Ecology, 66(2), 1985, pp. 606–611

USE OF AN IMPROVED STATISTICAL METHOD FOR GROUP COMPARISONS TO STUDY EFFECTS OF PRAIRIE FIRE[1]

GREGORY M. ZIMMERMAN
Department of Range Science, Colorado State University, Fort Collins, Colorado 80523 USA

HAROLD GOETZ
Department of Botany, North Dakota State University, Fargo, North Dakota 58105 USA

AND

PAUL W. MIELKE, JR.
Department of Statistics and Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado 80523 USA

Abstract. This paper describes an improved method for performing statistical comparisons among experimental groups. This technique, termed multi-response permutation procedures (MRPP), is similar in purpose to the *t* test and one-way analysis of variance *F* test. However, in contrast to these, the new method features very relaxed requirements on the data structure, is easily applied to multivariate problems, and makes it possible to relate the analysis visually to the perceived data space. The MRPP test statistic is based on the within-group average of pairwise distance measures between object response values in a euclidian data space. The null distribution of the test statistic is based on the collection of all possible permutations of the objects into groups having specified sizes. For large group sizes, this distribution is approximated by a continuous distribution satisfying three exact moments.

The advantages and applications of MRPP are illustrated using both artificial examples and empirical data on total August standing crop in mixed prairie following an October wildfire. The MRPP analyses of the empirical data revealed that there were no differences in standing crop between burned and unburned areas after the first postfire growing season, but that after two growing seasons, standing crop was significantly greater in the previously burned areas.

Key words: distribution free; euclidian distance; fire; mixed prairie; multivariate; nonparametric; permutation techniques; statistical methods.

When results of statistical comparisons among experimental groups do not agree with a researcher's predictions, the results are accepted and the researcher's hypothesis rejected. Ideally, this prevents unfounded opinion from becoming reified. However, seldom is either the validity or the underlying structure of the statistical comparison more than superficially questioned. Unfortunately, some widely used statistical procedures, including the *t* test and the one-way analysis of variance *F* test, have been shown to operate in a geometry which is not in conformity with a sensible euclidian perception of the data space (Mielke et al. 1982, Mielke and Berry 1983). When such statistical comparisons do not agree with a researcher's visual perception of the experimental results, the statistical results should be questioned.

A newly developed technique for statistical comparisons among experimental groups overcomes this difficulty. This technique, termed multi-response permutation procedures (MRPP), also features additional advantages over previously used techniques: ease of application to multivariate problems and relatively few assumptional requirements on the data. In this paper, the basic concepts and advantages of MRPP are illustrated with simple artificial examples and the utility of MRPP is demonstrated on empirical data measuring the effects of fire on the phytomass in a mixed prairie.

[1] Manuscript received 28 February 1983; revised 10 November 1983; accepted 27 January 1984.

MULTI-RESPONSE PERMUTATION PROCEDURES

The purpose of MRPP is to detect concentration within a priori groups (a similar purpose to that of the *t* and *F* tests). However, the applicability of MRPP does not depend on assumptions such as that the data are from a normal population or that there are homogeneous variances under the alternative hypothesis. Rather, MRPP depend only on the internal variability of the existing data. Early versions of MRPP are given in Mielke et al. (1976), and descriptions that include recently developed modifications are given in Mielke et al. (1981*a, b*) and Mielke (1984).

To illustrate the computations involved, we consider an initial example involving two measurements (for example, biomass of two plant species x and y) taken on each object (quadrat) in two groups. Groups could be sites, treatments, etc. Let the first group consist of three objects and the second group of four, for a total

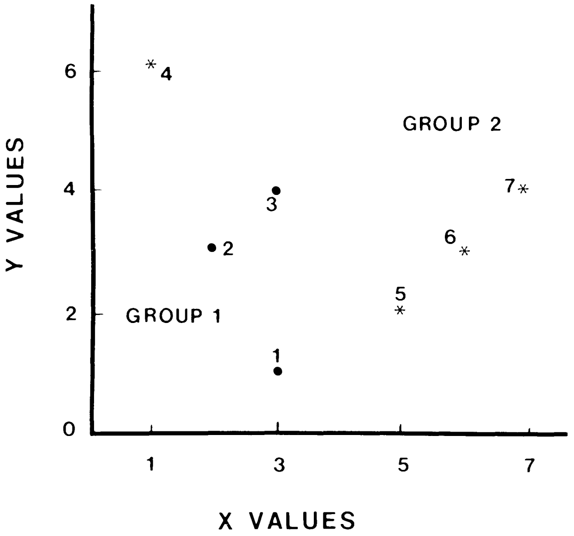

FIG. 1. Observed configuration of data in a two-group bivariate problem. The groups are distinguished by different symbols (* and ●).

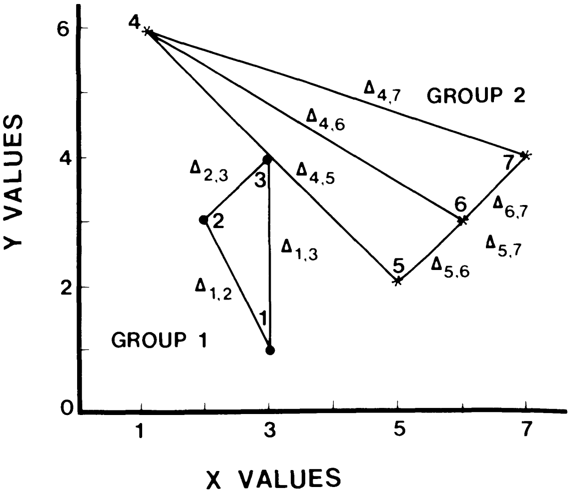

FIG. 2. Within-group euclidian distances in a two-group bivariate problem.

of seven objects. The biomass values define a space in which the objects can be visualized. An example of an observed configuration of the seven objects in a particular data space is shown in Fig. 1.

The basis of the MRPP statistic is the distance measure involving objects K and L defined as

$$\Delta_{K,L} = \left[\sum_{j=1}^{r} (X_{K,j} - X_{L,j})^2 \right]^{v/2},$$

where r is the number of measurements taken on the K^{th} object $(X_{K,1}, \ldots, X_{K,r})$ and $v > 0$. In our simple example, $r = 2$, and Fig. 2 shows the distances involved when the distance measure $v = 1$ (euclidian distance) is considered. Since it is just as easy to think of distances in a multivariate space as in a univariate (one-dimensional) space, the use of MRPP in multivariate problems is quite natural.

The choice of v determines the analysis space associated with the MRPP statistic. Thus the analyst must choose the distance measure carefully. Two common choices of the distance measure are based on $v = 1$, corresponding to the familiar metric known as euclidian distance, and $v = 2$, corresponding to a nonmetric (viz, squared euclidian distance). The presentation that follows includes both choices, as well as a discussion of issues concerning these choices.

For each of the groups, the average value of all the $\binom{n_i}{2}$ within-group distance measures for each pair of objects (indexed by K and L) in the i^{th} group is defined as

$$\xi_i = \binom{n_i}{2}^{-1} \sum_{K<L} \Delta_{K,L},$$

where n_i is the number of objects in the i^{th} group. Thus, for our example with $n_1 = 3$ and $n_2 = 4$,

$$\xi_1 = (\Delta_{12} + \Delta_{13} + \Delta_{23})/3,$$
$$\xi_2 = (\Delta_{45} + \Delta_{46} + \Delta_{47} + \Delta_{56} + \Delta_{57} + \Delta_{67})/6.$$

From the data of Fig. 1, $\xi_1 = 2.217$ and $\xi_2 = 3.913$ when $v = 1$, and $\xi_1 = 5.333$ and $\xi_2 = 19.667$ when $v = 2$.

The MRPP statistic is then defined as

$$\delta = \sum_{i=1}^{g} C_i \xi_i,$$

where C_i is a positive constant associated with the i^{th} of g groups, and $\sum_{i=1}^{g} C_i = 1$. Therefore, δ is a linear combination of average within-group distance measures for the g groups. A relatively small value of δ suggests a concentration of the object measurements within the groups. To insure that group comparisons are efficient (Mielke 1984), the choice for C_i is n_i/N for $i = 1, 2, \ldots, g$, where $N = \sum_{i=1}^{g} n_i$ is the total number of objects in the combined sample (not an unknown population size). This makes δ the weighted average of within-group pairwise distance measure means. In our example, $g = 2$, $N = 7$, and $\delta = {}^3\!/_7\, \xi_1 + {}^4\!/_7\, \xi_2$.

Thus, $\delta = 3.186$ for $v = 1$, and $\delta = 13.52$ for $v = 2$. A more general test for situations in which all the objects cannot be classified into g groups has also been considered (Mielke et al. 1981*a, b*).

It has been shown that the two-sample t test statistic and the one-way analysis of variance F test statistic are functionally identical to an MRPP statistic when $v = 2$ and $C_i = (n_i - 1)/(N - g)$ for $i = 1, 2, \ldots, g$ (Mielke et al. 1982). Similarly, the Mann-Whitney-Wilcoxon test and other nonparametric tests are also

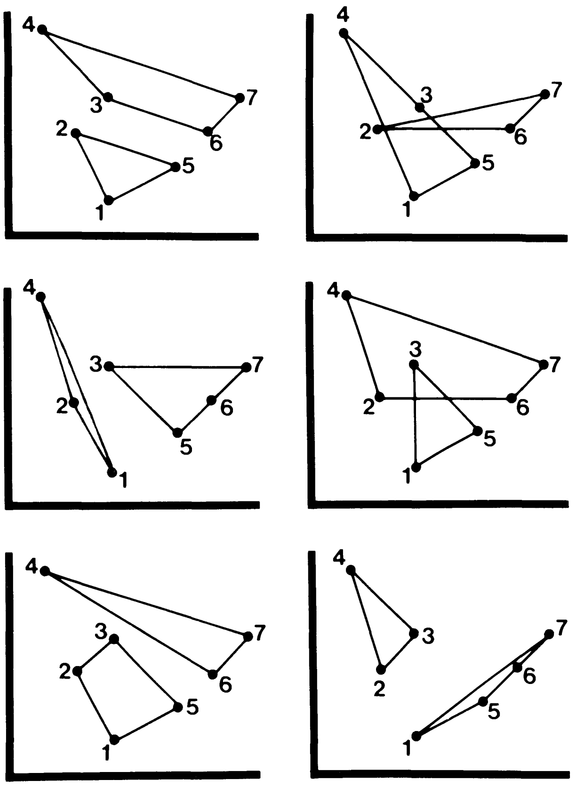

FIG. 3. Six of the 35 possible permutations of the data into the two groups.

functionally identical to an MRPP when $v = 2$, the rank transformations are used, and again $C_i = (n_i - 1)/(N - g)$ for $i = 1, 2, \ldots, g$ (Mielke et al. 1982). An MRPP analog for randomized blocks has been developed (Mielke and Iyer 1982), in which the permutation versions of the randomized block F test, the matched-pairs t test, and Pearson's correlation coefficient (and their nonparametric counterparts) are special cases when $v = 2$. However, if $v = 2$, the distance measure is nonmetric because the triangle inequality of a metric space does not hold (Mielke and Berry 1982, Mielke and Iyer 1982, Mielke et al. 1982).

Since the data space (i.e., the visualized collection of measurements in question) is viewed by most investigators as a euclidian space, the data space and the MRPP analysis space are congruent only when $v = 1$ (Mielke and Berry 1983). It should not be surprising, therefore, that there may be drastic differences between the statistical comparisons based on MRPP with $v = 2$, and an investigator's intuitive euclidian space comparisons. Thus the choice of an MRPP statistic based on $v = 1$ rather than one with $v = 2$ is compelling.

The P value of an MRPP statistic is derived through a permutation argument; thus, there are no distributional requirements on the data. A permutation is a specific arrangement or assignment of all N objects to

TABLE 1. The first 10 of the 35 possible values of the test statistic δ for the data illustrated in Fig. 1. The complete list of all 35 ordered values constitutes the null distribution. The statistic corresponding to the permutation actually observed (permutation number 1) is identified by an asterisk. Note that the P value associated with permutation 1 depends on whether $v = 1$ or $v = 2$.

$v = 1$			$v = 2$		
δ	Permutation number	P	δ	Permutation number	P
2.525	31	1/35	7.414	31	1/35
2.629	20	2/35	8.096	20	2/35
3.030	2	3/35	10.660	2	3/35
3.186*	1	4/35	11.269	35	4/35
3.211	8	5/35	12.381	19	5/35
3.246	19	6/35	12.570	8	6/35
3.305	35	7/35	12.667	14	7/35
3.308	14	8/35	13.048	9	8/35
3.335	3	9/35	13.333	3	9/35
3.356	34	10/35	13.520*	1	10/35

the specified groups. The null hypothesis for MRPP states that all $N!/\Pi_{i=1}^{g} n_i!$ permutations are equally likely. The ordered list of computed values of the test statistic for all permutations (the null distribution) depends strictly on the available data and thus avoids any extraneous assumptions. The P value for an observed statistic is obtained by finding the statistic's relative position on the ordered list (i.e., finding the proportion of all possible values of δ that are less than or equal to the observed value of δ).

In our example, $N = 7$, $n_1 = 3$, and $n_2 = 4$, so that there are 35 permutations. Six of these 35 possible permutations are shown in Fig. 3. Table 1 shows the first 10 of the 35 ordered values of δ for both the MRPP statistic for $v = 1$ and that for $v = 2$. Note that the rank of the observed value corresponding to permutation number 1 (illustrated in Fig. 1) is dependent on the distance measure used. Here $P = .114$ for $v = 1$, and $P = .286$ for $v = 2$. While the results for $v = 1$ and $v = 2$ in this example do not differ dramatically, this is not always so. A second (admittedly pathologic) example with $n_1 = 35$, $n_2 = 35$, and $N = 70$ yields $P = .00001$ for $v = 1$, and $P = 1.0$ for $v = 2$ (corresponding to the permutation version of a two-sided t test). The data for this second example are illustrated in Fig. 4.

The P value can be evaluated through exact enumeration if the number of possible permutations is not large (Berry 1982). However, even in moderately sized experiments, the number of permutations is so large that a continuous probability density function must be used to approximate the discrete null distribution. The standardized test statistic is defined as $T = (\delta - \mu_\delta)/\sigma_\delta$, where μ_δ and σ_δ are, respectively, the mean and standard deviation of δ under the null hypothesis. Since the distribution of δ is often highly skewed in the negative direction, a Pearson type III distribution (Mielke et al. 1981a) that accounts for the parameters μ_δ, σ_δ, and γ_δ

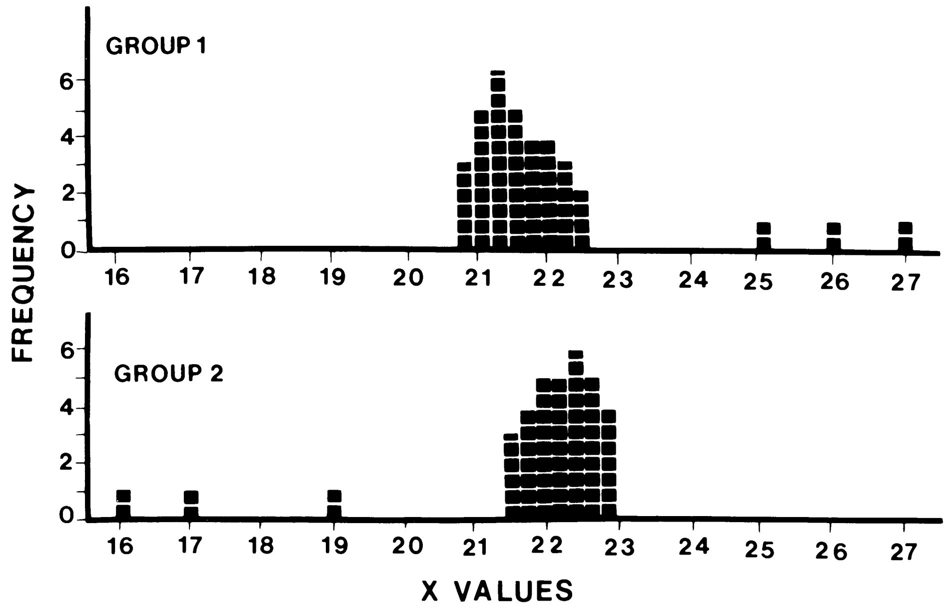

FIG. 4. Data in a univariate two-group experiment ($n_1 = 35$, $n_2 = 35$) illustrating the discrepancy between the two distance measures based on $v = 1$ and $v = 2$, respectively. If v is chosen to equal 1, $P = .00001$, while for $v = 2$, $P = 1.0$.

(the skewness of δ under the null hypothesis) is used. Efficient computer algorithms exist for finding the exact values of μ_δ, σ_δ, and γ_δ (Berry and Mielke 1983). The tail areas (P values) associated with an observed value of T, T_o, are then evaluated through numerical integration:

$$P(T < T_o) = \int_{-\infty}^{T_o} f(y)\, dy,$$

where $f(y)$ is the Pearson type III distribution. This approximation appears to be excellent (Mielke and Berry 1982, Mielke et al. 1982). A complete computer algorithm for MRPP is given by Berry and Mielke (1983).

APPLICATION

The data for this application were taken from a fire study in southwestern North Dakota (Zimmerman 1981) and consist of measures of total standing crop (grams per square metre) within 0.1-m^2 quadrats randomly placed in the area surrounding four 10–30 m long transects. The transects spanned unburned areas as well as areas burned in an October 1976 wildfire that swept across ≈2400 ha. Data are from the mid-August harvest period of the first and second postfire growing seasons (1977 and 1978).

The study site was on land administered by the United States Forest Service, in Slope County, North Dakota, 11 km from Medora and 17 km from Amidon. The climate of the area is characterized by warm summers and long, cold winters. The annual mean temperature is 5°C, with a July mean of 21° and a January mean of −10°. The frost-free season generally extends from early May to late August. Annual precipitation averages 40 cm, but is highly variable. The months preceding the fire were fairly dry, with total precipitation for several months in the lower percentiles of monthly totals for the years on record. However, the seasons following the fire were somewhat wet, with monthly precipitation totals during fall 1977 and spring 1978 generally in the upper percentiles.

Transects 1 and 2 were located in areas characterized by *Agropyron smithii, Bouteloua gracilis, Carex filifolia,* and *Stipa comata.* Transects 3 and 4 were in locally abundant patches of *Schyzachrium scoparium,* which also featured *Echinacea angustifolia* and *Helianthus* spp. (Nomenclature follows Kartesz and Kartesz 1980). Each transect included both treatments (burned and unburned). Thus there is a two-way classification of the quadrats into eight transect-by-treatment groups. Three quadrats were sampled in each such group in 1977 ($N = 24$) and five were sampled in each group in 1978 ($N = 40$).

For comparative purposes, two sets of MRPP analyses, one for $v = 1$ and one for $v = 2$, are presented. This application demonstrates that results for $v = 1$ and those for $v = 2$ are not always drastically different.

In the analyses of the 1977 data, we first looked for concentration of the quadrat values into two groups representing treatments ($g = 2$, $n_1 = n_2 = 12$, $N = 24$, $r = 1$). The MRPP analyses yielded $P = .88$ ($v = 1$) and $P = .74$ ($v = 2$). Such high P values may simply reflect a lack of influence of treatment, or may be caused by differences between the transects in the direction of response to the treatment. We thus tested for concentration within eight groups representing the treatment-

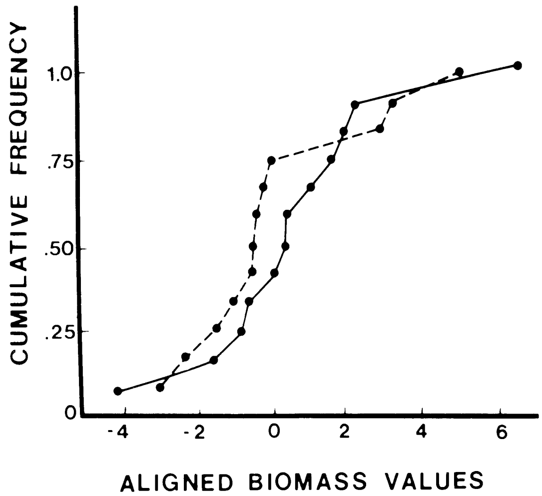

FIG. 5. Frequency ogives for aligned values from burned and unburned areas for August 1977 sampling (—— burned, - - - unburned).

by-transect classification ($g = 8$, $n_1 = \ldots = n_8 = 3$, $N = 24$, $r = 1$). In this case, the MRPP analyses yielded $P = .0017$ ($v = 1$) and $P = .0024$ ($v = 2$), indicating a meaningful grouping. However, before declaring that there were differential treatment responses, we tested for concentration within four groups representing the four transects ($g = 4$, $n_1 = \ldots = n_4 = 6$, $N = 24$, $r = 1$). These MRPP analyses yielded $P = .0017$ ($v = 1$) and $P = .0018$ ($v = 2$), showing substantial differences among the transects. The similarity between the MRPP results for the four-group and the eight-group analyses suggests that the only meaningful influence was that of transect.

For the purpose of discerning treatment influences, the differences among transects are not of direct interest. As a consequence, transects may be treated similarly to blocks in a randomized block experiment. Mielke and Iyer (1982) present methods for performing MRPP analyses on randomized block experiments, including a method known as "align and collapse" for situations involving more than one object in each treatment group in each block. This method involves removing the block differences by (1) subtracting the appropriate block median from each value, and (2)

TABLE 2. Group averages for total standing crop (g/m^2) in mixed prairie in August 1977, one growing season after a fire. Entries are means of values for three 0.1-m^2 quadrats.

Treatment	Transect 1	2	3	4	Average
Burned	3.9	6.1	11.0	7.7	7.2
Unburned	5.5	5.5	12.3	3.6	6.7
Average	4.7	5.8	11.6	5.6	6.9

TABLE 3. Group averages for total standing crop (g/m^2) in mixed prairie in August 1978, two growing seasons after a fire. Entries are means of values for five 0.1-m^2 quadrats.

Treatment	Transect 1	2	3	4	Average
Burned	25.5	20.2	22.6	23.8	23.0
Unburned	14.5	23.2	14.6	15.5	17.0
Average	20.0	21.7	18.6	19.7	20.0

analyzing the entire set of aligned values (i.e., collapsing the block structure) using the usual MRPP. Accordingly, from each value we subtracted the median of the combined burned and unburned groups comprising each transect. The MRPP analyses of the resulting aligned values showed no concentration within treatment groups ($g = 2$, $n_1 = n_2 = 12$, $N = 24$, $r = 1$): $P = .56$ ($v = 1$) and $P = .67$ ($v = 2$).

Group averages are shown in Table 2 to summarize the results under $v = 2$ for the 1977 data, since this choice of distance function corresponds to the analysis of variance. Such a summary is not appropriate for tests based on $v = 1$, which test concentrations rather than group average differences. Instead, the data are summarized by frequency ogives for each of the two treatment groups (Fig. 5); these ogives show the lack of separation between the aligned values in the two treatments.

For the 1978 data (Table 3, Fig. 6) the MRPP analyses indicated concentration within treatment groups ($g = 2$, $n_1 = n_2 = 20$, $N = 40$, $r = 1$): $P = .016$ ($v = 1$) and $P = .012$ ($v = 2$). The alignment method was not used because the test for transect grouping ($g = 4$, $n_1 = \ldots = n_4 = 10$, $N = 40$, $r = 1$) showed no dif-

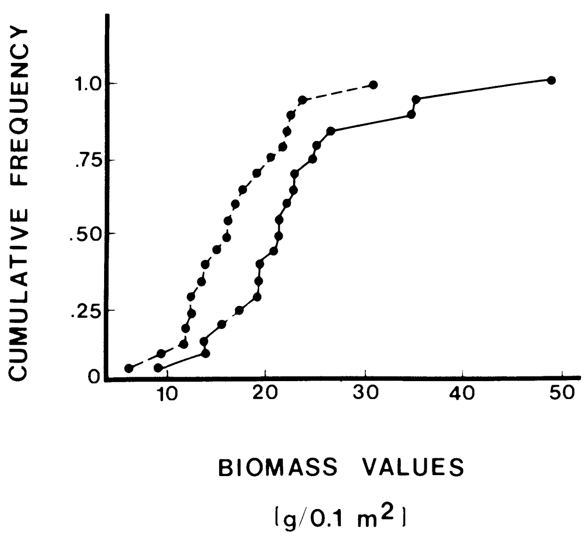

FIG. 6. Frequency ogives for burned and unburned areas for August 1978 sampling (—— burned, - - - unburned).

ferential treatment responses: $P = 44$ ($v = 1$) and $P = .88$ ($v = 2$).

In summary, MRPP analyses indicate that standing crop (phytomass) in previously burned and unburned areas did not differ significantly in 1977, one growing season after the fire, but that in 1978, after two growing seasons, standing crop was significantly greater on the previously burned areas.

Acknowledgments

We thank Carolyn Rajewski for assistance in field data collection and preparation of the figures. This research has been sponsored by a cooperative agreement between the North Dakota Agricultural Experiment Station, North Dakota Game and Fish Department, and the United States Department of Agriculture Forest Service, and has been supported in part by National Science Foundation Grant ATM81-07056. We also appreciate constructive comments made by the reviewers of this paper.

Literature Cited

Berry, K. J. 1982. Algorithm AS 179: enumeration of all permutations of multi-sets with fixed repetition numbers. Applied Statistics **31**:169–173.

Berry, K. J., and P. W. Mielke. 1983. Computation of finite population parameters and approximate probability values for multi-response permutation procedures (MRPP). Communications in Statistics-Simulation and Computation **B12**: 83–107.

Kartesz, J., and R. Kartesz. 1980. A synonymized checklist of the vascular flora of United States, Canada and Greenland. Volume II. The biota of North America. University of North Carolina Press, Chapel Hill, North Carolina, USA.

Mielke, P. W. 1984. Meteorological applications of permutation techniques based on distance functions. Pages 813–830 *in* P. R. Krishnaiah and P. K. Sen, editors. Handbook of statistics, Volume 4: Nonparametric methods. North-Holland Publishing, Amsterdam, The Netherlands.

Mielke, P. W., and K. J. Berry. 1982. An extended class of permutation techniques for matched pairs. Communications in Statistics-Theory and Methods **A11**:1197–1207.

Mielke, P. W., and K. J. Berry. 1983. Asymptotic clarifications, generalizations, and concerns regarding an extended class of matched pairs tests based on powers of ranks. Psychometrika **48**:483–485.

Mielke, P. W., K. J. Berry, and G. W. Brier. 1981*a*. Application of multiresponse permutation procedures for examining seasonal changes in monthly mean sea-level pressure patterns. Monthly Weather Review **109**:120–126.

Mielke, P. W., K. J. Berry, P. J. Brockwell, and J. S. Williams. 1981*b*. A class of nonparametric tests based on multiresponse permutation procedures. Biometrika **68**:720–724.

Mielke, P. W., K. J. Berry, and E. S. Johnson. 1976. Multiresponse permutation procedures for a priori classifications. Communications in Statistics-Theory and Methods **A5**: 1409–1424.

Mielke, P. W., K. J. Berry, and J. G. Medina. 1982. Climax I and II: distortion resistant residual analyses. Journal of Applied Meteorology **21**:788–792.

Mielke, P. W., and H. K. Iyer. 1982. Permutation techniques for analyzing multi-response data from randomized block experiments. Communications in Statistics-Theory and Methods **A11**:1427–1437.

Zimmerman, G. M. 1981. Effects of fire upon selected plant communities in the Little Missouri Badlands. Thesis. North Dakota State University, Fargo, North Dakota, USA.