The Impact of Experimental Design on the Application of Grazing Research Results—An Exposition

M.A. BROWN AND S.S. WALLER

Abstract

Funding limitations often restrict pasture replication in grazing research on rangeland. Consequently, subsample error has been used to estimate treatment effects or characterize populations. Assumptions associated with experimental designs which utilize subsample error to make inferences are discussed and an example evaluated. The appropriate experimental unit for inferential grazing research is the pasture. Animals or vegetation sampling within pastures must be considered as subsamples in inferential grazing research. Pasture replication must be used in intensive grazing trials to establish treatment differences or provide adequate characterization. Following intensive trials, extensive, unreplicated trials implemented by private producers can be effective in establishing broad-based applicability. Unreplicated pasture trials may also be used for screening several treatments.

Shrinking research budgets, increased competition for research dollars, and increased complexity of grazing research have contributed to efforts to enhance the efficiency of the research process. Additionally, range and pasture research competes with other research for land in organizations with finite land resources. Differentiating between efficiency of experimental design and the ability to make the proper characterization or inference is particularly important in range science. Minimizing cost in the experimental design process often is done at the expense of inference. Most books on experimental design or statistics provide few examples representing experimentation at the population, community, or ecosystem levels of organization (Hurlbert 1984). The principles of design for grazing studies are most often violated. It is the purpose of this discussion to explore some aspects of experimental design critical to evaluating efficiency and cost versus interpretive credibility of grazing research.

Hurlbert (1984) defined experimental research as mensurative (sampling studies) or manipulative (external treatment). Mensurative experiments involve only the making of measurements at one or more points in space or time; space or time is the only experimental variable or "treatment". Cochran (1977) provided an excellent review of sampling techniques for sampling studies. Manipulative experiments always involve two or more treatments and have as goals making one or more comparisons. Several references provide valuable information on the design of manipulative experiments (Federer 1955, Cochran and Cox 1957, Cox 1958). Measuring attributes of interest is important in both mensurative and manipulative experiments. There are several references available on measurement techniques (Brown 1954, U.S. Forest Service 1963, Greig-Smith 1964, Mueller-Dombois and Ellenberg 1974, 't Mennetge 1978, Stubbendieck and Schacht 1984, Cook and Stubbendieck 1986). Experiments can be inferential, providing information about a population; or they can be descriptive, providing information about specific individuals within a population.

Inferential Grazing Trials

Population of Inference

The focal point of inferential research is the characterization of or inference about some population. Steel and Torrie (1980) defined a population as all possible values of a variable. The selected population of interest becomes the population of inference from which individuals are selected for experimentation. The population could be all native pastures in eastern Nebraska for a grazing study or all silty range sites in Vegetative Zone IV of Nebraska for a vegetative survey. Explicit definition of the population of inference must be made prior to the selection of the experimental material. If the population is well defined; means, variances, covariances, probabilities and other statistics generated from a research project can be properly interpreted. Once the population is defined, sample units can be selected within the appropriate experimental material. Sample units could be termed observational units in descriptive research and experimental units in experimental research.

Experimental Unit

A proper experimental design can be destroyed by failing to recognize what constitutes the experimental unit (Nelson and Rawlings 1983). According to Cox (1958) the experimental unit corresponds to the smallest division of experimental material such that any two units may (independently) receive different treatments in the actual experiment. Because responses are not constant in biological research, sample units should be chosen at all levels which can affect characterizations or responses. For example, treatment, pasture, animal, year of measurement, year of treatment, the failure of main effects to respond consistently within the grazing season, and appropriate interactions are all potential sources of variability in a manipulative grazing trial.

Animals can be considered experimental units in mensurative and manipulative research trials in which the forage resource has no differential effect on the measured response. Breed comparisons, reproductive physiology, insect or parasite studies are examples of such studies. Prior knowledge is the basis for the assumption that pasture effects are negligible or of no interest. In identifying the animal as the experimental unit and not replicating pastures, the researcher foregoes the opportunity to evaluate potential pasture × treatment interactions.

Animals in manipulative grazing research on rangeland are generally used as a treatment (e.g., stocking rate study) or as measures of treatment effect (e.g., gain/head, gain/area). Free-grazing animals in range situations can seldom be considered as experimental units. Each animal must receive a treatment to be considered an experimental unit in manipulative research. In addition, each animal must be independent of other animals for the response measured. If forage availability or selection are important factors in the response variable, animals within a pasture cannot be considered independent. As an example, forage consumed by one animal cannot be consumed by another, implying a dependency. Thus, animals are actually repeated measures on the same experimental unit. The experimental unit of a production system in the measurement of animal production from grasslands must consist of an area of land and the animals grazing on it as well as auxiliary facilities for such management activities as supplemental feeding (Morley 1978). Pasture size and number is a critical consideration for research facilities with limited land resources. Pastures should be large enough to support adequate animal numbers to provide the precision in response measurements (Peterson and Lucas 1960). However, small pastures are more desirable than large, production-size pastures for intensive grazing trials since land is restricted. Additionally, initiation of new grazing studies should

Authors are research animal scientist, USDA, ARS, Rt. 2, Box 144A, Booneville, Ark. 72927; professor, Department of Agronomy, University of Nebraska-Lincoln, Lincoln 68583-0915.

Manuscript accepted 8 August 1985.

not be confounded with previous grazing trials, emphasizing the importance of "protecting" available land resources from experimentation when possible.

Replication

Relication is the observation of more than one experimental unit treated alike. It provides evidence of repeatability and is the only way to provide an estimate of experimental error (degree of variability among individuals treated alike). An estimate of experimental error is required for tests of significance and for confidence interval estimation (Steel and Torrie 1980). Without estimates of experimental error it is impossible to distinguish real differences from inherent variation.

In grazing trials, pastures and animals should be observed within a grazing season over a number of years to estimate these sources of variability for characterization or inference. To avoid bias, pastures and animals would be sampled and/or assigned treatments at random. For an estimate of variability, at least two independent units per source are needed; i.e., in manipulative research two pastures per treatment and two animals per pasture observed over the grazing season with the experiment conducted for at least two years. The actual numbers chosen would be a function of the desired precision in mensurative research or precision and the probability of acceptable decision error in manipulative research. Thus, it is apparent that replication of pastures, animals, and years is appropriate for proper characterization and inference.

Most evaluations of grazing trial research have been conducted on introduced pastures, having a greater degree of homogeneity than native range. Green et al. (1953) concluded that replication of both pastures and animals in pastures was essential for valid conclusions in grazing trials conducted on a plot scale. Mott and Lucas (1953) concluded that the animal component of variance was most important when measuring product per animal whereas the animal and pasture components of variance were of equal importance when measuring forage yield in research on cultivated pastures. They suggested using several small pastures per treatment with one to three animals per pasture for evaluating both animal performance and forage yield. Additionally they recommended using larger pastures (more animals) when evaluating animal performance. Peterson and Lucas (1960) reported the important sources of variability in animal performance to be the between-animal variance; the animal × time interaction and the pasture × time interaction. Important sources of variability for yield per area were the pasture × time and animal × time interaction.

It is important that grazing trials be designed to be able to estimate these sources of variability to ascertain which sources of variability are important for the responses being measured. These sources of variability can then be tested for significance and insignificant sums of squares can be pooled.

Pseudoreplication

Pseudoreplication can be defined as testing treatment effects with an error term inappropriate with the hypothesis being considered (Hurlbert 1984). Interpreting a significant difference as a "treatment effect" or real difference using the variability of subsamples is a form of pseudoreplication. Pasture replication is often omitted in research due to funding limitations. In such cases vegetation variables are measured with subsamples within a pasture (transects, quadrats, or exclosures for example) while animal responses are determined using animals within a pasture. The transect or the animal in unreplicated pastures must be considered the experimental unit to be provided an experimental error term. However, transects within a pasture do not have an equal opportunity to receive different treatments; therefore they cannot be considered experimental units (Cox 1958). If "replicates" are only samples from a single experimental unit (pasture) then replicates are not independent (Hurlbert 1984).

An Example Grazing Trial

An hypothetical experiment testing the null hypothesis that there is no difference in gain/head or vegetation composition between fertilized and unfertilized range can illustrate the limitations associated with unreplicated pastures (Table 1). In this example there were 8 pastures with a dividing fence in each. One half of each pasture was fertilized and the remaining half was left unfertilized. Twenty steers grazed in each treatment to document animal performance and 20 cages were placed at random in each treatment in each pasture to estimate kinds and amounts of vegetation. The experiment was run for four consecutive years without rerandomization. Year effects are repeated measurements of the same experimental units reflecting the cumulative effects of fertilization as well as weather and cannot be considered replications of the experiment. The linear model for this design is as follows:

$$Y_{ijkl} = \mu + P_i + F_j + PF_{ij} + T_k + PT_{ik} + FT_{jk} + PFT_{ijk} + e_{l(ijk)}$$

where:

Y_{ijkl} = response of the l^{th} subsample in the k^{th} year, j^{th} fertilizer regime and i^{th} pasture

μ = overall population mean

P_i = random effect of the i^{th} pasture

F_j = fixed effect of the j^{th} fertilizer regime

PF_{ij} = random interaction effect of the i^{th} pasture with the j^{th} fertilizer regime

T_k = fixed effect of the k^{th} year

PT_{ik} = random interaction effect of the i^{th} pasture with the k^{th} year

FT_{jk} = fixed interaction effect of the j^{th} fertilizer regime with the k^{th} year

PFT_{ijk} = random interaction of the i^{th} pasture, j^{th} fertilizer regime and k^{th} year

$e_{l(ijk)}$ = random effect of the l^{th} subsample in the k^{th} year, j^{th} fertilizer regime and i^{th} pasture.

Table 1. Sources of variation, degrees of freedom (df) and expected mean squares E (MS) for two response variables.

Gain/head			Vegetative composition		
Source	d.f.	E(MS)	Source	d.f.	E(MS)
Pasture (P)	7	$\sigma_A^2 + 160\sigma_P^2$	Pasture (P)	7	$\sigma_C^2 + 160\sigma_P^2$
Fertilizer (F)	1	$\sigma_A^2 + 80\sigma_{PF}^2 + 640\ \Sigma F^2$	Fertilizer (F)	1	$\sigma_C^2 + 80\sigma_{PF}^2 + 640\ \Sigma F^2$
P × F	7	$\sigma_A^2 + 80\sigma_{PF}^2$	P × F	7	$\sigma_C^2 + 80\sigma_{PF}^2$
Year (T)	3	$\sigma_A^2 + 40\sigma_{TP}^2 + 320\ \frac{\Sigma T^2}{3}$	Year (T)	3	$\sigma_C^2 + 40\sigma_{TP}^2 + 320\ \frac{\Sigma T^2}{3}$
T × P	21	$\sigma_A^2 + 40\sigma_{TP}^2$	T × P	21	$\sigma_C^2 + 40\sigma_{TP}^2$
T × F	3	$\sigma_A^2 + 20\sigma_{TPF}^2 + 160\ \frac{\Sigma[FT]^2}{3}$	T × F	3	$\sigma_c^2 + 20\sigma_{TPF}^2 + 160\ \frac{\Sigma[FT]^2}{3}$
T × P × F	21	$\sigma_A^2 + 20\sigma_{TPF}^2$	T × P × F	21	$\sigma_C^2 + 10\sigma_{TPF}^2$
Animal nested in P, F, T	1216	σ_A^2	Cage nested in P, F, T	1216	σ_C^2
Total	1279		Total	1279	

Thus the design is a split plot on time with subsampling with the main unit a randomized complete block design and the subunit time. This design is only one of a large number of designs that might be appropriate in grazing trials. It includes pasture effects, interactions with pasture, and animal or cage effects. For simplicity it has ignored repeated measurements on animals or cages within a grazing season.

The expected mean squares (Table 1) indicate the appropriate tests for fertilizer effects, year effects and year by fertilizer effects:

F_F = Fertilizer Mean Square (MS)/ Pasture × Fertilizer MS

F_T = Year MS/ Pasture × Year MS

F_{FT} = Fertilizer × Year MS/ Pasture × Fertilizer × Year MS

Using subsampling variance would be valid only if the pasture × fertilizer, pasture × year, and pasture × fertilizer × year effects were zero, depending on the particular test. These assumptions are questionable given the inherent variability of any biological system and the confounding effects of pre-treatment management on both pasture and animal. Hurlbert (1984) stated that in any field situation, two experimental units are different in every measurable property. According to Federer (1955) all fields of research have at least one feature in common; i.e., the variability of experimental material. Assuming no interaction with pasture is particularly unrealistic when evaluating range with large pastures possessing range site and species diversity.

Descriptive Grazing Trials

Osborne and Reid (1953) discussed the design of range grazing trials and recommended replication of pastures in intensive studies. They differentiated between intensive and extensive studies. Intensive studies would be used to plan extensive studies (large scale evaluation). Unreplicated pastures in an intensive study would confound treatment and pasture differences and make it impossible to estimate variances of treatment differences. The pasture variance component would not be available and the pasture × time interaction would be confounded with the treatment × time interaction. Since extensive trials are generally not possible at most range research institutions, intensive studies are directly extrapolated and implemented by private individuals. This places an extreme burden on the accuracy of inferences made from intensive grazing trials and emphasizes the importance of pasture replication.

Replication is often restricted by limitations in resources. The importance of replication should not prohibit the accumulation and publication of descriptive information gathered from unreplicated trials; it should only caution the researcher in the limits of any inference. Within unreplicated studies random variation is most often expressed as variability among subsamples (vegetation or animal measurements) determined in the pasture. This is adequate for descriptive and inferential purposes within the experimental area only. But as previously illustrated, this approach assumes no pasture to pasture variability. Any inference beyond the study pasture is unsubstantiated and the scientist is unsure if differences are real or the result of inherent variation. However, some researchers have, based on literature and prior experience, selected experimental areas considered to be "most representative" of the population of desired inference and made inferences. This does not allow quantitative estimates of variability and probabilities and relies totally on experimental information to identify "representative" areas and accurately anticipate the validity of the measured response within the area of extrapolation. A more logical sequence of inferential grazing research is:

1. Evidence in the replicated trials suggested a real treatment difference so that unreplicated extensive studies on larger areas are warranted to substantiate response to treatment. When research institutions do no have land resources for large scale studies, applicability should be estimated through limited implementation by private producers.
2. Results from the extensive studies suggested the same conclusions as those of the initial, replicated study. Treatment implementation appears feasible on an industry-wide basis. However, continued monitoring and evaluation of the treatment response across the broad range of industry use should be encouraged to provide continued refinement of the treatment.

The unreplicated pasture study has a place in the research process. It can serve as a screening trial for several treatments, after which intensive replicated trials are initiated with the most promising treatments. It can also follow replicated trials to establish a broad based applicability. Often it is necessary to conduct unreplicated extensive trials because the scale effects of such experiments are necessary to allow expression of treatment differences and the scale prohibits replication. If conducted properly, this information can add new knowledge and should be published. However, a single study should never constitute the only source of information regarding a treatment effect. Replication of large scale studies can be done in time and space (as well as is possible to replicate in time) by cooperative efforts of research groups over a number of years. A cooperative, regional approach to grazing trials can provide the proper inference without excessive cost assigned to any one project or institution. A regional approach does require common planning and agreement on hypotheses by the various researchers. This compromise and team effort among researchers results in much more reliable information for the producer.

Conclusions

Experimental design of comparative range and pasture grazing trials should include sufficient replication of land, animals and time to properly estimate variances at an acceptable level of precision for characterization or inference. Efficiency can be addressed in terms of insuring the experiment is not over-replicated, insuring that experimental units are not over-sampled, using experimental designs that can increase experimental precision (such as randomized block, split plot, incomplete block, lattice, etc.) insuring the size of the experimental units is not larger than needed for proper inference, increasing the efficiency of use of experimental material by more efficient treatment designs or overlaying nonconflicting research on the experimental material. Proper experimental design is a fundamental premise of successful research.

Literature Cited

Brown, D. 1954. Methods of surveying and measuring vegetation. Commonwealth Bur. Pastures and Field Crops. Hurley, Berkshire, England. Bull. 42.

Cochran, W.G. 1977. Sampling techniques. John Wiley and Sons, Inc. New York.

Cochran, W.G., and G.M. Cox. 1957. Experimental designs. John Wiley and Sons. New York.

Cook, C.W., and J.L. Stubbendieck (eds.) 1986. Range research: basic problems and techniques. Soc. Range Manage. Denver, Colo.

Cox, D.R. 1958. Planning of experiments. John Wiley and Sons. New York.

Federer, W.T. 1955. Experimental design: theory and application. Oxford and IBH Pub. Co. Calcutta, India.

Green, J.O., H.J. Langer, and T.E. Williams. 1953. Sources and magnitudes of experimental errors in grazing trials. Proc. Sixth Internat. Grassl. Congr. 2:1374-1379.

Greig-Smith, P. 1964. Quantitative plant ecology. Butterworths, Inc. Washington, DC.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monog. 54:187-211.

Morley, F.H.W. 1978. Animal production studies on grassland, p. 103-162. *In:* L. 't Mannetje (ed.) Measurement of grassland vegetation and animal production. Commonwealth Bur. Pasture and Field Crops. Hurley, Berkshire, England. Bull. 52.

Mott, G.O., and H.L. Lucas. 1953. The design, conduct and interpretation of grazing trials on cultivated and improved pastures. Proc. Sixth Internat. Grassl. Congr. 2:1380-1385.

Mueller-Dombois, D., and H. Ellenberg. 1974. Aims and methods of vegetation ecology. John Wiley and Sons. New York.

Nelson, LA., and J.O. Rawlings. 1983. Ten common misuses of statistics in agronomic research reporting. J. Agron. Educ. 12:100-105.

Osborne, J.G., and E.H. Reid. 1953. The design, conduct and interpretation of range grazing studies. Proc. Sixth Internat. Grassl. Congr. 2:1386-1391.

Peterson, R.G., and H.L. Lucas. 1960. Experimental errors in grazing trials. Proc. Eighth Internat. Grassl. Congr. 8:747-750.

Steel, R.G.D., and J.H. Torrie. 1980. Principles and procedures of statistics. McGraw-Hill Book Co. New York.

Stubbendieck, J., and W. Schact. 1984. Rangeland analysis. Univ. Nebraska-Lincoln. Lincoln, NE.

't Mennetge, L. (ed). 1978. Measurement of grassland vegetation and animal production. Commonwealth Bur. Pasture and Field Crops. Hurley, Berkshire, England. Bull. 52.

U.S. Forest Service. 1963. Range research methods, Misc. Pub. No. 940. USDA. U.S. Government Printing Office, Washington, DC.