

## Mantel Test

### Background

The Mantel Test is used to test the null hypothesis of no relationship between two square symmetrical matrices. The Mantel test was originally developed to evaluate spatial and temporal clusterings of diseases, such as leukemia (Mantel 1967). Sokal (1979) introduced the technique to systematics and quantitative biogeography. Douglas and Endler (1982) gave the mathematical details in a palatable form.

The Mantel test evaluates correlation between distance (or similarity or correlation or dissimilarity) matrices. Normally each matrix is calculated from a different set of variables, measured for the same sample units. The Mantel test is an alternative to regressing one matrix against the other, circumventing the problem of partial dependence within each matrix. Because the cells of distance matrices are not independent of each other, we cannot accept the  $p$ -values from standard techniques that assume independence of the observations (for example, Pearson correlation). Nevertheless, the Pearson correlation ( $r$ ) can be used as a measure of the strength of relationship between two distance matrices. In this context  $r$  is called the **standardized Mantel statistic** (Sokal & Rohlf 1995), and  $r$  ranges from -1 to 1.

### When to use it

Use a Mantel test to evaluate the congruence between two distance (or similarity) matrices of the same dimensions. The two matrices must refer to the same set of entities in the same order. For example, we may wish to evaluate the correspondence between:

- two groups of organisms from the same set of sample units (e.g., plants and invertebrates),
- dissimilarity matrices representing community structure before and after a disturbance,
- geographic distance and ecological distance, or
- genetic distance and geographic distance.

In contrast to MRPP, which compares multiple groups of various sizes, each group consisting of separate sample units, Mantel tests are appropriate when two distance matrices are to be compared from the same set of sample units. While Mantel tests seek linear relationships between two matrices, the ability to construct those matrices from any distance measure,

similarity measure, or design variable lends great power and flexibility to the approach. Mantel tests are, for example, flexible enough to test almost the same hypothesis as with MRPP (see example below).

### How it works

Mantel's method tests the significance of the correlation between matrices by evaluating results from repeated randomization. The basic question is, "How often does a randomization of one matrix result in a correlation as strong or stronger than the observed correlation?" If randomizations frequently result in correlations between matrices that are as strong as the correlation with the original data, then we have little confidence that the correlation meaningfully differs from zero. On the other hand, a strong correlation structure between matrices will rarely be preserved or enhanced if one of the matrices has been shuffled.

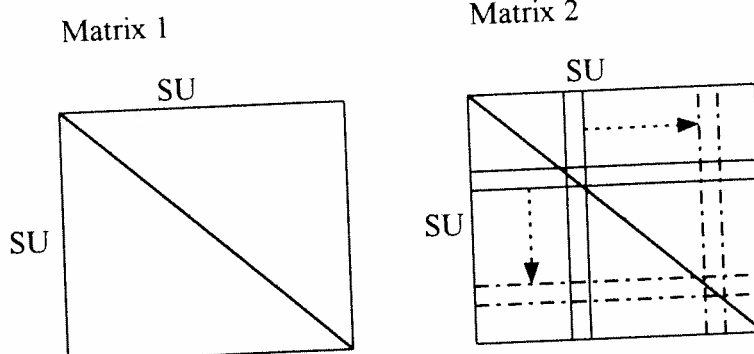
The exact nature of the randomization is important. For the Mantel test, we shuffle the order of the rows and columns of one of the two matrices (it doesn't matter which matrix). The rows and columns are permuted simultaneously, such that for each  $i$ , the  $i$ th row and  $i$ th column will correspond to the same case. The diagonal elements always remain in the diagonal (but in different positions), but nondiagonal elements appear with equal probability in each of the  $n(n-1)$  nondiagonal positions (Box 27.1). After each permutation, Mantel's  $Z$  statistic is calculated and the resulting values provide an empirical distribution that is used for the significance test. The  $Z$  statistic from the nonrandomized data is compared with the distribution of the  $Z$  statistic from the shuffled matrix. Note that  $Z$  is simply the sum of the product of corresponding non-redundant elements of the two matrices, excluding the diagonal.

$$Z = \sum_{i=1}^{n-1} \sum_{j=i+1}^n x_{ij} y_{ij}$$

The standardized Mantel statistic ( $r$ ) is calculated as the usual Pearson correlation coefficient between the two matrices. This is  $Z$  standardized by the variances in the two matrices.

## Box 27.1. Method of permuting a matrix for the Mantel test.

Schematic showing the simultaneous permutation of rows and columns



The starting matrix for the second of two symmetrical matrices is shown below. In this example, each matrix has four sample units. Only one matrix is permuted. For clarity, the contents of the matrix have been replaced with two digits indicating the original row and column in the matrix. For example, "23" originated in row 2, column 3.

	SU1	SU2	SU3	SU4
SU1	11	12	13	14
SU2	21	22	23	24
SU3	31	32	33	34
SU4	41	42	43	44

Select SU3 at random to swap with first item. Swap rows 1 and 3

	SU1	SU2	SU3	SU4
SU3	31	32	33	34
SU2	21	22	23	24
SU1	11	12	13	14
SU4	41	42	43	44

Swap columns 1 and 3

	SU3	SU2	SU1	SU4
SU3	33	32	31	34
SU2	23	22	21	24
SU1	13	12	11	14
SU4	43	42	41	44

This completes one step in the permutation. The step is repeated for items 2, 3, and 4.

You can choose **Mantel's asymptotic approximation** method (much faster with large data sets) or perform a **randomization (Monte Carlo) test** by conducting the actual permutations of one of the distance matrices as described above. Like other randomization tests, a  $p$ -value is calculated from the number of randomizations that yielded a test statistic equal to or more extreme than the observed value. Alternatively, Mantel's asymptotic approximation transforms the  $Z$  statistic into  $t$ , which is asymptotically normal (Mantel 1967, Douglas & Endler 1972). The asymptotic approximation is most useful with a large sample size, because the calculation is much faster than performing the permutations and the approximation is closer than with a small sample size.

Just like a correlation coefficient, the association between matrices can be either negative or positive, as indicated by the sign of the standardized Mantel statistic. If  $r$  is negative, then the randomization test counts the number of randomized runs with  $Z$  less than or equal to the observed value.

## What to report

- ☐ Content and size of the two distance (or similarity) matrices.
- ☐ Method for evaluating test statistic (Mantel's asymptotic approximation or randomization).
- ☐ Standardized Mantel statistic ( $r$ ) as a measure of "effect size."
- ☐  $p$ -value.

## Examples

### Literature

All of the uses of Mantel tests in community ecology that we know have used nonexperimental data, but there is no reason why Mantel tests could not be used for experimental research. See Sokal and Rohlf (1995) for applications to other areas in biology.

McCune and Allen (1985) used Mantel tests to compare site difference matrices with compositional dissimilarity matrices. Although the site variables could be combined into a single composite site difference matrix, they did not do this, because site variables vary greatly in their importance to vegetation, and the inclusion of weak factors with strong on an equal basis can only weaken the site difference matrix.

Burgman (1987) compared four community matrices (total data, common plants only, rare plants

only, and guilds) against each of three environmental matrices.

Tuomisto et al. (1995) compared similarity matrices based on different families of plants and trees in the Amazonian rainforest. They used the "R package" software (Legendre and Vaudor 1991).

Ritchie et al. (2000) compared microbe community profiles with two types of molecular markers, fatty acid methyl esters (FAME) and length-heterogeneity PCR (LH-PCR). Both methods showed differences in soil microbe communities related to agricultural practices. They used a Mantel test on Sørensen distance matrices to demonstrate the fundamental similarity of community structure indicated by FAME vs. LH-PCR ( $r = 0.79$ ,  $p = 0.001$ ). This evidence helped to convince them that the newer method, LH-PCR, gave reliable results.

The Mantel tests offers a very useful solution to the perplexing problem of incorporating *spatial proximity* into models of ecological communities (Urban et al. 2002). By combining geographic coordinates (such as UTM's) into a distance matrix, we can use Mantel tests to evaluate the relationship between distances in species space and distances in geographic space. Urban et al. distinguished between two kinds of spatial models: "distance apart" and "location." A Mantel test using a geographic distance matrix evaluates distance apart, consistent with an interpretation of local spatial structure due to autocorrelated environmental variables or a contiguous spatial process. On the other hand, by incorporating geographic coordinates in CCA and RDA (redundancy analysis), we represent location. Urban et al. (2002) gave the example:

...consider a case where vegetation pattern is controlled by a strong directional gradient (say, elevation) in addition to a local spatial process (say, seed dispersal). As an explanatory variable, *distance apart* would capture the dispersal process, while *location* might not. In fact, for the residuals from elevation (i.e., variation in species composition *not* predicted by elevation) to be correlated with *location*, the residuals would have to either increase or decrease monotonically with location (say, longitude or UTM easting). But we would expect seed dispersal to act similarly over the elevation gradient, an expectation consistent with *distance apart* as an explanatory variable.

### Comparison of asymptotic and randomization methods

1. We have a data set of vegetation (A) and associated site characteristics (E) for 19 sample units.

In the first example, we test the hypothesis of no relationship between vegetation and site characteristics ( $H_0$ : A is unrelated to E).

The problem is approached in two ways, first with Mantel's asymptotic approximation, then with a randomization procedure (see example output below). The standardized Mantel statistic  $r$  is, of course, the same, because the choice of method only pertains to evaluating the statistical significance of  $r$ . In both cases we can soundly reject the null hypothesis. Note that the  $p$ -value is considerably higher with the

randomization test. Yet it is the smallest possible  $p$ -value given the number of randomizations selected. If  $n$  is the number of randomized runs with  $Z \geq Z_{\text{obs}} = 28.3756$ , and  $N$  is the number of randomized runs, then

$$p = \frac{1+n}{1+N} = \frac{1+0}{1+1000} = 0.001$$

As in this example, the smallest possible  $p$ -value is always  $1/(1+N)$ .

**Example Output** (from PC-ORD, McCune & Mefford 1999):

With Mantel's asymptotic approximation the results are:

```
DATA MATRICES
Main matrix:
  19 STANDS (rows)
  50 SPECIES (columns)
Distance matrix calculated from main matrix.
Distance measure = SORESENSE

Second matrix:
  19 STANDS (rows)
  6 ENVIRON (columns)
Distance matrix calculated from second matrix.
Distance measure = SORESENSE

TEST STATISTIC: t-distribution with infinite degrees of freedom
using asymptotic approximation of Mantel (1967).
If t < 0, then negative association is indicated.
If t > 0, then positive association is indicated.

STANDARDIZED MANTEL STATISTIC: .481371 = r
OBSERVED Z = .2838E+02
EXPECTED Z = .2645E+02
VARIANCE OF Z = .1222E+00
STANDARD ERROR OF Z = .3496E+00
t = 5.4969
p = .00000005
```

With the randomization method, the form of the results is somewhat different:

```
MANTEL TEST RESULTS: Randomization (Monte Carlo test) method
.481371 = r = Standardized Mantel statistic
.283756E+02 = Observed Z (sum of cross products)
.264450E+02 = Average Z from randomized runs
.120814E+00 = Variance of Z from randomized runs
.255765E+02 = Minimum Z from randomized runs
.278301E+02 = Maximum Z from randomized runs
1000 = Number of randomized runs
0 = Number of randomized runs with Z > observed Z
.001000 = p (type I error)

-----
p = proportion of randomized runs with Z ≥ observed Z; i.e.,
p = (1 + number of runs ≥ observed) / (1 + number of randomized runs)
Positive association between matrices is indicated by observed
Z greater than average Z from randomized runs.
```

### Comparison of two species groups

Using the same data set, we now partition the species matrix ( $A$ ) into two parts, grasses ( $A_g$ ) and other components of the vegetation ( $A_v$ ). Our new hypothesis is that the distribution of grasses is independent of the other components of the vegetation. ( $H_0$ :  $A_g$  is unrelated to  $A_v$ . Test statistic:  $Z$ .)

Dissimilarity matrices are calculated separately for  $A_g$  and  $A_v$  using the quantitative version of the Sorensen coefficient. The relationship between the two matrices is shown in a scatterplot (Fig. 27.1). There is considerable scatter in the relationship, though a weak positive relationship is apparent. This appearance is supported by a smallish standardized Mantel statistic ( $r = 0.24$ ) that is, nevertheless, statistically significant ( $p = 0.032$ ) based on the randomization test. We can conclude that dissimilarity in grasses is positively, but weakly related to dissimilarity in other components of the vegetation.

The randomization test yielded the following results: Of the 1000 randomizations, 31 had a  $Z$  as large or larger than the observed  $Z$  (Fig. 27.2).

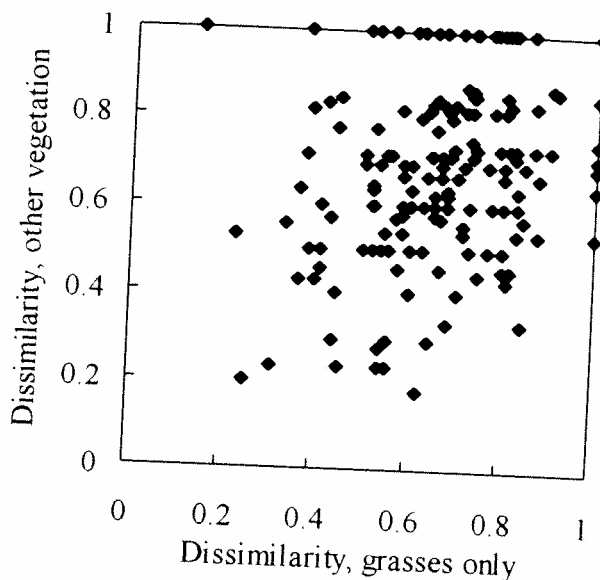


Figure 27.1. Scatterplot of dissimilarity of plots based on grasses against dissimilarity of plots based on other species.

### Example output, comparison of two species groups:

```
Method chosen is a randomization (Monte Carlo) test.
Monte Carlo test: null hypothesis is no relationship between matrices
No. of randomized runs:      1000
Random number seeds:        1217
```

```
MANTEL TEST RESULTS: 'Randomization (Monte Carlo test) method
```

```
-----
0.238500 = r = Standardized Mantel statistic
0.158331E+03 = Observed Z (sum of cross products)
0.155539E+03 = Average Z from randomized runs
0.171850E+01 = Variance of Z from randomized runs
0.152315E+03 = Minimum Z from randomized runs
0.160216E+03 = Maximum Z from randomized runs
1000 = Number of randomized runs
31 = Number of runs with Z > observed Z
0 = Number of runs with Z = observed Z
969 = Number of runs with Z < observed Z
0.032000 = p (type I error)
-----
```

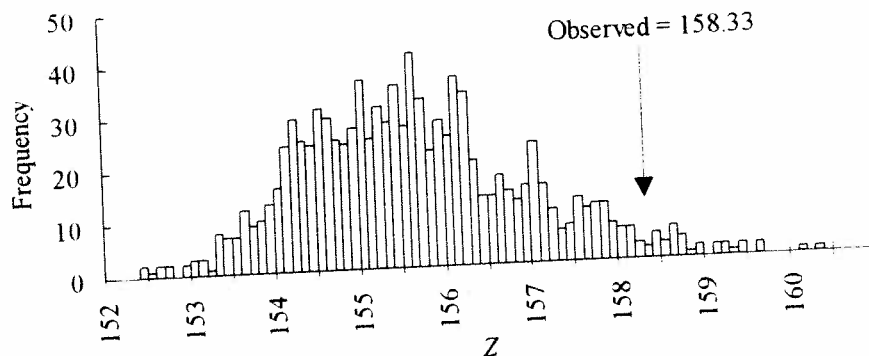


Figure 27.2. Frequency distribution of  $Z$ , the Mantel test statistic, based on 1000 randomizations for the same example as in Figure 27.1.

### Comparison of Mantel test and MRPP

Refer to Chapter 24 on MRPP (p. 190) for a small example comparing three groups of five plots. MRPP tested the hypothesis of no difference between groups in species space. A similar comparison can be made with a Mantel test (though less conveniently), by constructing a design matrix (Table 27.1), then comparing the design matrix to a matrix of Sørensen distances in species space (Fig. 27.3).

The design matrix in this case is a matrix of  $15 \times 15$  plots, each element indicating whether the comparison is a between-group comparison or a within-group comparison (Table 27.1). The design matrix is treated as a distance matrix, with zero "distance" for a within-group comparison and ones indicating a between-group comparison.

The standardized Mantel statistic equals 0.44. The  $Z$  statistic in this case is 107.2, which is compared to an expected value of 93.7 with a variance of 6.9. Mantel's asymptotic approximation yields  $t = 5.1$  and  $p \approx 10^{-7}$ . Using the randomization method with 9999 trials yielded  $p = 0.0003$ , which is quite similar to the result from MRPP ( $p = 0.0007$ ) using the same distance measure.

Although in this case the results are quite similar, MRPP should be the preferred approach for this problem. For one-way classifications, it is easier to explain and more convenient to run than the Mantel test.

### Variations

Rank-transformed distance matrices (Dietz 1983, Lefkovich 1984) should be explored as a useful transformation of species data prior to a Mantel test.

judging by the success of ranked distance matrices in nonmetric multidimensional scaling of community data (Ch. 16). This would be analogous to using a Spearman rank correlation coefficient instead of a Pearson correlation coefficient.

Smouse et al. (1986) extended the method to deal with more than two matrices at once (analogous to multiple regression). They calculated a partial Mantel statistic, expressing the relationship between two matrices while controlling for a third matrix. The calculation of the partial Mantel statistic is essentially the same as a partial correlation coefficient (see Ch. 30). The significance of the partial Mantel statistic is tested by permutation. Two permutation methods are given by Smouse (1986) and Legendre and Legendre (1998).

# Mantel Test

Table 27.1. Design matrix for Mantel test of same hypothesis as MRPP: no multivariate difference among three groups. Within-group comparisons are assigned a zero, between group comparisons a one.

	Plot														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
plot1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
plot2	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
plot3	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
plot4	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
plot5	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
plot6	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1
plot7	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1
plot8	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1
plot9	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1
plot10	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1
plot11	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
plot12	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
plot13	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
plot14	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
plot15	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0

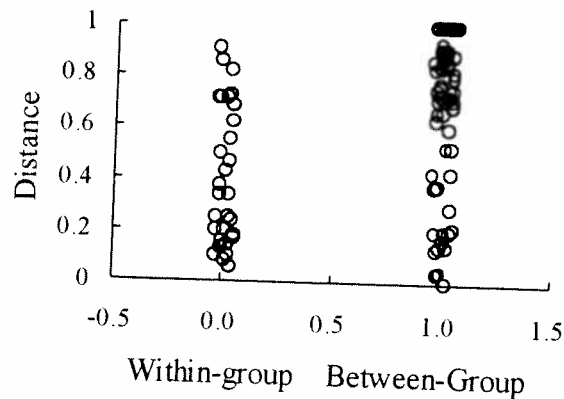


Figure 27.3. Scatterplot of Sørensen distances in species space against values in the design matrix (0 = within group, 1 = between group). The 0/1 values in the design matrix have been jittered by adding a small random number.