Application of Multi-Response Permutation Procedures for Examining Seasonal Changes in Monthly Mean Sea-Level Pressure Patterns

PAUL W. MIELKE, JR.,[1] KENNETH J. BERRY[2] AND GLENN W. BRIER[3]

Colorado State University, Fort Collins 80523

(Manuscript received 9 May 1980, in final form 25 July 1980)

ABSTRACT

This paper considers the examination of possible differences in monthly sea-level pressure patterns. The satisfactory examination of such differences requires appropriate multi-response parametric methods based on unknown multivariate distributions (i.e., an appropriate parametric technique is probably nonexistent). In order to avoid the likely insurmountable difficulties involving parametric methods, the application of multi-response permutation procedures (MRPP) is suggested as an appropriate approach for the examination of such differences.

1. Introduction

The purpose of this paper is to both describe and motivate the use of a recently developed class of statistical techniques termed multi-response permutation procedures (MRPP).

The MRPP (Mielke *et al.*, 1976; Mielke, 1979) are used to compare *a priori* classified groups of objects where measurements of r responses ($r \geq 1$) are obtained from each object and each response measurement is at least ordinal (i.e., all object measurements corresponding to any singly specified response must at least be orderable from smallest to largest). For $r \geq 2$, distinct response measurements should be commensurate with one another (i.e., some of the response measurements may require appropriate scaling).

In contrast with well-known parametric statistical techniques such as univariate and multivariate analysis of variance (ANOVA and MANOVA) where individual object measurements are the primary analysis units, normed distances (e.g., Euclidean distances) between the r-dimensional points (associated with the r measurements of each object) are the primary analysis units of MRPP. Since MRPP yield inferential results which are solely dependent on a realized data set in question, they avoid certain difficult problems inherent in parametric techniques. Specifically avoided are 1) the initial choice of a univariate or multivariate distribution that is both reasonable (provides a good fit) and tractable and 2) the selection (choice of technique) and analytic development of a parametric inference technique based on a *chosen* univariate or multivariate distribution.

The intuitive concept of MRPP is initially introduced and motivated with a simple illustrative example of MRPP in Section 2. A more complete technical description of MRPP is then presented in Section 3 which includes the essential large sample approximations for statistical inferences. Examples illustrating the application of MRPP with real data are then given in Section 4. The first example involves a comparison between September and October mean sea-level pressures. This comparison is based on indices derived from four key areas which (based on a previous investigation) showed the largest average differences in mean sea-level pressure between September and October. The second and third examples (September and October considered separately) use these same indices to compare mean sea-level pressures for years when the winter season (Northern Hemisphere) tropical stratospheric winds are westerly with mean sea-level pressures for immediately following years when the corresponding winds are easterly (known as the quasi-biennial oscillation).

2. Illustrative example of MRPP

The general concepts behind the MRPP can best be illustrated by considering a comparison between two mutually exclusive subgroups of objects (A and B) where two measured responses (x_1 and x_2) have been obtained from each object in the two subgroups. Fig. 1 shows how these responses could be represented in a two-dimensional diagram where the responses of the three objects in subgroup A are plotted

[1] Department of Statistics.
[2] Department of Sociology.
[3] Department of Atmospheric Science.

0027-0644/81/010120-07$05.75

as crosses and the responses of the four objects in subgroup B are plotted as circles. Although a visual impression suggests that the subgroups A and B are separated, a more rigorous and objective characterization of this separation is needed before a quantitative evaluation or inference can be made. A classical approach would involve using the two-sample Hotelling T^2 test which has the disadvantage of requiring the assumption that the response measurements of the two subgroups are distributed as the multivariate normal distribution with equal variances and covariances. Since these conditions are *never* met in practice, it is desirable to consider alternative procedures.

One way of doing this is by first examining the distances

$$\Delta_{I,J} = \|\mathbf{x}_I - \mathbf{x}_J\| = [\sum_{i=1}^{2} (x_{iI} - x_{iJ})^2]^{1/2}$$

between *all* distinct pairs of points in the diagram. The seven points of Fig. 1 imply there are $\binom{7}{2} = 21$ distinct pairs of points and consequently 21 distances must be computed. These 21 distances are listed in Table 1 and ordered from the lowest to highest value. Table 1 confirms the visual impression of clustering since the distances between points of a *common* subgroup tend to be smaller than the distances between points of *different* subgroups. A natural way of considering this clustering tendency is by forming the average of the between-point distances for each subgroup. Thus, for the three distances of subgroup A, the average is

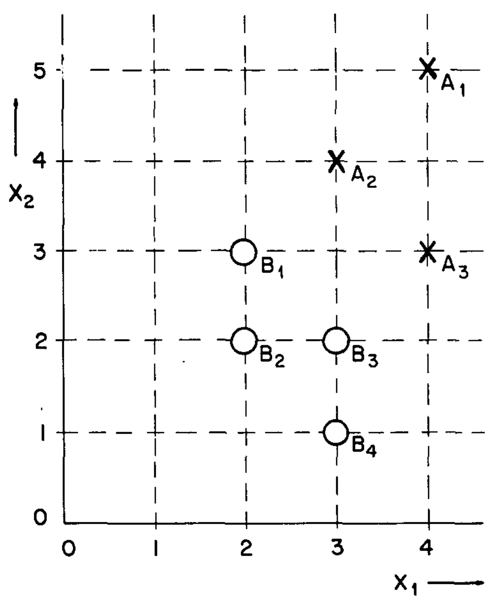

FIG. 1. Scatter diagram showing the points of the two subgroups (A and B) plotted as a function of the measured responses x_1 and x_2.

TABLE 1. Ordered distances between all 21 pairs of the seven points shown in Fig. 1 where distances between points in either subgroup A or subgroup B are indicated by crosses or circles, respectively.

Rank	Points	Distance
1	B_1B_2	1.000 (○)
2	B_2B_3	1.000 (○)
3	B_3B_4	1.000 (○)
4	A_1A_2	1.414 (×)
5	A_2A_3	1.414 (×)
6	A_2B_1	1.414
7	A_3B_3	1.414
8	B_1B_3	1.414 (○)
9	B_2B_4	1.414 (○)
10	A_1A_3	2.000 (×)
11	A_2B_3	2.000
12	A_3B_1	2.000
13	A_2B_2	2.236
14	A_3B_2	2.236
15	A_3B_4	2.236
16	B_1B_4	2.236 (○)
17	A_1B_1	2.828
18	A_2B_4	3.000
19	A_1B_3	3.162
20	A_1B_2	3.606
21	A_1B_4	4.123

$$\xi_A = (1/3) \sum_{A} \Delta_{I,J} = 1.6095$$

and for the six distances of subgroup B the average is

$$\xi_B = (1/6) \sum_{B} \Delta_{I,J} = 1.3441.$$

A measure or statistic which describes the separation between the points of subgroups A and B is the simple weighted mean given by

$$\delta = (3/7)\xi_A + (4/7)\xi_B = 1.4578.$$

Smaller values of δ would indicate a tendency for clustering while larger values of δ would indicate a lack of clustering. The problem is to determine whether the observed statistic ($\delta = 1.4578$) for this particular partition (A and B) is unusual with respect to other possible partitions with the same size structure that could have been made with these seven objects. Now N objects can be partitioned into two subgroups A and B with fixed numbers of points n_A and n_B, respectively, in precisely

$$M = N!/(n_A!n_B!)$$

ways. Since $M = 35$ for this example, 35 values of δ can be obtained by enumerating all the possible 35 partitions. These 35 values of δ are listed in Table 2 and ordered from the lowest to highest value. We see that the observed statistic ($\delta = 1.4578$) obtained for the realized partition (A and B) is indeed unusual since each of the remaining 34 values is greater. If all partitions could have occurred with equal chance (the null hypothesis), then the observed significance

TABLE 2. Ordered values of δ for all 35 partitions of the seven points shown in Fig. 1 into two subgroups (A and B) having fixed sizes $n_A = 3$ and $n_B = 4$.

Rank	Value	Rank	Value
1	1.4578	19	2.1381
2	1.5421	20	2.1480
3	1.6939	21	2.1591
4	1.7505	22	2.1646
5	1.8389	23	2.1709
6	1.8547	24	2.1740
7	1.8935	25	2.1769
8	1.9898	26	2.1891
9	1.9915	27	2.1939
10	1.9988	28	2.2025
11	2.0060	29	2.2169
12	2.0157	30	2.2258
13	2.0176	31	2.2280
14	2.0522	32	2.2470
15	2.0575	33	2.2518
16	2.0829	34	2.2812
17	2.0944	35	2.2935
18	2.1158		

level or P value is $1/35 = 0.0286$. Thus, we would accept the realized partition (A and B) as being significant at the 5% level of significance.

When M is large (e.g., $M = 1.55 \times 10^8$ when $N = 30$ and $n_A = n_B = 15$), it is obviously impractical to generate the discrete probability distribution of δ illustrated in Table 2. It is then necessary to approximate the distribution of δ by a continuous distribution in order to determine whether a P value for an observed value of δ is smaller than some prescribed level of significance. The following discussion in Section 3 describes how this can be done for a more general version of MRPP.

3. Technical description of MRPP

Following the notation of Mielke *et al.* (1976) and Mielke (1979), let

$$\Omega = \{\omega_1, \ldots, \omega_N\}$$

be a finite population of N objects, let

$$\mathbf{x}_I' = [x_{1I}, \ldots, x_{rI}]$$

denote r commensurate response measurements (and also could involve functions of response measurements or residuals adjusted by predictors) for object ω_I ($I = 1, \ldots, N$), and $S_1, \ldots, S_{g+1}$ represent an exhaustive partitioning of the N objects comprising Ω into $g + 1$ disjoint subgroups. In addition, let

$$\Delta_{I,J} = \|\mathbf{x}_I - \mathbf{x}_J\|$$

designate the normed distance of response measurements between objects ω_I and ω_J. The normed distance used in the subsequent examples of this paper will be the Euclidean distance given by

$$\Delta_{I,J} = [\sum_{k=1}^{r} (x_{kI} - x_{kJ})^2]^{1/2}.$$

The MRPP are based on the statistic given by

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

where $n_i \geq 2$ is the number of (classified) objects in subgroup S_i ($i = 1, \ldots, g$), $K = \sum_{i=1}^{g} n_i$, $n_{g+1} = N - K$ is the number of remaining (unclassified) objects in subgroup S_{g+1} (which is an empty subgroup in many applications), $C_i > 0$ ($i = 1, \ldots, g$), $\sum_{i=1}^{g} C_i = 1$, $C_i = n_i/K$ (cf. Mielke, 1979) is the choice of C_i in this paper [note that Mielke *et al.* (1976) used $C_i = \binom{n_i}{2}/\sum_{j=1}^{g}\binom{n_j}{2}$, an inefficient choice of C_i for detecting location shifts when $n_1, \ldots, n_g$ are not all equal, $g \geq 2$ and $K = N$],

$$\xi_i = \binom{n_i}{2}^{-1} \sum_{I<J} \Delta_{I,J} I_{S_i}(\omega_I) I_{S_i}(\omega_J)$$

is the average between-object distance for all objects within subgroup S_i ($i = 1, \ldots, g$), $\sum_{I<J}$ is the sum over all I and J such that $1 \leq I < J \leq N$, and $I_{S_i}(\omega_I)$ is 1 if ω_I belongs to S_i and zero otherwise for $I = 1, \ldots, N$. The underlying permutation distribution of δ implies that each of the

$$M = N!/(\prod_{i=1}^{g+1} n_i!)$$

possible allocation combinations of the N objects to the $g + 1$ subgroups will have the same chance of occurring. With the underlying distribution of δ assumed, let μ_δ, σ_δ^2 and γ_δ, denote the mean, variance and skewness of δ, respectively. If δ_j denotes the jth value among the M possible values of δ, then

$$\left.\begin{aligned} \mu_\delta = M^{-1} \sum_{j=1}^{M} \delta_j, \quad \sigma_\delta^2 = M^{-1} \sum_{j=1}^{M} \delta_j^2 - \mu_\delta^2 \\ \gamma_\delta = (M^{-1} \sum_{j=1}^{M} \delta_j^3 - 3\mu_\delta\sigma_\delta^2 - \mu_\delta^3)/\sigma_\delta^3 \end{aligned}\right\}.$$

Efficient computational techniques for obtaining μ_δ, δ_δ^2 and γ_δ for a realized set of data are described elsewhere (Mielke *et al.*, 1976; Mielke, 1979). An intuitive interpretation of MRPP results is simply that small values of δ imply a concentration of response measurements within the g subgroups. For the illustrative example of Section 2, we have $N = K = 7, r = 2, g = 2, n_1 = 3, n_2 = 4, \xi_1 = 1.6095$, $\xi_2 = 1.3441$, $\delta = 1.4578$, $\mu_\delta = 2.0547$, $\sigma_\delta^2 = 0.0396$ and $\gamma_\delta = -1.3523$.

Since the calculation of the M possible values of δ is seldom computationally feasible, an adequate approximation of the underlying permutation distribution of δ is essential for determining a P value (the probability of obtaining a value of δ which is not larger than a realized value of δ given the underlying permutation distribution of δ). The standardized test statistic given by

$$T = (\delta - \mu_\delta)/\sigma_\delta$$

is approximately distributed as the Pearson type III distribution. In particular, the Pearson type III distribution compensates for the fact that the underlying permutation distribution of δ is often very skewed in the negative direction (i.e., $\gamma_\delta < 0$). The P-value approximation for a realized value of δ (say, δ_0) is given by.

$$P[\delta \leq \delta_0] = \int_{-\infty}^{T_0} f(y)dy,$$

where

$$f(y) = \frac{(-2/\gamma)^{4/\gamma^2}}{\Gamma(4/\gamma^2)} [-(2 + y\gamma)/\gamma]^{(4-\gamma^2)/\gamma^2} e^{-2(2+y\gamma)/\gamma^2},$$

$-\infty < y < -2/\gamma$, $T_0 = (\delta_0 - \mu_\delta)/\sigma_\delta$ and $\gamma = \gamma_\delta \leq -0.001$ (a P value approximation based on the normal distribution is reported if $\gamma_\delta > -0.001$). The P value approximation based on the Pearson type III distribution is evaluated with Simpson's rule over the interval $(T_0 - 9, T_0)$. Incidentally, the asymptotic (as $N \rightarrow \infty$) distribution of δ (which may or may not be a normal distribution) is considered elsewhere (Mielke, 1979; O'Reilly and Mielke, 1980).

4. Examples and discussion

The following applications of MRPP involve sea-level pressure comparisons between groups of indices selected 1) on the basis of calendar month and 2) on the basis of the direction (easterly or westerly) of the tropical stratospheric 50 mb winds during the winter season. The four indices (X_1, X_2, X_3 and X_4) used as variables for this study were suggested in a previous investigation (Brier, 1948) as being effective in distinguishing between September and October mean sea-level pressure maps. This previous investigation was based on the Historical Daily Map Series (Joint Meteorological Committee) and showed four areas (corresponding to the four indices) in the Northern Hemisphere where September and October involved pronounced average differences in mean sea-level pressures. Brier (1948) also indicated that the maximum differences in mean sea-level pressures occurred in different areas for other seasons of the year (the present study is confined to September and October). Five geographical points represent each area and the specific locations of these points are listed in Table 3 and shown in Fig. 2 by crosses connected with straight lines. These four indices of circulation patterns were derived simply by summing the monthly mean sea-level pressure values for the five points of each area. The data of the present study are based on the years 1946–78 and thus are independent of the data used by Brier (1948) that prescribed the four areas. The observed values of the indices for the months of September and October during these 33 years are presented in Table 4.

TABLE 3. Location of points used in determining indices of circulation patterns.

Variable	Latitude	Longitude
X_1	55°N 60°N	0°, 10°W, 20°W 0°, 10°W
X_2	40°N 45°N, 50°N, 55°N, 60°N	140°E 130°E
X_3	40°N, 45°N, 50°N, 55°N, 60°N	150°W
X_4	35°N 40°N 45°N	60°E 60°E, 70°E 70°E, 80°E

The first comparison examines the possible difference between the months of September and October. Visual inspection of the data in Table 4 shows that X_4 by itself appears quite effective in separating these months. The advantage of MRPP is that it can take a number of variables into account and provide appropriate inferential statements. Examination of Fig. 3 suggests that X_1 and X_3 also should provide further information on separating the months, especially since X_1 and X_3 do not appear to be highly correlated. From the climatological point of view, the areas represented by X_1 and X_3 reflect the deepening and movement of the Aleutian and Icelandic low systems with the onset of the winter season. On the other hand, X_2 and X_4 represent areas with increasing pressure as the monsoonal circulation changes from summer to winter. The following summary describes the MRPP results for this first comparison (associate 1 and 2 with September and October, respectively): $N = K = 66$, $r = 4$, $g = 2$, $n_1 = n_2 = 33$, $\xi_1 = 40.5315$, $\xi_2 = 43.6078$, $\delta = 42.0696$, $\mu_\delta = 54.7535$, $\sigma_\delta^2 = 0.1697$, $\gamma_\delta = -1.8346$, $T = -30.7916$ and the P value is less than 10^{-14}.

The second comparison investigates the possible difference between eight September months classified as westerly (W) and eight October months classified as easterly (E), according to the previously defined W and E categories. The objective is to determine a possible association between the tropical stratospheric 50 mb winds and these particular measures (indices) of the sea-level circulation patterns. However, the sample size of only eight cases in each subgroup for this example is quite small compared

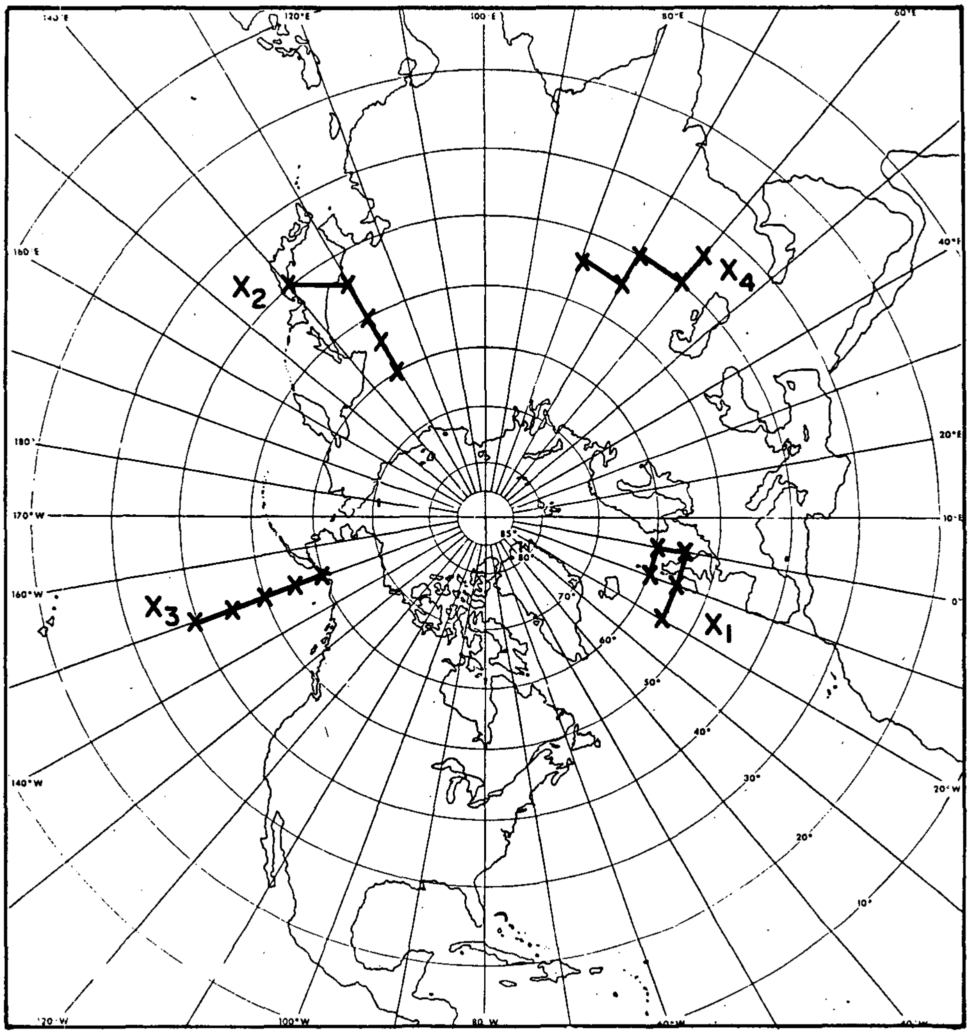

FIG. 2. Points selected for pressure indices shown by crosses and connected by straight lines.

with the sample size of 33 cases in each subgroup of the first comparison. The following summary describes the MRPP results for this second comparison (associate 1 and 2 with W and E, respectively): $N = K = 16, r = 4, g = 2, n_1 = n_2 = 8, \xi_1 = 41.4325$, $\xi_2 = 40.2511$, $\delta = 40.8418$, $\mu_\delta = 41.8533$, $\sigma_\delta^2 = 1.3297$, $\gamma_\delta = -1.2948$, $T = -0.8772$ and the P value is 0.1685. In addition, the third comparison is an analogous investigation of a possible difference between eight October months classified as W and eight October months classified as E. The following summary describes the MRPP results for this third comparison (again associate 1 and 2 with W and E, respectively): $N = K = 16$, $r = 4$, $g = 2$, $n_1 = n_2 = 8$, $\xi_1 = 42.8183$, $\xi_2 = 39.7836$, $\delta = 41.3010$, $\mu_\delta = 40.1692$, $\sigma_\delta^2 = 1.4366$, $\gamma_\delta = -1.6137$, $T = 0.9442$ and the P value is 0.8740.

The MRPP results of the first comparison imply (with very little doubt) that the indices of the type presently considered can be used as a classification tool to define seasonal characteristics of the sea-level circulation. In spite of large year-to-year variations, the circulation pattern for a particular month appears to have definite characteristics which distinguish it from adjacent months. For information on other months of the year, the reader should refer to the previous investigation by Brier (1948). In contrast, the MRPP results of the second and third comparisons indicate that these same indices are not successful in separating the W and E subgroups for either the month of September or October. This is not surprising since the sample size is small and it is known that the signal associated with the quasi-biennial oscillation in tropospheric weather data is much less pronounced than the corresponding signal in the tropical stratosphere.

The applications of MRPP in this paper were used to investigate associations among indices of sea-

TABLE 4. Pressure indices (mb) used for MRPP comparisons (years labeled W or E indicate that winter season 50 mb winds at Balboa were westerly or easterly, respectively).

	X_1		X_2		X_3		X_4	
Year	Sep	Oct	Sep	Oct	Sep	Oct	Sep	Oct
1946	5025.9	5090.7	5053.8	5088.5	5056.9	5074.9	5054.6	5082.2
1947	5064.3	5086.2	5067.9	5101.0	5080.5	5003.8	5056.4	5097.9
1948	5041.7	5048.7	5063.1	5089.5	5072.5	5025.1	5062.7	5099.4
1949	5074.7	5035.5	5062.8	5091.4	5057.5	5052.5	5064.8	5112.9
1950	5010.1	5040.4	5063.1	5087.1	5053.4	5024.4	5064.5	5093.9
1951	5035.7	5070.4	5068.6	5081.8	5053.8	5064.0	5064.9	5088.0
1952	5072.4	5026.4	5060.3	5081.8	5060.8	5040.2	5072.4	5101.2
1953	5040.6	5058.5	5061.7	5087.8	5067.5	5015.1	5058.2	5109.2
1954	5018.3	5019.4	5056.5	5098.3	5091.4	5023.0	5058.9	5090.2
1955W	5045.0	5060.4	5061.3	5087.7	5078.3	5033.4	5064.5	5104.9
1956E	5056.7	5072.1	5058.2	5085.0	5093.1	5061.5	5065.4	5102.3
1957W	5049.7	5040.9	5061.9	5074.9	5049.2	5050.0	5071.2	5109.4
1958E	5053.4	5059.0	5066.0	5098.6	5073.5	5046.6	5062.9	5106.8
1959W	5092.1	5042.8	5063.6	5091.9	5064.4	5042.6	5045.1	5095.3
1960E	5060.7	5038.2	5058.5	5102.6	5075.7	5025.6	5066.9	5097.6
1961	5031.0	5012.6	5053.8	5095.4	5096.8	5069.4	5052.3	5115.0
1962	5047.8	5070.9	5062.4	5098.6	5060.6	5028.4	5074.9	5108.4
1963	5060.8	5038.2	5052.2	5090.8	5032.3	5012.1	5066.2	5098.5
1964W	5053.5	5060.3	5070.1	5094.1	5079.2	5039.9	5071.9	5117.1
1965E	5038.8	5071.4	5057.3	5090.5	5103.3	5015.4	5076.0	5099.8
1966	5060.0	5044.7	5063.2	5086.1	5035.8	5039.7	5069.4	5114.0
1967W	5037.3	4986.7	5072.6	5088.1	5016.1	5024.0	5066.0	5110.3
1968E	5039.4	5031.5	5076.1	5098.8	5061.9	5030.7	5073.2	5105.4
1969W	5072.3	5060.6	5064.7	5087.1	5055.9	5027.3	5075.3	5094.5
1970E	5039.3	5059.4	5058.9	5087.9	5067.5	5060.4	5065.6	5098.7
1971W	5090.2	5066.9	5072.8	5086.8	5079.8	5064.9	5073.9	5101.8
1972E	5106.0	5080.8	5073.9	5082.4	5091.3	5077.2	5076.9	5105.0
1973W	5065.1	5089.3	5070.5	5092.8	5063.6	5053.3	5079.1	5105.7
1974E	5016.8	5083.9	5058.0	5077.8	5062.2	5033.1	5072.0	5109.9
1975	5037.8	5071.9	5067.1	5095.5	5080.6	5048.3	5064.7	5099.3
1976	5078.2	5003.9	5073.4	5082.5	5039.4	5047.2	5067.6	5098.0
1977	5080.5	5017.9	5061.7	5095.5	5090.5	5015.8	5071.3	5105.9
1978	5053.4	5078.9	5058.1	5094.0	5058.5	5042.5	5058.6	5104.4

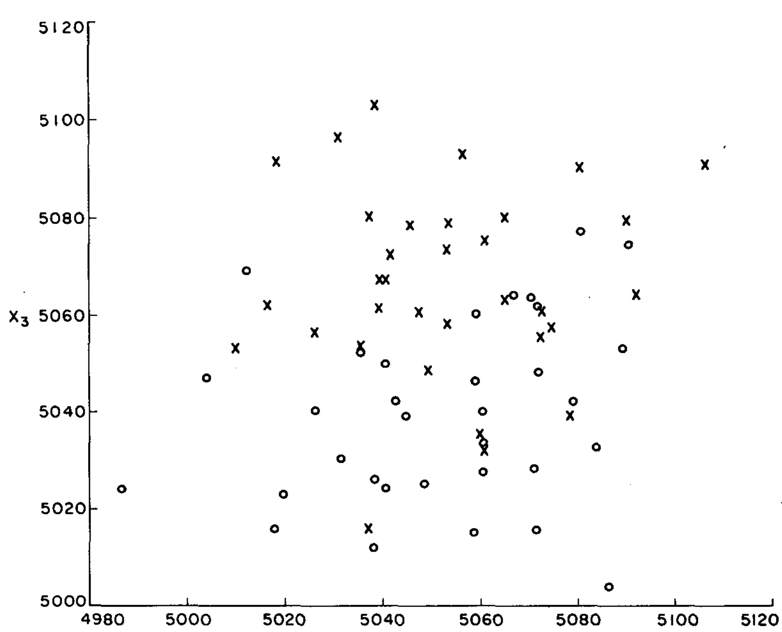

FIG. 3. Scatter diagram showing September (crosses) and October (circles) plotted as a function of X_1 and X_3.

level circulation patterns. However, the potential applications of MRPP for investigating a broad variety of univariate and multivariate comparisons in the atmospheric sciences (e.g., comparative investigations associated with climatology, tropical storms, and weather modification) are seemingly unlimited. Incidentally, an examination of practical nonparametric references such as Hollander and Wolfe (1973) and Mosteller and Rourke (1973) will indicate the relation between MRPP and the more well-known nonparametric statistical methods.

Acknowledgments. This work has been supported by the Office of Atmospheric Resources Research, Water and Power Resources Service, U.S. Department of the Interior, under Contract 8-07-83-V0009. The authors are also indebted to a reviewer for very helpful suggestions which improved the clarity of this paper.

REFERENCES

Brier, G. W., 1948: Changes in the Northern Hemisphere sea-level circulation in the 40-year period 1899–1939. *Trans. Amer. Geophys. Union*, **29**, 789–795.

Hollander, M., and D. A. Wolfe, 1973: *Nonparametric Statistical Methods*. Wiley, 503 pp.

Mielke, P. W., 1979: On asymptotic non-normality of null distributions of MRPP statistics. *Commun. Statist. Theor. Meth.*, **A8**, 1541–1550.

——, K. J. Berry and E. S. Johnson, 1976: Multi-response permutation procedures for *a priori* classifications. *Commun. Statist. Theor. Meth.*, **A5**, 1409–1424.

Mosteller, F., and R. E. K. Rourke, 1973: *Study Statistics*. Addison-Wesley, 395 pp.

O'Reilly, F. J., and P. W. Mielke, 1980: Asymptotic normality of MRPP statistics from invariance principles of U-statistics. *Commun. Statist. Theor. Meth.*, **A9**, 629–637.