability to use the weight-estimate method. While this singular incidence may suggest that the reference unit method can be vulnerable to an inability to fully distinguish living from dead plant materials, especially when they are in large quantities, all other equations proved otherwise. With increased awareness and experience in visually sorting live stems from standing dead in high biomass plots, observers should be able to correct nonlinearity in regressions.

Comparisons of functional group slopes for all three observers revealed that observer calibrations were generally scattered around unity ($y = 1x$), which is a 1:1 relationship of estimated to actual biomass (Figure 5). The slope provides information on the accuracy of estimates and, ideally, would be very close to unity (Hutchings and Schmautz 1969). If a slope is >1, then that functional group biomass tends to be underestimated by observers, while a slope <1 indicates a tendency to overestimate biomass. Slopes of AGI demonstrated accurate estimates of biomass by all observers, as slopes were close to unity. AGI was comprised of only two winter annual species, Japanese brome (*Bromus avensis* L.) and downy brome (*Bromus tectorum* L.), which have wispy stems that support large drooping panicles. Since these species are winter

annuals, during the early sampling season they were more phenologically advanced than most other species. Since the panicle contributed more biomass relative to the stem and the developed panicles were visually prominent in the plot, the observers could make more accurate estimations.

Calibration slopes for some other functional groups, such as PFI, were more variable. Interestingly, the PFI regression for Observer 1 had the largest slope (2.06) of all functional groups. Reese et al. (1980), demonstrated that if calibrations had points (x,y) that were tightly clumped together, they may have a less defined relationship. Actual biomass for PFI was less than 1.5g for all transects for Observer 1. Although this calibration was still fairly precise ($r^2 = 0.91$), slope of the regression may have been closer to unity (y=1x) had there been a greater diversity in the spread of data points (x,y).

All observers tended to overestimate biomass of SHORT (buffalograss and blue grama) (Figure 5). Since buffalograss is stoloniferious, observers had to determine whether to include a stolon while making a SHORT reference unit. Given that stolons have a relatively greater mass density than the leaf blades, reference units that did not contain a stolon were more likely to under-represent biomass. Because a stolon was generally found within the 5% area, observers tended to include a stolon in their 5% area reference unit. If a stolon was included in the reference unit, estimations may have included more stolons than were actually present on a plot, causing an overestimation of biomass. These overestimations may be further compounded by the occasional changes in proportions of buffalograss to blue grama. Furthermore, overestimations of SHORT may have been complicated by short *Carex* spp. that were intermixed with the buffalograss and blue grama. Because short *Carex* spp. were difficult to find, they may have been

included in estimations, but were sorted out into 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 may have underestimated AFI and AFN because of the difficulties of finding all the annual forb species in a given plot when intermixed with other species. In a similar study, Hutchings and Schmautz (1969) had issues using the relative weight method because it was difficult to recognize or see all species (as many as 15) in a plot. In the current study, above-average precipitation (Figure 2) and secondary succession in prairie dog colonies may have increased the height of grasses (Osborn and Allan 1949), which lead to grasses dominating over annual forb species. Consequently, many individuals of AFI and AFN may not have been found or referenced correctly.

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. Despite relatively precise and accurate estimations of individual species and functional groups, observers tended to underestimate small values of biomass and overestimate large values of biomass (Figure 6). 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, data from the current study (Figure 6) suggests that estimated data with the reference unit method may have greater dispersion than actual biomass. These trends are often not noticed on individual species and functional groups. However, the cumulative impacts of these small errors are noticeable when individual estimates are added together to estimate total biomass. Since

sampling duties rotated from transect to transect in the current study, observers often did not harvest the plots they referenced. Pechanec and Pickford (1937), and Wilm et al. (1944) indicated that observations made during the harvest of double sampled plots can aid in reducing personal errors in subsequent estimates. In the future, if observers harvested plots directly after referencing them, they may gain a fuller understanding of how standing dead influences live biomass appearance and which species are commonly hidden under or intermixed with others. When these groups or species are identified, additional efforts by observers can be made to assure more accurate estimations. However, information on the regressions of actual to estimated biomass should not be provided to observers until after sampling is completed to avoid unintended influence on the consistency of personal bias (Shoop and McIlvain 1963). For example, Observer 1 typically underestimated by an average of 32%, and Observers 2 and 3 overestimated by an average of 9% and 6%, respectively (Figure 7 and Table 2). This demonstrates that the reference unit method is subject to inherent personal bias that has been reported by other authors exploring estimation techniques (Hughes 1959; Pechanec and Pickford 1937; Wilm et al. 1944; and others). Despite these obvious biases, observers had high precision (average r^2 for Observer 1 = 0.88, Observer 2 = 0.80, Observer 3 = 0.88). Although estimates may be inaccurate, as long as personal bias of estimates are consistent, they can be corrected with a calibration regression created through double sampling (Reese et al. 1980).

Biologically, calibrations should pass through or closely to the origin (Carpenter and West 1987). If the value of the intercept is too high or negative, estimates of mean biomass may be too large or even negative in value. The implications of this error could

be especially important when estimating pastures where a species or functional group exhibits a lower mean biomass than typically found in the area. Forcing the intercept through the origin can be done to prevent possible negative estimates for low biomass or the return of a positive value for a species that is not present (Ahmed et al. 1983; Cabral and West 1986). However, Ahmed et al. (1983) recommended against forcing the intercept through zero because forced intercepts, in their study, had larger variance estimations and the coefficient of determinations were difficult to interpret. Ideally, it would not be necessary to force calibrations through the origin and the intercepts should not be significantly different than zero (Hutchings and Schmautz 1969). This ideal situation was demonstrated in the current study, where 25 of the 26 regressions had intercepts that were not significantly different from the origin (Table 2).

Determining composition of functional groups and reference units may impact the precision of estimates, as seen by the various coefficients of determination for functional groups, when observer coefficients of determination were averaged. 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 ($r^2 = 0.73$) (Table 2). This particular functional group contained both tall grass species and short *Carex* (grass-like) species. Most of the remaining functional groups had species with less disparity between growth characteristics, which 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. In addition, errors could have been caused by short statured *Carex* spp. that were intermixed with other similar species (e.g. SHORT [as mentioned previously]) or hidden by taller vegetation. In the future, it would

be beneficial to further separate PGN and functional groups with similar issues based upon growth habits (e.g. short *Carex* spp. and taller grass species) or weight densities. Interestingly, the two groups of species that shared the highest average coefficient of determination (0.91), PASSMI and SHORT, were also two of the simplest functional groups. Since PASSMI consisted of a single species and SHORT contained two very similar species, it appears that simple functional groups are easier to estimate using reference units than diverse functional groups. In addition, Kirmse and Norton (1985) found 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.

If observers consistently used the reference unit method, the high precision of the calibration regressions should be reflected by non-significant differences between actual and calibrated estimation biomass for the validation set. In the current study, estimations made throughout an entire season had high predictive value. For example, Observer 1 and 3 had similar mean coefficients of determination (mean $r^2 = 0.88$ [range 0.67 to 0.95] and mean $r^2 = 0.88$ [range 0.81 to 0.95], respectively) while Observer 2 was only slightly lower (mean $r^2 = 0.80$ [range 0.61 to 0.97]) than the other two (Table 2). Furthermore, these coefficients of determination are congruent with previous studies evaluating the reference unit method (Andrew et al. 1979; Andrew et al. 1981; Cabral and West 1986; Carpenter and West 1987; Kirmse and Norton 1985). In addition, all calibration regressions except for AGI, for Observer 2, were validated based upon time, which

implies that the relationships between estimated biomass and actual biomass for most calibrations were consistent from early to late season (Table 3). The lack of validation for AGI is especially interesting because the AGI calibration equation for Observer 2 demonstrated high precision ($r^2 = 0.77$) and accuracy (slope = 1.03). As previously mentioned, AGI was comprised of only two species, Japanese brome and downy brome. These species produce large drooping panicles that shatter and leave an erect stem standing as the summer progresses. If visual estimation of AGI is overly dependent upon the presence of a panicle, then a substantial number of stems may not have been accounted for during estimations as the panicles began to shatter. This issue was recognized before sampling began and observers were warned of this potential source of error. Nonetheless, Observer 2 had two distinct relationships between estimated to actual biomass for early and late season estimations (Figure 8). Unexpectedly, this late season relationship overestimated biomass rather than underestimated biomass. Since the relationship between estimated and actual biomass for AGI did not change as the season progressed for Observers 1 and 3, it is possible that Observer 2 may have begun overcompensating after the warning of the potential to err in estimation. An increased personal understanding of the relative distribution of the weight of seed heads to stems could greatly benefit observers in the estimation of species similar to these annual brome grasses.

Even with precise calibrations, individual plot or transect estimates in previous studies appeared to be inaccurate (Andrew et al. 1981; Carpenter and West 1987; Hutchings and Schmautz 1969). These authors recommended averaging estimates over an area. Although the standard deviation for the mean differences between actual and

calibrated biomass in the current study was rather high, mean differences remained relatively small (AGI for Observer 2 excluded; range -1.3 to 1.6 g/0.25m^2, $\bar{x}$ = 0.27 ± 0.14 g/0.25m^2; Table 3). This supports the previous recommendations that estimated biomass is not accurate at the plot or transect level, but must be averaged over an area to obtain accurate results. Interestingly, mean differences for Total biomass of Observers 1 and 3 were similar to that of functional groups (0.8 g/0.25m^2 and -1.6 g/0.25m^2 respectively). Total biomass for Observer 2 was not significantly different between calibrated and actual biomass (P = 0.06, mean -11.1 g/0.25m^2), however, the mean difference between calibrated and actual was much higher relative to Observers 1 and 3. This difference was probably heavily influenced by the curvilinear relationship of PASSMI for Observer 2 and its relative importance to the plant communities of the area.

Previous research has focused on creating daily regressions for a single reference unit to estimate a species (Andrew et al. 1979; Andrew et al. 1981; Cabral and West 1986; Carpenter and West 1987; Harrington and John 1990; Kirmse and Norton 1985; Noble et al. 2009; Woodland 2004). However, validation based on time for 25 of 26 calibration regressions in the current study suggests that observers can produce consistent estimations through time. Furthermore, creating calibrations from many reference units of different heights and weights throughout the season does not impact the precision and accuracy of estimates. In their study of the dry weight-estimation method, Tadmor et al. (1975) concluded that local differences in species composition, growth forms or habits, and phenological stages can decrease the consistency of estimates made in a given time period. However, the results of the current study demonstrate that the reference unit method maintains consistency and is less vulnerable to spatial differences, possibly

because the reference units reflected the current phenological states. The ability to maintain consistency in estimation through a variety of sampling locations can decrease the sample costs by allowing the pooling of data, which can increase the precision of estimates (Laca et al. 1989), and achieve a degree of stability in coefficients (Reppert et al. 1962). Therefore, reference data can be pooled through a season, rather than the single unit calibrations used in previous research.

While using the dry-weight rank method, Walker (1970) reported difficulties in estimating weight contributions because individually insignificant species had combined weights that were "undoubtedly significant". Mannetje and Haydock (1963), and Hutchings and Schmautz (1969) suggested collectively measuring uncommon or minor species to improve estimates. Despite their reference unit regressions being composed of only a single species, Kirmse and Norton (1985) suggested that there is a potential to apply regression equations to groups of species with similar foliage characteristics. In the current 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 too time consuming, especially given the patchy or uncommon nature of many of the species. Although previous research has not thoroughly indicated the potential for using multiple species calibration equations for reference units, it appears that the reference unit method is capable of accurate and precise collective estimates of multiple species. This appears to be the case for the present study because almost all of the multiple species groups were validated (22 of 23 calibrations). However, Reese et al. (1980) demonstrated difficulties in comparing species that have different weight densities; therefore, I recommend that estimates of

multiple species be based on a reference unit that contains only the species of the functional group that is most dominate within the plot.

Several functional groups were not found in many transects, resulting in an inadequate number of points for the creation and validation of calibration regressions. Within all of these functional groups, non-calibrated estimated biomass was not significantly different than actual biomass (Table 4). Subsequently, these non-calibrated groups would minimally change outcomes during rangeland monitoring. However, if a more accurate estimation of biomass is required, separate plots containing the uncommon species could be placed discriminately to assure adequate points for the creation of a calibration. Alternatively, if subsequent sampling is not necessary, these plants could be harvested in all cases, dried, and weighed to provide an accurate measurement of biomass.

An observer's experience using a method could impact the precision and accuracy of their estimates. Gillen and Smith (1986) noted that the amount of error in an estimate was dependent on the experience of the observers. Several authors suggested intensive training periods before implementing an estimation method (Hughes 1959; Mannetje and Haydock 1963; Pechanec and Pickford 1937; Shoop and McIlvain 1963). However, authors using methods dependent upon relative weight comparisons reported no significant difference between experienced and inexperienced observers (Haydock and Shaw 1975; Hutchings and Schmautz 1969; Hutchinson et al. 1972; Kirmse and Norton 1985; Reese et al. 1980). For this study, observer experience was variable. For instance Observer 1 had significant experience using the reference unit method, while Observer 2 had no experience, and Observer 3 had been briefly introduced to the method. An

analysis of covariance (ANCOVA) was used to test for coincident lines and resulted in no significant difference between observers (Table 5). Collectively, this evidence suggests that individuals may have a greater natural aptitude for relative estimates than direct estimates. However, Tadmor et al. (1975) suggested the ability to estimate using the dry-weight estimation method, a direct estimation technique, was a matter of natural aptitude. Regardless, it seems that individuals can become capable estimators of relative weight in a brief amount of time compared to direct estimation. This is most likely the result of individuals making many more relative comparisons during their everyday life than they would direct estimations.

Consistency of estimates through time is important for an individual observer; however, consistency among observers can be more critical because monitoring assessments are likely to be performed over several years by different observers (Friedel et al. 1988). This is especially important given the probable high turnover rate of most technician positions. While Friedel et al. (1988) reported significant differences among observers using the comparative yield method, the current study suggests that the reference unit method estimations are consistent among observers. Friedel and Bastin (1988) suggested using a single observer for vegetation monitoring programs, if possible, to maintain the most consistent results. However, results from this study suggest that relatively inexperienced observers may be as consistent as a single well experienced observer. It may be more efficient for a researcher to hire several relatively inexperienced observers rather than a single, highly experienced observer, because a greater number of samples could be collected from a broader area and labor costs for inexperienced observers would be comparatively less. A cost benefit analysis found that one out of

every four transects sampled using reference units should be double sampled. Given that it took approximately five times longer to clip a transect than to estimate one, in the time it would take to clip 9 transects, 20 transects could be sampled using the reference unit method. This is a >120% net gain in information compared to direct harvest.

As previously mentioned, pooling data can decrease sample costs and increase the precision of biomass estimates (Laca et al. 1989). In fact, Kirmse and Norton (1985) found no significant difference in the slopes and intercepts among three observers, and subsequently pooled observer estimation data. At first glance, data from this study indicated that observer estimations could be pooled because there was no difference among observers (Figure 9). However, further analysis revealed a significant difference among functional groups; therefore, pooling functional group data is not appropriate (Figure 10).

Most regression slopes for observers, pooled by functional groups, approached unity, with the exception of SHORT, which appeared to be slightly overestimated (slope = 0.85) (Table 6). Despite high precision of regressions with slopes approaching unity, inherent personal bias is noticeable, with specific observer points tending to be biased above, near, or below the regression (Figure 11 and 12). This is fairly evident in AFN, PASSMI, PFI, PFN, and Total where data for Observer 1 tended to be above the regressions and data for Observer 2 below. Despite observer bias, precision of the functional group equations were similar to that of individual observer calibration equations ($\bar{x}$ = 0.82, range 0.64-0.90). With high precision and slopes approaching unity, it appears that overestimations by an observer were balanced, at least partially, by underestimations of another. Haydock and Shaw (1975) reported similar results and

suggested this balance between over- and under-estimation is expected if a method is unbiased. This 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.

For the most part, calibrated estimates approximated actual biomass values for all models but AGI and AGN (Table 7). This implies that the compiled observer relationships between estimated biomass and actual biomass for AGI and AGN changed with time. Seasonal bias from the overestimation of AGI by Observer 2 (Figure 8) heavily contributed to this significant difference. Similar to AGI, as time progressed the annual grasses of AGN were overestimated as they matured and senesced. Additionally, the mean difference between actual and calibrated biomass for PASSMI was high ($3.0g/0.25m^2$). Collectively, this evidence demonstrates that pooled estimates may be vulnerable to compounding errors. Although errors may be relatively unimportant individually, they become an issue when combined. Although two models were not validated, the average mean difference between calibrated and actual values of the validation sets was 0.76 ± 0.37 $g/0.25m^2$ for the nine functional groups.

Although previous research indicated an ability to combine observer data (Kirmse and Norton 1985), observers will inevitably differ in their abilities to estimate, with some consistently over- or underestimating to varying degrees (Friedel et al. 1988). As demonstrated from the current study, changes in estimation relationships due to

phenological changes may be further compounded by individual observer bias in calibrations. Although seven of the nine calibrations were validated, the potential interaction between phenological changes and personal bias may inhibit the practicality of pooling observer data. Because of this, I do not recommend combining regressions across observers, unless there are an inadequate number of samples to create individual observer regressions.

MANAGEMENT IMPLICATIONS

The reference unit method is capable of precise and accurate estimates of biomass for herbaceous forbs and grasses. Although the method was susceptible to personal bias, double sampling can be used to create a calibration equation that corrects this issue, provided personal bias is consistent. 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. Conscientious observers can maintain consistency in estimations throughout a season and within a variety of ecological conditions. Because of this, a single calibration can be made for each functional group using data from estimations made with numerous reference units. Experience using the method has little impact on estimations and training can be completed in one hour. Although estimates were consistent among observers, each observer should create their own calibration regression. Sampling using reference units in mixed-grass prairies of southwestern South Dakota can increase efficiency in sampling by >120% and is highly recommended for large scale monitoring projects.

WORKS CITED

Ackerman, C. J., H. T. Purvis, G. W. Horn, M. E. Payton, L. T. Vermeire, S. I. Paisley, and T. N. Bodine. 1999. Estimation of forage mass of Old World bluestem using a visual obstruction measurement technique. *Animal Science Research Report* P-973:194-199.

Ahmed, J., and C. D. Bonham. 1982. Optimum allocation in multivariate double sampling for biomass estimation. *Journal of Range Management* 35:777-779.

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.

Alexander, C. W., J. T. Sullivan, and D. E. McCloud. 1962. A method for estimating forage yields. *Agronomy Journal* 54:468-469.

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.

Back, H. L. 1968. An evalutation of an electronic instrument for the pasture yield estimation. Part 1. General relationships. *Journal of the British Grassland Society* 23:216-222.

Back, H. L., F. E. Alder, and B. G. Gibbs. 1969. An evaluation of an electronic instrument for pasture yield estimation. 2. use with double sampling for regression estimation. *Journal of the British Grassland Society* 24:168-172.

Baker, B. S., T. V. Eynden, and N. Boggess. 1981. Hay yield determinations of mixed swards using a disk meter. *Agronomy Journal* 73:67-69.

Benkobi, L., D. W. Uresk, G. Schenbeck, and R. M. King. 2000. Protocol for nonitoring standing crop in grasslands using visual obstruction. *Journal of Range Management* 53:627-633.

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

Bransby, D. I., A. G. Matches, and G. F. Krause. 1977. Disk meter for rapid estimation of herbage yield in grazing trials. *Agronomy Journal* 69:193-196.

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 an herbs. *Vegetatio* 72:75-79.

Castle, M. E. 1976. A simple disc instrument for estimating herbage yield. *Journal of the British Grassland Society* 31:37-40.

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

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

Crafts, E. C. 1938. Height-volume distribution in range grasses. *Journal of Forestry* 36:1182-1185.

Currie, P. O., T. O. Hilken, and R. S. White. 1987. Evaluation of a single probe capcitance meter for estiamting herbage yield. *Journal of Range Management* 40:537-541.

Currie, P. O., M. J. Morris, and D. L. Neal. 1973. Uses and capabilities of electronic capcitance instruments for estimating standing herbage. *Journal of the British Grassland Society* 28:155-160.

Daubenmire, R. 1959. Canopy coverage method of vegetation analysis. *Northwest Science* 33:43-64.

Douglas, J. T., and C. E. Crawford. 1994. An evaluation of the drop-disc technique for measurements of herbage production in ryegrass for silage. *Grass and Forage Science* 49:252-255.

Earle, D. F., and A. A. McGowan. 1979. Evaluation and calibration of an automated rising plate meter for estimating dry yield pasture. *Austrialan Journal of Experimental Agriculture and Animal Husbandry* 19:337-343.

Emlen, J. T., Jr. 1956. A method for describing and comparing avian habitats. *Ibis* 98:565-576.

Engel, R. K., L. E. Moser, J. Stubbendieck, and S. R. Lowry. 1987. Yield accumulation, lear area index, and light interception of smooth bromegrass. *Crop Science* 27:316-321.

Evans, R. A., and M. B. Jones. 1958. Plant height times ground cover versus clipped samples for estimating forage production. *Agronomy Journal* 50:504-506.

Fletcher, J. E., and M. E. Robinson. 1956. A capacitance meter for estimating forage weight. *Journal of Range Management* 9:96-97.

Friedel, M. H., and G. N. Bastin. 1988. Photographic standards for estimating yield in arid rangelands. *Australian Rangeland Journal* 10:34-38.

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.

Green, J. O. 1949. Herbage sampling errors and grazing trails. *Journal of the British Grassland Society* 4:11-16.

Harmoney, K. R., K. J. Moore, J. R. George, E. C. Brummer, and J. R. Russell. 1997. Determination of pasture biomass using four indirect methods. *Agronomy Journal* 89:665-672.

Harrington, G. N., and G. G. John. 1990. Herbaceous biomass in *Eucalyptus* savanna woodland after removing tress and/or shurbs. *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.

Higgins, K. F. 1986. A comparison of burn season effects on nesting birds in North Dakota mixed grass prairie. *Prairie Naturalist* 18:219-228.

High Plains Regional Climate Center. 2011. *Historic climate data summaries*. Available at: http://www.hprcc.unl.edu/index.php. Accessed 6 Jun 2011.

Holmes, C. W. 1974. The Massey grass meter. Dairy Farming Annual. Palmerston North, New Zealand: Massey University. p. 26-30.

Hughes, R. H. 1959. The weight-estimate method in herbage production determination. Techniques and methods of measuring understorey vegetation; Tifton, Georgia: USDA Forest Service, South and Southeast Forest Experimental Stations.

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.

Jackson, R. D., and L. K. Paine. 2006. Revisiting Robel's visual obstruction method for estimation of standing crop in grasslands. Forage and Grazinglands: Plant Management Network.

Johnson, J. R., and G. E. Larson. 1999. Grassland plants of South Dakota and the Northern Great Plains. Brookings, SD.

Jones, R. M., and J. N. G. Hargreaves. 1979. Improvements to the dry-weight-rank method for measuring botanical composition. *Grass and Forage Science* 43:181-189.

Jones, R. M., R. L. Sandland, and G. A. Bunch. 1977. Limitations of the electronic capacitance meter in measuring yields of grazed tropical pastures. *Journal of the British Grassland Society* 32:105-113.

Karl, M. G., and R. A. Nicholson. 1987. Evaluation of the forage-disk method in mixed-grass rangelands of Kansas. *Journal of Range Management* 40:467-471.

Kirmse, R. D., and B. E. Norton. 1985. Comparision 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). Laramie, WY: University of Wyoming. 102 p.

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.

Martin, R. C., T. Astatkie, J. M. Cooper, and A. H. Fredeen. 2005. A Comparison of Methods Used to Determine Biomass on Naturalized Swards. *Journal of Agronomy and Crop Science* 191:152-160.

Mazaika, R., and P. R. Krausman. 1991. Use of dry-weight rank multipliers for desert vegetation. *Journal of Range Management* 44:409-411.

Michalk, D. L., and P. K. Herbert. 1977. Assessment of four techniques for estimating yield on dryland pastures. *Agronomy Journal* 69:864-868.

Michell, P., and R. V. Large. 1983. The estimation of herbage mass of perennial ryegrass swards: a compartive evaluation of a rising-plate meter and a single probe capacitance meter calibrated at and above ground level. *Grass and Forage Science* 38:295-299.

Miller-Goodman, M. S., L. E. Moser, S. S. Waller, J. E. Brummer, and P. E. Reece. 1999. Canopy Analysis as a Technique to Characterize Defoliation Intensity on Sandhills Range. *Journal of Range Management* 52:357-362.

Neal, D. L., P. O. Currie, and M. J. Morris. 1976. Sampling herbaceous native vegetation with an electronic capacitance instrument. *Journal of Range Management* 29:74-77.

Neal, D. L., and J. L. Neal. 1973. Uses and Capabilities of electronic capcitance instruments for estimating standing herbage. *Journal of the British Grassland Society* 28.

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.

O'Donovan, M., P. Dillion, M. Rath, and G. Stakelum. 2002. A comparison of four methods of herabage mass estimation. *Irish Journal of Agricultural and Food Research* 41:17-27.

Ogura, S., Y. Nagatomo, and M. Hirata. 2005. Estimation of herbage mass in a bahia grass (*Paspalum notatum*) and a centipede grass (*Eremochloa ophiuroides*)

pasture using a capacitance probe, a sward stick and a rising plate. *Tropical Grasslands* 39:22-30.

Osborn, B., and P. F. Allan. 1949. Vegetation of an abandoned prairie-dog town in tall grass prairie. *Ecology* 30:322-332.

Pasto, J. K., J. R. Allison, and J. B. Washko. 1957. Ground cover and height of sward as a means of estimating pasture production. *Agronomy Journal* 49:407-409.

Pearce, R. B., R. H. Brown, and R. E. Blaser. 1965. Relationships between leaf area index, light interception, and net photosynthesis in orchardgrass. *Crop Science* 5:553-556.

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* 44:140-147.

Rayburn, E. B., and S. B. Rayburn. 1998. A standardized plate meter for estimating pasture mass in on-farm research trials. *Agronomy Journal* 90:238-241.

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.

Reppert, J. N., M. J. Morris, and C. A. Graham. 1962. Estimation of herbage on California annual-type range. *Journal of Range Management* 15:318-323.

Robel, R. J., J. N. Briggs, A. D. Dayton, and L. C. Hulbert. 1970. Relationships between visual obstruction measurements and weight of grassland vegetation. *Journal of Range Management* 23:295-297.

Sanderson, M. A., C. A. Rotz, S. W. Fultz, and E. B. Rayburn. 2001. Estimating forage mass with a comercial capacitance meter, rising plate meter, and pasture ruler. *Agronomy Journal* 93:1281-1286.

Santillan, R. A., W. R. Ocumpaugh, and G. O. Mott. 1979. Estimating forage yield with a disk meter. *Agronomy Journal* 71:71-74.

SAS Institute. 2008. The SAS system for Windows. Release 9.2. Cary, NC: SAS Institute, Inc.

Serrano, J. M., J. O. Peca, M. d. Silva, and S. Shahidian. 2011. Calibration of a capacitance probe for measurement and mapping of dry matter yield in Mediterranean pastures. *Precision Agriculture* 12:860-875.

Sharrow, S. H. 1984. A simple disc meter for measurement of pasture height and forage bulk. *Journal of Range Management* 37:94-95.

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

Snee, R. D. 1977. Validation of regression models: methods and examples. *Technometrics* 19:415-428.

Stockdale, C. R., and K. B. Kelly. 1984. A comparison of a rising-plate meter and an electronic capacitance meter for estimating the yield of pastures grazed by dairy cows. *Grass and Forage Science* 39:391-394.

Stoltenberg, M. B. 2004. Effects of prairie dogs on plant community composition and vegetation disappearance in mixed-grass prairie. Brookings, SD: South Dakota State University. 72 p.

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.

Terry, W. S., D. H. Hunter, and B. F. Swindel. 1981. Herbage Capacitance Meter: An Evaluation of Its Accuracy in Florida Rangelands. *Journal of Range Management* 34:240-241.

Trott, J. O., K. J. Moore, V. L. Lechtenberg, and K. D. Johnson. 1988. Light penetration through tall fescue in relation to canopy biomass. *Journal of Production Agriculture* 1:137-140.

Tucker, C. J. 1980. A critical review of remote sensing and other methods for non-destructive estimatino of standing crop biomass. *Grass and Forage Science* 35:177-182.

Uresk, D. W., and T. A. Benzon. 2007. Monitoring with a modified Robel pole on meadows in the central Black Hills of South Dakota. *Western North American Naturalist* 67:46-50.

Uresk, D. W., and T. M. Juntti. 2008. Monitoring Idaho fescue grasslands in the Big Horn Mountains, Wyoming, with a modified Robel pole. *Western North American Naturalist* 68:1-7.

Uresk, D. W., D. E. Mergen, and T. A. Benzon. 2009. Monitoring meadows with a modified Robel pole in the northern Black Hills, South Dakota. *The Prairie Naturalist* 41:121-125.

USDA NRCS. 2010. *Web soil survey*. Available at: http://websoilsurvey.nrcs.usda.gov/app. Accessed April 23 2010.

USDA U.S. Forest Service. 2006. Buffalo Gap National Grassland: America's great outdoors.

Vartha, E. W., and A. G. Matches. 1977. Use of a weighted-disk measure as an aid in sampling the herbage yield on tall fescue pastures grazed by cattle. *Agronomy Journal* 69:888-890.

Vermeire, L. T., A. C. Ganguli, and R. L. Gillen. 2002. A robust model for estimating standing crop across vegetation types. *Journal of Range Management* 55:494-497.

Vermeire, L. T., and R. L. Gillen. 2001. Estimating herbage standing crop with visual obstruction in tallgrass prairie. *Journal of Range Management* 54:57-60.

Vickery, P. J., J. L. Bennett, and G. R. Nicol. 1980. An improved electronic capcitance meter for estimating herbage mass. *Grass and Forage Science* 35:247-252.

Volesky, J. D., W. H. Schacht, and P. E. Reece. 1999. Leaf area, visual obstruction, and standing crop relationships on Sandhills rangeland. *Journal of Range Management* 52:494-499.

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

Welles, J. M., and J. M. Norman. 1991. Instrument for indirect measurement of canopy architecture. *Agronomy Journal* 83:818-825.

Whitbeck, M., and J. B. Grace. 2006. Evaluation of non-destructive methods for estimating biomass in marshes of the upper Texas, USA coast. *Wetlands* 26:278-282.

Williamson, S. C., J. K. Detling, J. L. Dodd, and M. I. Dyer. 1987. Nondestructive estimation of shortgrass aerial biomass. *Journal of Range Management* 40:254-256.

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.

Woodland, R. D. 2004. Influence of fall grazing by sheep on plant productivity, shrub age class structure and herbaceous species diversity in Sagebrush Steppe. Logan, Utah: Utah State University. 146 p.

Zar, J. H. 2010. Biostatistical Analysis. Upper Saddle River, New Jersey: Pearson. 944 p.