



Spatial heterogeneity in Chihuahuan Desert vegetation: implications for sampling methods in semi-arid ecosystems

Laura F. Huenneke^{*}, Dennis Clason[†] & Esteban Muldavin[‡]

^{*}*Department of Biology, New Mexico State University, Las Cruces, NM 88003, U.S.A.*

[†]*University Statistics Center, New Mexico State University, Las Cruces, NM 88003, U.S.A.*

[‡]*New Mexico Natural Heritage Program, University of New Mexico, Albuquerque, NM 87131, U.S.A.*

(Received 25 February 2000, accepted 30 June 2000, published electronically 26 January 2001)

Patchiness of above-ground vegetation, such as that in semi-arid grasslands and shrublands, can pose problems in sampling plant cover, biomass and productivity. We present a method of measuring above-ground plant biomass and production that can be applied consistently among vegetation types and that generates seasonal, spatially-explicit results. Results from 15 sites within the Jornada Basin (Chihuahuan Desert, New Mexico, U.S.A.) confirm considerable patchiness and non-normal distributions of plant biomass, even in grasslands. However, tests of adequacy of sample size and of sample error associated with the regression-based estimates of biomass confirmed that the estimates of above-ground net primary productivity are sufficiently precise to be useful in comparisons of both shrub-dominated and grassland sites.

© 2001 Academic Press

Keywords: net primary productivity; semi-arid ecosystems; Chihuahuan Desert; spatial heterogeneity; *Prosopis glandulosa*; *Bouteloua eriopoda*; *Larrea tridentata*; Jornada LTER program

Introduction

Semi-arid ecosystems are characterized by low plant biomass and vegetative cover, relative to more mesic systems. Moreover, most arid and semi-arid ecosystems are highly variable in both spatial and temporal dimensions. Long-lived perennials (woody plants and succulents) may form patches of high biomass, interspersed by nearly bare areas. During occasional wet periods, average biomass values may increase due to the lush growth of annuals and other short-lived or opportunistic plants. This extreme variability poses challenges for any attempt to quantify plant cover, biomass or productivity. Both sparseness (low mean values) and patchiness of semi-arid vegetation demand extensive and careful sampling to determine mean values and to compare means and variability between sites or over time. However, few published reports describe the

^{*}Corresponding author. Fax: + 505 646 5665. E-mail: LHuennek@nmsu.edu

patchiness of these systems in quantitative terms, or document the adequacy of sampling regimen or sample size relative to the heterogeneity of the systems under study. Hence comparisons between published values are difficult to interpret with confidence.

The Jornada Basin Long-Term Ecological Research (LTER) program is investigating spatial and temporal patterns of above-ground net primary production (ANPP) in ecosystems of varying structure, from semi-desert grasslands to shrub-dominated systems, as part of an integrated assessment of the causes and consequences of desertification. In the south-western U.S.A., as in several other semi-arid regions of the world, the desertification of perennial grasslands concerns land managers. Of particular importance is understanding how these structural changes in ecosystems have altered the magnitude of productivity, the seasonal timing of plant production and the spatial distribution of production. One hypothesis is that redistribution of soil resources, exacerbated by the presence of long-lived woody plants, has resulted in the accumulation of resources under shrubs and the loss of resources from inter-shrub openings—the ‘resource redistribution’ hypothesis (Schlesinger *et al.*, 1990; Reynolds *et al.*, 1999). This hypothesis has generated interest in the spatial patterns of plant biomass and productivity.

Primary productivity, the fixation of carbon dioxide into organic molecules by photosynthetic organisms, is a key feature of ecosystem function. However, net primary production (NPP) has proven a difficult process to quantify and understand. First, some portion of NPP is realized as growth of below-ground plant parts, and in terrestrial communities this portion can be difficult to quantify. Second, even investigations of above-ground NPP (ANPP) have been problematic. Conventional approaches in grasslands have differed from those in forests, woodlands and shrublands. Some methods introduce substantial multiplicative errors, leading to low confidence in the measurements; other approaches involve clipping or other manipulations, that in themselves can alter plant growth and allocation patterns. Most methods cannot provide repeated estimates of productivity for specific plots of ground. This has limited our ability to study spatial variation, which is critical to our understanding of temporal variation in production patterns.

We sought to develop a method for measuring ANPP that can be applied consistently over space and time, in vegetation of different types, and that can quantify the relative heterogeneity of different ecosystems. Our criteria for such a method included: field and lab methods efficient enough to allow effective sampling over a large number of field sites; field and lab measures straightforward enough that trained crews can apply them consistently, even with turnover of personnel over time; methods equally effective or sensitive in ecosystems of very different structure; both spatial and temporal patterns detectable with some resolution; sampling non-destructive, to permit long-term monitoring in a particular study site; powerful enough to detect differences of ecological significance (e.g. 10% or greater).

This paper describes such a method, which we have used successfully in long-term studies in the northern Chihuahuan Desert, New Mexico, U.S.A. We present here a description of our method and an assessment of its strengths and weaknesses compared with conventional methods of measuring ANPP. We use the method to describe the spatial patchiness of grasslands and shrub-dominated ecosystems at the Jornada Basin LTER site, and then discuss the implications for estimating and comparing plant community properties between sites or over time. This approach provides the basis for the Jornada Basin LTER program’s ANPP monitoring.

Materials and methods

Study site

The Jornada Basin LTER program carries out research on the Jornada Experimental Range, administered by the U.S. Department of Agriculture’s Agricultural Research

Service (ARS), and on the Chihuahuan Desert Rangeland Research Center (formerly the College Ranch), administered by New Mexico State University. The sites are located 37 km north-east of Las Cruces, NM, U.S.A. Together they encompass about 104,000 ha devoted entirely to research. The Experimental Range was established in 1915, and the College Ranch in 1927; research efforts since then have been carried out by university staff, ARS scientists, other federal agencies, the International Biological Programme (IBP), and, most recently, the U.S. National Science Foundation-funded LTER program.

The Jornada Basin of southern New Mexico has been the site of a dramatic and well-documented conversion of vegetation from semi-desert grassland (dominated by *Bouteloua eriopoda*, black grama) to shrublands of creosote bush, mesquite, and tarbush. Long-term data sets from the Jornada Experimental Range document this change in composition and structure (Buffington & Herbel, 1965; Gibbens & Beck, 1988), with increasing numbers and sizes of shrubs accompanying a decrease in the abundance and cover of perennial grasses. The flora of the Jornada Basin is well-known. The Jornada LTER program maintains a plant species checklist (available via the World Wide Web at <http://jornada.nmsu.edu/>), with voucher specimens in the New Mexico State University Herbarium. Plant nomenclature follows that of Correll and Johnston (1970), except for grasses which are named according to Allred (1993). Those few non-grass taxa not found in Texas are named according to Kearney & Peebles (1964).

Fifteen sites were selected for study, three in each of five ecosystem types: *Bouteloua*-dominated grasslands; *Larrea tridentata* or creosote bush shrub stands on bajada slopes; *Prosopis glandulosa* or mesquite shrublands; *Flourensia cernua* or tarbush shrub areas on alluvial flats; and grass-dominated playas or ephemeral lakebeds. Livestock grazing has been excluded from all sites since at least 1989.

Sampling design and methods

The three sites of each ecosystem type were explicitly selected to encompass the range of natural variation within ecosystem types. For example, very small, medium, and large playas are each represented in the playa type; differences in size are undoubtedly associated with differences in frequency of flooding, nutrient status, etc. Shrub-dominated sites (creosote, mesquite and tarbush) range from those with short or small shrubs to those with large and well-developed shrub canopies. Because sampling effort at each site is so intensive, there were real limitations on the number of sites of each type that could be studied. We considered three sites per type to be the maximum number feasible, but $n = 3$ is too few for random selection to adequately represent the range of site conditions present. Our intention in the selection of sites was to represent as large a range of variation as possible within the southern portion of the Jornada Basin, and to make any statistical test of differences among vegetation types as rigorous as possible.

Within each site, we established a systematic grid of permanent quadrats. The 1×1 m quadrats are located on a grid, with 10 m between the equivalent corners of quadrats. The grids contain 49 quadrats laid out in a square pattern of 7×7 (Fig. 1). The one exception is the College Playa, where the small area available for experimental work forced us to arrange quadrats in a 3×16 pattern, yielding only 48 quadrats. Buffer strips lie between the outermost quadrats and any nearby fences or other studies; in the smallest study sites these buffer zones are a minimum of 5-m wide, but in most sites such buffers are several tens of meters in width. Quadrats are marked by large aluminum nails with washers, two nails marking the diagonal corners of each quadrat.

Plant biomass within these quadrats is estimated by making non-destructive measurements of the dimensions (cover and height) of individual plants or plant parts.

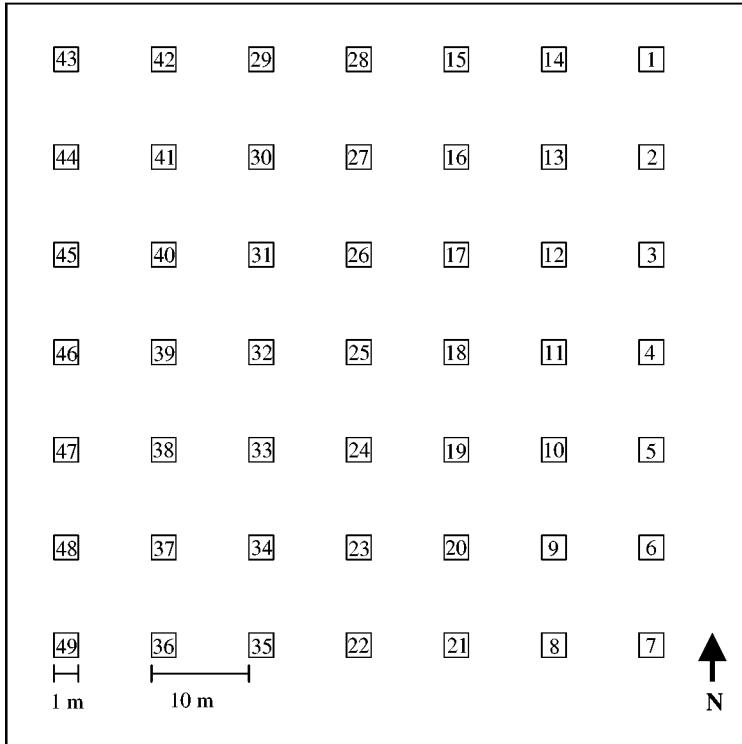


Figure 1. Diagram of arrangement of grid of permanent 1 m² quadrats at each site. Quadrats (not shown to scale) are located on a 10 m × 10 m grid.

A portable square frame with internal area 1 m² is mounted with hose clamps on four tubular PVC legs, which fit over the nails marking the corners of the quadrat. The interior of the frame is gridded with twine into 100 10 cm × 10 cm sections (each constituting 1% of the quadrat's area). The frame slides up or down the corner legs to allow placement as close as possible to the plant canopy; the frame is kept horizontal even if the ground surface is uneven or sloped. Cover or projected surface area for a plant or a plant part is estimated by counting the grid squares or portions of squares occupied by that plant or part. Very small plants or plant parts are measured with a hand-held clear acrylic plate on which are marked squares of 0.5, 0.1, 0.05 and 0.01% area (that is, one-half, one-tenth, one-twentieth, and one-one-hundredth of the area of a single 10 cm × 10 cm grid square). Vertical height is measured with a hand-held tape measure, to the nearest cm. When a plant part extends into the quadrat from outside, or grows out of the quadrat from the inside, only the portion of the plant physically encompassed in the rectangular volume above the quadrat is measured.

During the first three years of work, reference specimens of each species encountered in the quadrats of a site were harvested from plants adjacent to that site. Plants or plant parts were measured in the same way as in the permanent quadrats; samples were taken of plants or plant parts encompassing the range of sizes encountered in the quadrats. Sample size was ten for most species, but 15 or 20 for dominant shrubs and perennial grasses. Sparse or infrequently-encountered species were harvested as available; sample sizes for these were often quite small. Harvested material was returned to the laboratory, where dead material was removed and living biomass (or in the case of plants where portions may have senesced, the current season's growth) was dried at 60°C.

Sampling and harvests take place three times per year. Winter sampling occurs in February, at a time when most species are dormant but some winter annuals have germinated and are present as rosettes. Spring sampling occurs in May, when most shrubs have leafed out and winter and spring annuals have reached full size and reproduction has commenced. Fall sampling takes place in September–October, when shrubs, perennial grasses, and summer annuals have reached peak biomass, and reproductive activity associated with late-summer rains has begun.

Analysis and interpretation

Harvest data were used to construct regressions of plant dimensions (plant volume) *vs.* live biomass. We evaluated several different ways of measuring plant dimension (e.g. diameter for small rosettes), but regressions were most satisfactory using the measurements of cover and height as described above. Consistent field methods also minimized within- and between-observer variation relative to special protocols for different species. Where plants were large or plant volumes diffuse (as for *Bouteloua eriopoda*, a stoloniferous grass that forms large patches) or irregular (as for the many branch systems of large shrubs), greater precision of cover estimates and tighter regressions were obtained by treating the plant as a series of small parts, each measured individually. Harvests were made by site and by sampling date, and initial regressions were calculated separately. After 4 years (1989–1992), we calculated composite regressions using the total data set for each species.

Volume (cm³) for a plant or plant part (observation) is calculated as the product of cover (where cover = 100 cm² multiplied by the % cover measured), and height in cm. Volume thus is not actual plant volume but instead the volume of a rectangular solid of the same linear dimensions as the plant or plant part, used as an index. Linear regressions of live biomass (g) against volume were calculated for each species; each regression was tested for significant improvement due to separate consideration of sites or of harvest dates (using PROC GLM of SAS System for Windows; SAS Institute Inc., 1989–1996). The true relationship between plant volume and biomass is probably not linear, but this regression approach proved robust enough to estimate plant biomass adequately, and simple enough to allow application of the method to the diverse species and structures of Jornada vegetation.

Regressions were calculated for 154 species (plus pooled small forb seedlings, treated as a single group for each vegetation type). Of these, only two species demonstrated significant differences among individual harvest dates (*Pleuraphis* [= *Hilaria*] *mutica* and *Panicum obtusum*, both playa grasses; year-to-year differences in playa flooding probably shape inter-annual variability in greening of the perennial grasses). Regressions for ten other species demonstrated significant differences among seasons, when harvests for different years were combined. These include some of the important dominants of the area (e.g. *Larrea tridentata*, *Bouteloua eriopoda*, and *Muhlenbergia porteri*). Significant differences among sites were encountered for 13 other species, requiring the use of separate regressions for each site in which that species was encountered; *Prosopis glandulosa* or mesquite was among these. All other species were represented adequately by a single regression covering all seasons and sites (that is, no improvement in r^2 noted when season and/or site was included in the analysis). In all, 240 regressions were used to calculate biomass for plant species encountered on the study sites.

Error is minimized using quality control procedures at each step: field measurements and sampling, data recording, sorting and analysis. First, field sampling is performed by teams familiar with the local flora and with measurement protocols. Group calibration of measurements and review of methods are integrated into the field-sampling program. All lab-sorting of live from dead biomass is done by a single

individual to maintain consistency in the measurements from which the biomass regressions are derived.

For each quadrat, biomass is estimated from the appropriate regression for each observation (plant or plant part) for each species; then the estimated biomass values are summed by species to create an output data set containing the calculated biomass for each species in each quadrat. Biomass is summed across all species to obtain total biomass per m² quadrat. ANPP is measured as the positive increment of biomass for a species in a quadrat, summed for all species in a quadrat, over a time interval. This is certainly an underestimate, as plant parts might be produced and lost (e.g. to herbivory or senescence) within a time interval. However, sample dates are timed to minimize senescence or failure to detect short-lived plants; dates of sampling are sometimes shifted in response to unusual weather conditions that hasten or retard the customary phenological pattern.

Results of this sampling regimen are quantitative estimates of biomass and ANPP of all species in the plant community for each quadrat. After summing across all species in a quadrat, standard statistical techniques can be used to describe means, ranges, and other measures of central tendency and of variability for the quadrats within a site. Alternatively, the spatially-explicit values for each grid can be used in geostatistical analyses to model the distribution of biomass or of productivity across the study area. In this paper we present descriptions of the variability associated with our regression estimates. We also document the distribution of biomass and of production values for individual grassland and shrub-dominated sites in both spring and late-summer growing seasons. Finally, we document the adequacy of sampling to allow characterization and comparison of biomass and production among sites and among ecosystem types.

Results

Regressions of biomass vs. volume for individual species

The general approach of calculating a rectangular index of plant volume and using a linear regression between live mass and volume appears to be satisfactory for most species. Harvests over the first 4 years of data collection resulted in sufficient sample sizes for most species to perform a rigorous test of whether regressions were significantly improved by considering site or season. The linear regressions for most species had associated *p*-values of 0.001 or less, with *r*² values of 0.6–0.99 (Fig. 2). Regressions for dominant species (those shrubs, grasses, and forbs contributing substantial biomass on any site) generally yielded *r*²-values greater than 0.75. Those species with poor fits (low values of *r*²) were annual or perennial forbs of low biomass (low slopes) and therefore of relatively little importance at the community level. Regressions for the important perennial grasses had excellent fits for the growing season (spring and fall) but the poorest fits for winter seasons, when the amount of above-ground material remaining green was small and highly variable.

Visual inspection of residuals for individual regressions revealed that fits for larger plant volumes (greater biomass values) were unbiased. Allowing the regressions to have non-zero intercepts facilitated good fits for the larger volumes, but sometimes resulted in poor fits for the smallest volumes (negative intercepts); in the analysis this problem is handled by substituting a very small positive (trace) biomass value for any negative value calculated from the regression. We were more willing to accept errors in the estimation of very small biomass values, since these constitute such a small proportion of the total for the community, than of the larger values, especially for dominant species. Figure 3 presents examples of regressions for two dominant plant species, *Larrea tridentata* (fall season biomass, *n* = 144, *r*² = 0.9168) and *Bouteloua eriopoda* (fall season biomass, *n* = 198, *r*² = 0.8134).

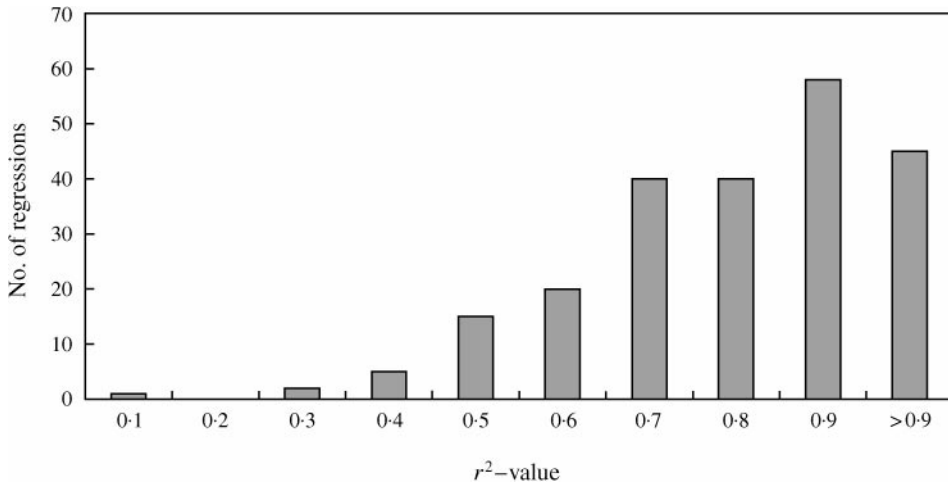


Figure 2. Frequency distribution of r^2 -values for the 240 regressions of live biomass (dry mass, g) against plant volume index.

Distribution of values of biomass and of productivity

We plotted the frequency distribution of values of total biomass for individual 1-m² quadrats, for spring 1991 and fall 1991 (Fig. 4). Strong differences in the spatial distribution of biomass were apparent between shrub-dominated ecosystems and grasslands. The modal (most frequent) values of aboveground biomass in shrub-dominated systems were the lowest values ($0 < 25.0 \text{ g m}^{-2}$), but creosote, mesquite, and tarbush communities also include many quadrats with very high biomass. In contrast, *Bouteloua*-dominated grasslands have modal values in the range of 50–200 g m^{-2} , with most quadrat values clustered near the mode (and mean). However, even grasslands include a small number of patches of extremely high biomass (generally clumps of large succulents such as *Yucca elata* or *Opuntia phaeocantha*); hence patchiness is an important characteristic of Chihuahuan Desert grasslands as well as shrublands. Overall, biomass is clearly not distributed homogeneously across the landscape; nor are biomass values normally distributed in a statistical sense.

Variability at different spatial scales

The extreme patchiness shown within individual sites far exceeds the range of variation between sites. Calculations of standard deviation or of coefficients of variation are misleading when individual values are not normally distributed; hence we chose not to use these as measures of variability in comparing biomass and production between ecosystem types. Parameters such as range or (better) interquartile range (the spread among the central 50% of data values) are more robust indicators of variability when data are distributed non-normally, as is the case for these patchy semi-arid systems. We used simple ranges of values to compare variation within sites, between sites within an ecosystem type, and across the five ecosystem types in the Jornada Basin. The range of biomass values within a single site was clearly much larger than the range of mean biomass between sites within an ecosystem type, or even the range of mean biomass across the five ecosystem types (e.g. fall 1991 values in Table 1).

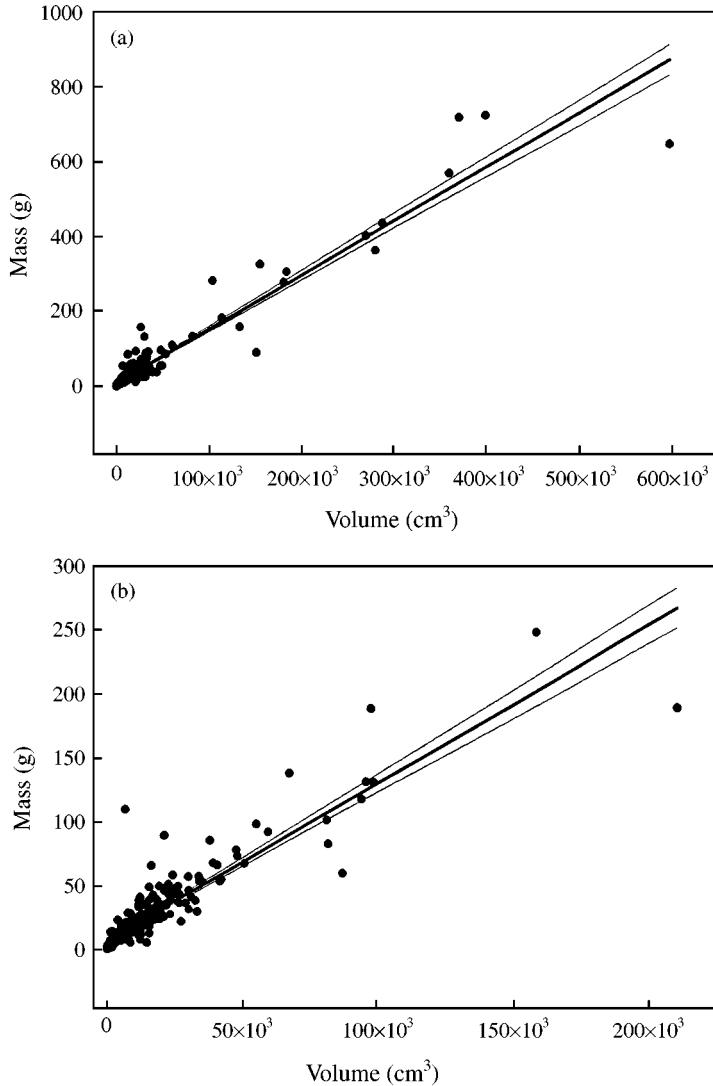


Figure 3. Sample regression lines and 95% confidence intervals for two dominant species. (a) fall biomass vs. volume for *Larrea tridentata*; (b) fall biomass vs. volume for *Bouteloua eriopoda*.

Adequacy of sample size

We used a jack-knife or resampling technique to assess the adequacy of sample size (number of quadrats) for estimates of mean biomass per site and mean ANPP per time interval per site. We assumed that sites with the greatest 'patchiness' of biomass would be the most difficult to sample adequately, and therefore selected the most patchy shrubland, the North Mesquite site, and the most patchy grassland, the IBP Grassland, for investigation. We chose 1991 as a 'typical' year for plant production (neither drought nor unusually wet conditions), and used data from spring 1991 and fall 1991 for the assessment. These data consisted of estimated total biomass for each of the 49 quadrats per site, and estimated net above-ground production over the February–May (spring) and May–October (fall) intervals for each of the 49 quadrats. The 49 values for a site

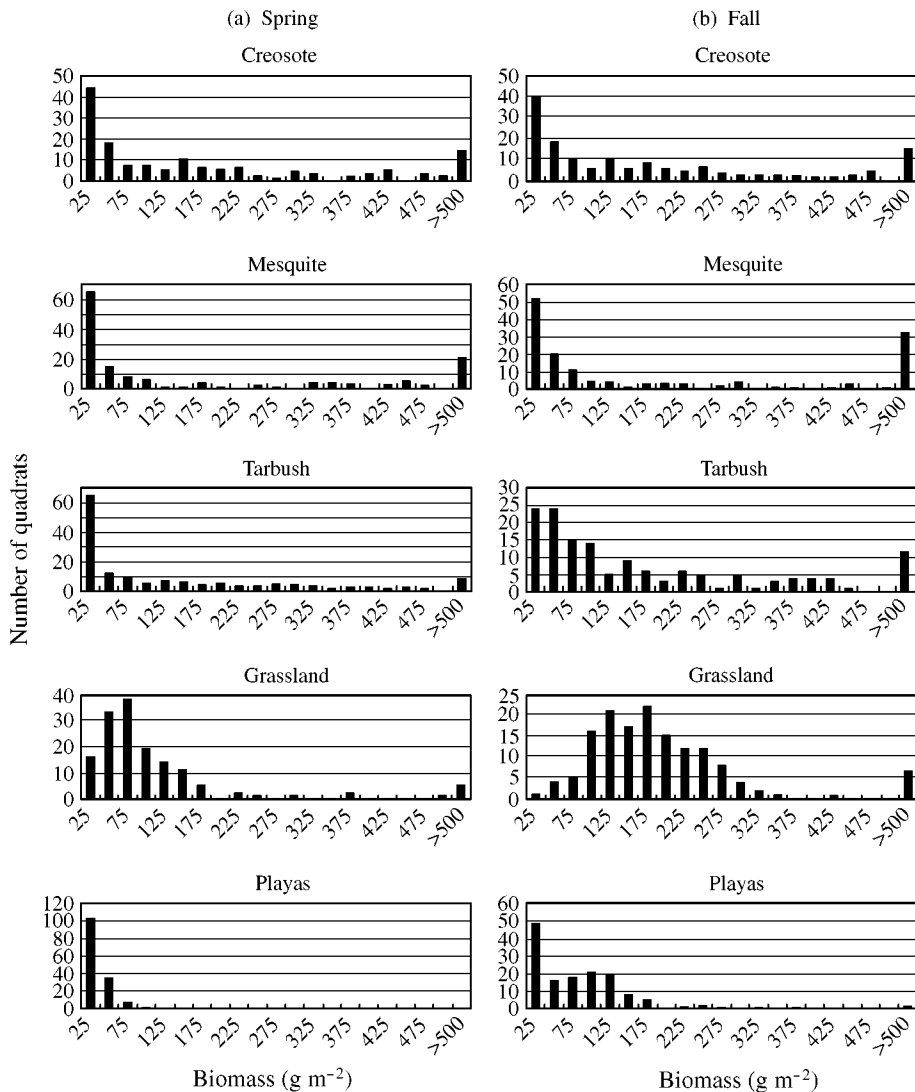


Figure 4. Frequency distributions of total biomass values (g m^{-2}) for individual quadrats from five ecosystem types at the Jornada LTER site. All quadrats from the three sites per ecosystem type are pooled. (a) spring 1991; (b) fall 1991.

were randomly ordered; then a running mean was calculated for samples of the range 1–49 quadrats. A given sample size was considered ‘adequate’ when the running mean reached (and stayed within) the interval of the final mean $\pm 10\%$. The randomization procedure and running average calculations were repeated 100 times for each data set. We inspected both the plots of running means *vs.* cumulative sample size (number of quadrats), and the frequency distribution of adequate sample sizes for the 100 runs.

Mean biomass for the IBP grassland site was 106 g m^{-2} in spring 1991 and 237 g m^{-2} in fall 1991. Most of the resampling runs reached adequate sample size (running mean consistently within $\pm 10\%$ of the final mean) at fewer than 45 quadrats (Fig. 5a). There were a few sample runs, especially in spring when mean biomass was relatively low, in which adequate sample size was not reached until nearly 49 quadrats. The patchiness of

Table 1. Ranked ranges of variation in values of biomass and of seasonal production for fall 1991 data from 15 sites at the Jornada Basin LTER

Site or scale	Range of biomass values	Site	Range of production values
P-SMAL	33	<i>T-ALL</i>	13
<i>T-ALL</i>	77	P-SMAL	29
<i>P-ALL</i>	113	<i>C-ALL</i>	30
<i>C-ALL</i>	142	C-CALI	64
<i>G-ALL</i>	144	<i>G-ALL</i>	69
<i>ALL SITES</i>	206	<i>ALL SITES</i>	80
<i>M-ALL</i>	236	<i>P-ALL</i>	107
P-COLL	323	<i>M-ALL</i>	113
P-TOBO	516	P-TOBO	175
C-CALI	527	G-SUMM	192
G-BASN	620	C-SAND	227
T-WEST	645	C-GRAV	248
T-TAYL	739	M-RABB	261
C-GRAV	943	G-BASN	269
M-RABB	1060	T-WEST	303
M-WELL	1087	T-EAST	321
C-SAND	1119	P-COLL	365
T-EAST	1245	T-TAYL	437
M-NORT	2078	M-WELL	696
G-IBPE	3445	M-NORT	1082
G-SUMM	3546	G-IBPE	2226

Ranges for individual sites (capital font) represent the total range of values (g m^{-2}) for the 49 quadrats per site. Ranges for ecosystem types (bold italic font) represent the range of the mean values for the 3 sites per ecosystem type. C, creosote sites; G, grassland sites; M, mesquite sites; P, playas; T, tarbush sites; four-letter acronyms following letter indicate specific site name. Ranges for all sites (italic font) represent the range of the mean values for the five ecosystem types.

biomass in the IBP grassland was caused by the occurrence of just two large clumps of *Yucca elata*; these large woody succulents caused unusually high values of biomass in two quadrats relative to the other 47.

The North mesquite site is dominated by large 'coppice dunes' of *Prosopis glandulosa*, where blown soil and organic matter have accumulated at the base of the low multi-stemmed clumps, separated by extensive, nearly bare inter-dune spaces. Mean above-ground biomass for this site was 253 g m^{-2} in spring 1991 and 412 g m^{-2} in fall 1991. Again, most sampling runs reached adequate sample size in fewer than 45 quadrats (Fig. 5b). A substantial portion of the sample runs reached adequate size in as few as 30 quadrats.

Net productivity for the interval February–May 1991 (spring 1991) averaged 55 g m^{-2} for the IBP grassland, compared with 151 g m^{-2} for the same site in fall 1991 (May–Oct interval). Adequate sample sizes for the spring sampling runs ranged from 40 to as high as 49 (running mean not stabilizing until the final quadrat sampled), but again the majority of sample runs reached adequate sample size at fewer than 45 quadrats (Fig. 5c). In fall 1991, adequate sampling sizes ranged from fewer than 10 quadrats to 46 quadrats.

Productivity for the North Mesquite site increased from an average of 34 g m^{-2} for the spring 1991 interval, to 164 g m^{-2} for fall 1991. Adequate sample sizes ranged from 20 quadrats to roughly 45 quadrats for the large majority of sample runs (Fig. 5d).

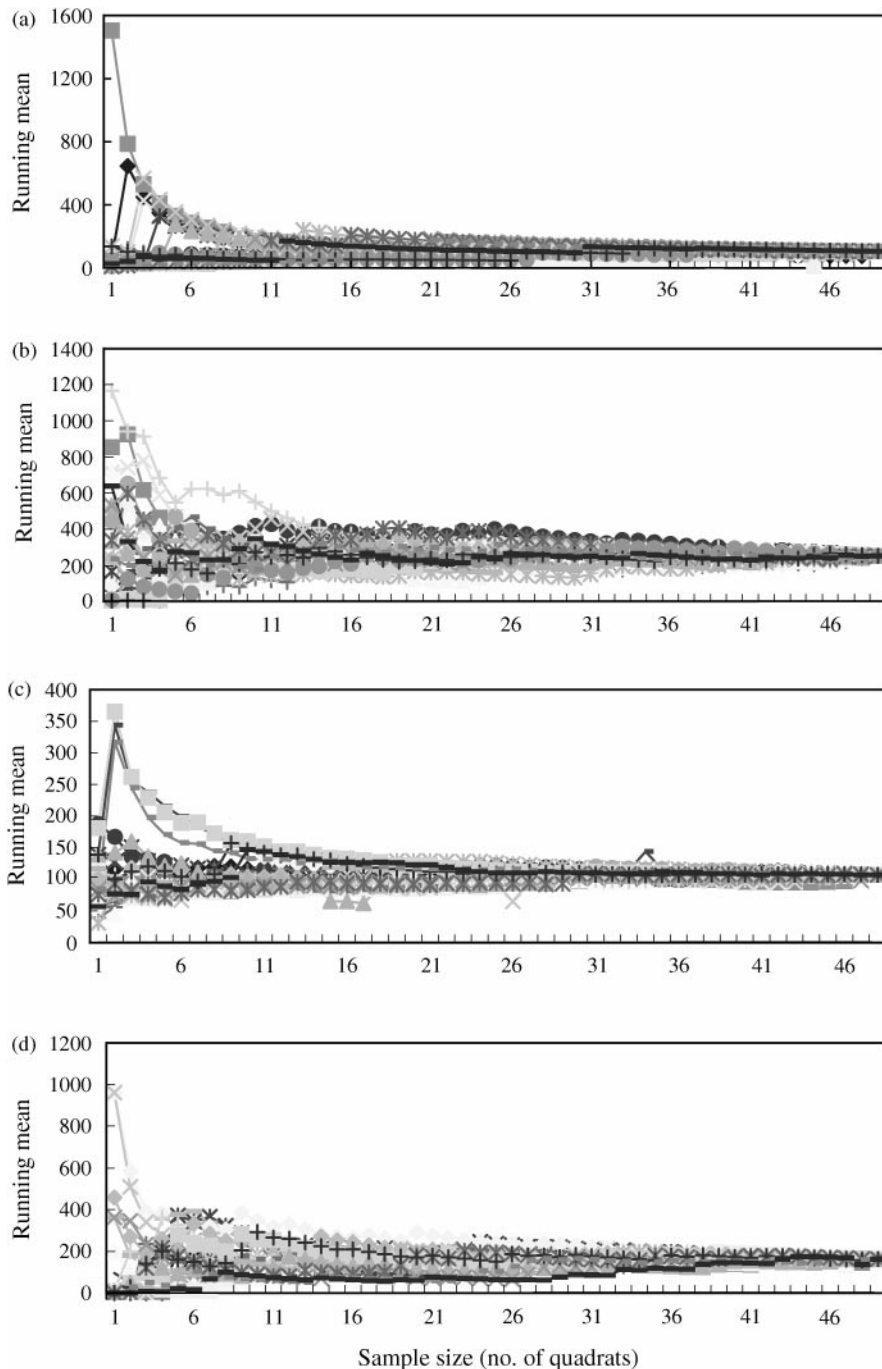


Figure 5. Running means against sample sizes for jack-knifed calculations of estimated mean biomass and mean productivity for two Chihuahuan Desert sites. The 49 quadrats per site were resampled in random order and the running mean for the parameter of interest was calculated, for 100 runs per site. For sake of clarity, only the first and the last run to converge to within 10% of the final mean value, and 25 other randomly selected runs of the 100, are shown. (a) running mean of biomass (g m^{-2}) in spring 1991 in the IBP grassland site; (b) running mean of biomass (g m^{-2}) in spring 1991 in the North Mesquite site; (c) running mean of ANPP (g m^{-2}) for the summer-fall interval in 1991 in the IBP grassland site; (d) running mean of ANPP (g m^{-2}) for the summer-fall interval in 1991 in the North Mesquite site.

Results of field sampling program

Results of the first few years of data collection demonstrated the feasibility of our approach for estimating biomass and productivity in these semi-arid ecosystems. Differences between sites in a season, or between seasons for a given site, for mean above-ground biomass typically exceeded 10% of the average value by a comfortable margin. Often the differences were in the order of 50% or more of the mean value for the three sites in an ecosystem type; for example, mean above-ground biomass values for the three *Larrea*-dominated shrublands in fall 1991 were 103, 192, and 245 g m⁻², while the three grassland values for that sample period were 172, 237, and 315 g m⁻². The same pattern holds true for seasonal productivity estimates. Winter season values for most ecosystem types range from 0 to 30 g m⁻²; spring (February–May) values, usually highest in shrub-dominated systems, range from 50 to 150 g m⁻²; and late summer productivity (May–October) values are in that same range, usually highest for grasslands. Summing across an entire year (1991), mean annual ANPP ranged from 48 g m⁻² yr⁻¹ in one playa site to around 250 g m⁻² yr⁻¹ in two of the grasslands and nearly 325 g m⁻² year⁻¹ in another playa. Our results confirm that ANPP can be quite variable between years in these semi-arid ecosystems, subject as they are to considerable inter-annual differences in precipitation. However, preliminary results do not suggest dramatic and consistent differences between ecosystem types in ANPP.

Discussion

Our system for estimating ANPP has a number of advantages. Chief among these are the comparability and consistency of methods and sampling intensity in ecosystems of varying structure; the relative efficiency of field sampling, allowing the study of a large number of sites (15) multiple times per year; and, after initial harvests are completed, minimal destructive sampling and impact on long-term study sites. Resulting data sets contain valuable information on both the relative abundance of species in the community (biomass estimates) and the functional contributions of those species (seasonal productivity estimates).

Our methodology, like any other, is a trade-off between advantages and deficiencies. Quadrat size, sampling frequency, and plant dimensions used as measures of volume are all compromises; each could be optimized for particular species or species groups, but we have foregone this optimization in favor of efficiency and consistency of field sampling. Our 15 dispersed field sites can be sampled with a 4–6 person team in 5–15 days of field work, depending on the biomass and diversity of annuals present in a given season. Like most other measures of ANPP based on physical samples or measurements of plants, it is growth of above-ground tissues that is estimated, not productivity in the sense of photosynthetic carbon fixation.

As reviews of available methodologies suggest, measuring net primary productivity is not a straightforward exercise (Singh *et al.*, 1975; Whittaker & Marks, 1975; McNaughton *et al.*, 1996). Major approaches might be classified as plot-based, plant-based, and physiological; each of these has limitations in the context of a long-term program aimed at quantifying spatial and temporal variability. The typical approach for grasslands is the repeated harvest of replicate plots over time. This is logistically difficult in slow-growing perennial vegetation, as one quickly exhausts the supply of undisturbed replicate plots in a long-term study. Spatial variability exacerbates the problem, as a large number of replicates is needed at each sample date to provide an adequate sample. Harvest or clipping has its own limitations of interpretation, even in grasslands (Singh *et al.*, 1975; Lauenroth & Whitman, 1977), especially if grazing or protection from grazing must be taken into consideration (McNaughton *et al.*, 1996).

Conventional measures of productivity in western U.S. rangelands are usually based on visual estimates, with calibration data rarely presented. Use of these estimates for

comparative purposes is problematic since one must take on faith consistency between observers and over time. Also, reported rangeland productivity values often include only forage species, omitting those plants not often used by livestock; thus it is difficult to compare these values with those obtained in any other way for the entire plant community.

Conventional plant-based methods such as dimension analysis can lead to large multiplicative errors in scaling up from individual plants to some area-based estimate of production. In practice, the labor (and the destructive harvests) required to make such measurements limits investigators to no more than a handful of dominant species in an ecosystem. This is a real problem in diverse systems where there is value in understanding the productivity of a large number of species. Finally, these plant-based measures make it difficult to relate productivity estimates to specific areas or to understand spatial patterns of productivity.

Seasonal measurements of photosynthetic rates and carbon fixation (or models of those processes) will not necessarily be well correlated in time with the appearance of above-ground growth, as much fixed carbon can be stored in below-ground tissues. We acknowledge that much of the above-ground 'growth' we measure comes from stored reserves and not from current photosynthesis. Annual totals, though, should be more closely correlated. Detailed physiological approaches are also difficult to apply to the diverse mix of species in these semi-arid communities, and are not easily adapted to assessing spatial patterns of productivity.

Any technique involving episodes of sampling at discrete dates will lead to errors of underestimation, for example, where leaf turnover is rapid, damage by herbivores is severe, or some annual species mature and senesce in between sample dates. We presume these errors of underestimation to be considerable in this as in other environments, but have attempted to minimize these by sampling seasonally and by adjusting sampling dates according to phenology.

Errors of overestimation come from the inherent bias of using increments or differences between sample dates as the basic measure of productivity. Negative increments, whether true declines in the standing crop of a plant or simple measurement error, are ignored, while all positive ones (some of which might be error) are counted as production (Sala *et al.*, 1988; Biondini *et al.*, 1991). Biondini *et al.* (1991) asserted that these overestimation errors will usually outweigh underestimations. The statistical technique of Sala *et al.* (1988) uses the observed variances of the biomass estimates at each sample date to adjust the mean differences to eliminate the overestimation error; application of the statistical correction to every increment for each species in our 15 sites would be unwieldy, though possible. For our purposes, though, we are satisfied with comparative estimates. Our greatest interest is in the patterns of variability between sites and between sample periods; given that all our values are subject to the same limitation, we can compare and contrast values obtained for different sites or for different years with confidence.

Our approach allows a quantitative characterization of spatial heterogeneity in semi-arid ecosystems, with important lessons for researchers. First, patchiness within individual sites is tremendous, far exceeding the differences in mean biomass between sites or even between ecosystem types of vastly differing structure. This patchiness poses severe challenges for sampling; nearly 50 quadrats (or 50 m²) per site are sometimes necessary to arrive at any confidence in estimates of mean above-ground properties. Many studies in the literature have presented estimates of cover or biomass based (apparently) on less intensive sampling, and sample size adequacy is rarely documented. We suggest caution, therefore, in using published values from different studies to compare cover, biomass or productivity estimates unless such documentation is available.

Second, the dramatic non-normality of distribution of biomass values means that summary statistics such as the coefficient of variation or the standard deviation or

standard error (so commonly presented) are not useful measures of the real variability of estimates of biomass and production. These parameters are quite sensitive to non-normality in the underlying distributions, and the distribution of biomass across the surface of these ecosystems is definitely non-normal. Simple transformations (e.g. a logarithmic transformation) of our data were not sufficient to arrive at normality. Hence we recommend strongly that researchers use caution and select more appropriate measures of variability when comparing means or variation between sites or in the same site over time.

We thank Dave Lightfoot and John Anderson for many hours of assistance in development and testing of field methods. John Anderson has supervised the hiring, training and work of scores of NPP field workers since 1981, and has borne heavy responsibility for quality control and assurance of consistency in plant identification, measurement methods and data management. We thank Neuza Buerkle for her consistent and meticulous work in sorting plant material in the laboratory. The NMSU Statistics Center provided assistance with programs for data analysis. Jeffrey E. Herrick, John Anderson, and Debra Coffin Peters provided extremely helpful reviews of an earlier version of the manuscript; we appreciate the assistance of Jeff Herrick and Michelle Buonopane with graphics. The NPP sampling program was established and initiated under Jornada LTER-II funding (to NMSU, NSF BSR-88-11160, continued under Duke University as DEB-92-40216). Manuscript preparation was supported by the Jornada LTER-III program, NSF DEB-94-111971 to Duke University.

References

- Allred, K.W. (1993). *A Field Guide to the Grasses of New Mexico*. Las Cruces, NM: Dept. of Agricultural Communications, New Mexico State University. 258 pp.
- Biondini, M.E., Lauenroth, W.K. & Sala, O.E. (1991). Correcting estimates of net primary production: Are we overestimating plant production in rangelands? *Journal of Range Management*, **44**: 194–198.
- Buffington, L.C. & Herbel, C.H. (1965). Vegetational changes on a semidesert grassland range from 1858 to 1963. *Ecological Monographs*, **35**: 139–164.
- Correll, D.S. & Johnston, M.C. (1970). *Manual of the Vascular Plants of Texas*. Renner, TX: Texas Research Foundation. 1881 pp.
- Gibbens, R.P. & Beck, R.F. (1988). Changes in grass basal area and forb densities over a 64-year period on grassland types of the Jornada Experimental Range. *Journal of Range Management*, **41**: 186–192.
- Kearney, T.H. & Peebles, R.H. (1964). *Arizona Flora*. Berkeley, CA: University of California Press. 1085 pp.
- Lauenroth, W.K. & Whitman, W.C. (1977). Dynamics of dry matter production in a mixed-grass prairie in western North Dakota. *Oecologia*, **27**: 339–351.
- McNaughton, S.J., Milchunas, D.G. & Frank, D.A. (1996). How can net primary production be measured in grazing ecosystems? *Ecology*, **77**: 974–977.
- Reynolds, J.F., Virginia, R.A., Kemp, P.R., de Soyza, A.G. & Tremmel, D.C. (1999). Impact of drought on desert shrubs: effects of seasonality and degree of resource island development. *Ecological Monographs*, **69**: 69–106.
- Sala, O.E., Biondini, M.E. & Lauenroth, W.K. (1988). Bias in estimates of primary production: an analytical solution. *Ecological Modelling*, **44**: 43–55.
- SAS Institute, Inc. 1989–1996. SAS System for Windows, v. 6.12. SAS Institute, Cary, NC, USA.
- Schlesinger, W., Reynolds, J., Cunningham, G., Hueneke, L.F., Jarrel, W., Virginia, R. & Whitford, W. (1990). Biological feedbacks in global desertification. *Science*, **247**: 1043–1048.
- Singh, J.S., Lauenroth, W.K. & Steinhorst, R.K. (1975). Review and assessment of various techniques for estimating net aerial primary production in grasslands from harvest data. *Botanical Review*, **41**: 181–232.
- Whittaker, R.H. & Marks, P.L. (1975). Methods of assessing terrestrial production. In: Lieth, H. & Whittaker, R.H. (Eds), *Primary Production of the Biosphere*, pp. 55–118. Ecological Studies 14. New York: Springer-Verlag. 339 pp.