



# **User Manual for Blossom Statistical Package for R**

By Marian K. Talbert and Brian S. Cade

Open-File Report 2005–1353

**U.S. Department of the Interior**  
**U.S. Geological Survey**

**U.S. Department of the Interior**  
SALLY JEWEL, Secretary

**U.S. Geological Survey**  
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2013

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <http://www.usgs.gov> or call 1–888–ASK–USGS

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov> The Blossom user manual and software for the original Fortran version is available for download from the web page of the Fort Collins Science Center, <http://www.fort.usgs.gov/products/software/blossom/blossom.asp>

Although this program has been used by the U.S. Geological Survey (USGS), no warranty, expressed or implied, is made by the USGS or the U.S. Government as to the accuracy and functioning of the program and related program material nor shall the fact of distribution constitute any such warranty, and no responsibility is assumed by the USGS in connection therewith.

Suggested citation:

Talbert, M.K., and Cade, B.S., 2013, User manual for Blossom statistical package for R: U. S. Geological Survey Open-File Report 2005–1353, 81 p., <http://pubs.usgs.gov/of/2005/1353/>.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

# Contents

Introduction.....	1
Overview of Statistical Concepts .....	4
Statistical Commands.....	6
Multiresponse Permutation Procedure (MRPP) .....	7
Multiresponse Randomized Block Procedure (MRBP).....	25
Permutation Tests for Matched Pairs (PTMP).....	31
Multiresponse Sequence Procedure (MRSP) .....	37
Least Absolute Deviation (LAD) and Quantile Regression.....	39
Regression Quantiles .....	52
Ordinary Least Squares Regression (OLS).....	66
G-sample and 1-sample Goodness-of-fit Coverage Tests .....	68
Appendix 1. Common Statistical Tests Embraced by the MRPP Command .....	73
Appendix 2. Compilation of the internal Blossom code.....	74
Appendix 3. Blossom Statistics Program Installation, Configuration, Requirements.....	75
Blossom Requirements and Program Limits.....	75
Appendix 4. Blossom Development and Testing .....	77
Appendix 5. Acknowledgments.....	77
References Cited.....	77

## Figures

<b>Figure 1.</b> The observed sample for 2 groups with bivariate response $Y_{\text{Coord}}$ and $X_{\text{Coord}}$ .....	7
<b>Figure 2.</b> One of the possible other 34 permutations of the data in figure 1 .....	8
<b>Figure 3.</b> Two-group example from Mielke, 1986 (no outliers) .....	10
<b>Figure 4.</b> Two-group example from Mielke, 1986 (one outlier in Group 2) .....	12
<b>Figure 5.</b> Migration distance and elevation change for 9 male and 12 female blue grouse (from Cade and Hoffman, 1993).....	15
<b>Figure 6.</b> Pattern of immigration and emigration for female striped newts at Breezeway Pond, Florida, 1985–1990.....	25
<b>Figure 7.</b> Proportion of basal area and canopy cover for lodgepole pine in 31 stands .....	28
<b>Figure 8.</b> Latitude and longitude coordinates (m) for blue grouse (no. 162) locations on 12 dates November through March .....	39
<b>Figure 9.</b> Lodgepole pine canopy cover as a linear function of basal area and tree density estimated with LAD regression for 31 sample stands.....	43
<b>Figure 10.</b> Engelmann spruce canopy cover as a quadratic function of basal area and linear function of tree density estimated with LAD regression for 31 sample stands .....	44
<b>Figure 11.</b> Soap scrap as a linear function of production speed for line 1 (open circles) and line 2 (solid circles) .....	47
<b>Figure 12.</b> Average annual biomass of acorns and acorn suitability indices based on oak forest characteristics in 43 0.2-ha sample plots in Missouri .....	57
<b>Figure 13.</b> Glacier lily seedling counts, lily flower numbers, and rockiness index for 256 2 x 2 m quadrats in subalpine meadow of western Colorado .....	59

<b>Figure 14.</b> Lahontan cutthroat trout $m^{-1}$ and width:depth ratios for small streams sampled 1993 to 1999 ( $n = 71$ ); exponentiated estimates for 0.90, 0.50, and 0.10 regression quantiles for the weighted model $w(\ln y) = w(\beta_0 + \beta_1 X_1 + (\gamma_0 + \gamma_1 X_1)\epsilon)$ , $w = (1.310 - 1.017X_1)^{-1}$ .....	63
<b>Figure 15.</b> Quantile plots of migration distances for 9 male and 12 female blue grouse .....	70
<b>Figure 16.</b> Quantile plots of elevation changes (m) made by 9 male and 12 female blue grouse when migrating from breeding to winter areas .....	71

# User Manual For Blossom Statistical Package for R

By Marian K. Talbert and Brian S. Cade

## Introduction

Blossom is an R package with functions for making statistical comparisons with distance-function based permutation tests developed by P.W. Mielke, Jr. and colleagues at Colorado State University (Mielke and Berry, 2001) and for testing parameters estimated in linear models with permutation procedures developed by B. S. Cade and colleagues at the Fort Collins Science Center, U.S. Geological Survey. This manual is intended to provide identical documentation of the statistical methods and interpretations as the manual by Cade and Richards (2005) does for the original Fortran program, but with changes made with respect to command inputs and outputs to reflect the new implementation as a package for R (R Development Core Team, 2012). This implementation in R has allowed for numerous improvements not supported by the Cade and Richards (2005) Fortran implementation, including use of categorical predictor variables in most routines.

Statistical procedures documented in this manual include:

- A permutation testing version of ordinary least squares (OLS) regression that parallels the least absolute deviation (LAD) and quantile regression permutation tests;
- A permutation and asymptotic chi-square approximation of  $P$ -values for a rank score statistic for regression quantiles;
- Double permutation (`hypothesis.test` with `double.permutation`) procedures for linear model tests (OLS, LAD regression, and quantile rank score tests) when null models are either implicitly or explicitly constrained through the origin, that is, no intercept models (Cade, 2003; Cade and others, 2005; Cade and others, 2006);
- Dropping all but a single zero residual in LAD and quantile regression permutation tests of subsets of variables in multiple regression models (Cade, 2005; Cade and Richards, 2006);
- Computation of all quantile regression estimates (`lad` with the option `all.quantiles = TRUE`);
- Empirical coverage tests for univariate goodness-of-fit and  $g$ -sample comparisons that are extensions of the Kolmogorov-Smirnov family of statistics for comparing cumulative distribution functions, including an option for testing goodness-of-fit for a random uniform distribution on a circle;
- The multiresponse permutation procedures (MRPP) family of statistics including two options for standardizing multiple dependent variables (average Euclidean distance or Hotelling's commensuration based on variance/covariance);

- Computing exact probabilities by complete enumeration of all possible combinations for small block and treatment designs in MRPP and multiresponse randomized block permutation procedures (MRBP);
- A Monte Carlo resampling approximation alternative for all the MRPP family of statistics (the `mrpp` commands with options `number.perms`); and
- Multivariate medians and distance quantiles (MEDQ) to be used as descriptive statistics with MRPP analyses. In addition, we offer the option to store (`save.test = TRUE`) the vector of permuted test statistic values from Monte Carlo resampling approximations of probabilities.

The permutation procedures in Blossom can be used for comparing data obtained in familiar survey sampling and comparative experimental designs.

1. Multiresponse permutation procedures (MRPP) are used for univariate and multivariate analyses of grouped data in a completely randomized one-way design. MRPP are used for comparing equality of treatment groups analogous to one-way analysis of variance (or  $t$ -test) for univariate data, or multivariate analysis of variance (or Hotelling's  $T^2$ ) for multivariate data. The default Euclidean distance function in MRPP provides an omnibus test of distributional equivalence among groups or a test for common medians if the assumption of equal dispersions is applicable. Options allow MRPP to perform permutation (randomization) versions of  $t$ -tests, one-way analysis of variance, Kruskal-Wallis tests (for ranked data), Mann-Whitney Wilcoxon tests (for ranked data), and one-way multivariate analysis of variance. Options in MRPP also allow truncation of distances to evaluate multiple clumping of data, establish an excess group, and select arc distances to compare circular distributions of grouped data. Multivariate data are commensurated (standardized) to a common scale but an option allows for no commensuration. Commensuration can be done by using average Euclidean distance (default) or the variance/covariance matrix for the dependent variables. Multivariate medians and distance quantiles (MEDQ) are provided as estimates to be used in describing distributional changes detected by MRPP analyses.
2. Multiresponse permutation procedures for randomized blocks (MRBP) are used for univariate and multivariate analyses of grouped data in a complete randomized block design. Again, the default Euclidean distance function provides an omnibus test for equivalence of distributions or common medians if the assumption of equal dispersions is satisfied. Univariate comparisons are analogous to analysis of variance or Friedman's test (for ranked data) for complete randomized block designs. Options allow MRBP to perform permutation versions of these two tests. Options also allow for aligned or unaligned data analyses and to commensurate or not commensurate multivariate data. MRBP also can be used to calculate agreement measures among blocks. A linear transform of Pearson's correlation coefficient and a permutation test of significance also can be calculated in MRBP.
3. The permutation test for matched pairs (PTMP) is a special case of MRBP, with univariate data in two groups and  $n$  blocks, used for paired comparisons. Options allow PTMP to perform permutation versions of paired  $t$ -tests and Wilcoxon's signed rank test (for ranked data).
4. Multiresponse sequence procedure (MRSP) is a special case of MRPP where first-order sequential pattern of data is tested against the null hypothesis of no sequential pattern. Univariate analyses are analogous to the Durbin-Watson test for first-order serial pattern and bivariate analyses are analogous to Schoener's  $t^2/r^2$  statistic (Solow, 1989).

Permutation versions of these two tests can be done. Options allow selection of the sequencing variable and to turn off multivariate commensuration.

5. Least absolute deviation (LAD) regression is an alternative to ordinary least squares (OLS) regression that has greater power for thick-tailed symmetric and asymmetric error distributions (Cade and Richards, 1996). LAD regression estimates the conditional median (a conditional 0.50 quantile) of a dependent variable given the independent variable(s) by minimizing sums of absolute deviations between observed and predicted values. Options allow for testing all slope parameters (full model) equal to zero or to test subsets of parameters (partial models) equal to zero by Monte Carlo resampling of the permutation distribution (Cade and Richards, 1996, 2006). LAD regression can be used anywhere OLS regression would be used but is often more desirable because it is less sensitive to outlying data points and is more efficient for skewed error distributions as well as for some symmetric error distributions.
6. Regression quantiles are a natural extension of LAD regression to estimate any conditional quantile and provided as an option in LAD regression. Regression quantiles allows estimation of any conditional quantile (say  $\tau$ ,  $0 < \tau < 1$ ) of a dependent variable given the independent variable(s) by minimizing the asymmetrically weighted sum of absolute deviations, where the weights are  $\tau$  for positive residuals and  $1 - \tau$  for negative residuals. A 0.50 regression quantile is LAD regression. Regression quantiles are useful in ecological applications involving limiting factors where it is desirable to estimate functional changes along boundaries of distributions (Terrell and others, 1996; Cade and others, 1999; Cade and Guo, 2000; Dunham and others, 2002; Cade and others, 2005) and for general modeling of rates of change associated with heterogeneous variation in linear models. Cade and Noon (2003) provide a primer on quantile regression for ecological applications. The LAD permutation tests of Cade and Richards (1996) have been extended to regression quantiles, including improvements to handle multiple variables under alternative hypotheses and estimates forced through the origin (Cade, 2003; Cade and Richards, 2006). Another permutation testing alternative also is provided that is based on the quantile rank score functions for regression quantiles (Koenker, 1994; Cade and others, 1999; Koenker and Machado, 1999), which maintains better Type I error rates than the Cade and Richards (1996) procedure when there are heterogeneous errors. The permutation approximation of  $P$ -values for the quantile rank score test statistic was evaluated in Cade, 2003; Cade and others, 2005; and Cade and others, 2006. The  $P$ -value based on the asymptotic Chi-square approximation of Koenker (1994) is also reported and was also evaluated by Cade, 2003; Cade and others, 2005; and Cade and others, 2006. Both test statistics require weighted estimates to maintain correct Type I error rates with heterogeneous distributions. It is possible to estimate all possible regression quantiles and save the estimates by quantiles to a specified data frame.
7. G-sample and goodness-of-fit tests based on empirical coverages (COV) are for univariate comparisons of grouped data similar to the Kolmogorov-Smirnov family of statistics for comparing cumulative distribution functions (Mielke and Yao, 1988, 1990). These statistics are appropriate for continuous univariate responses with no or few tied values. Options allow for testing goodness-of-fit to a uniform distribution on the unit circle, which is equivalent to a permutation version of Rao's spacing test (Rao, 1976).

Appendix 1 lists common statistical tests encompassed by these permutation procedures.

The methods contained in Blossom are presented by example. Most of the examples are from ecology, but of course the procedures in Blossom can be used on many other sorts of data.

## Overview of Statistical Concepts

The statistical procedures in Blossom are distribution free in the sense that probabilities of obtaining extreme test statistic values given the truth of the null hypothesis (Type I errors) are based on permutations of the data from randomization theory and are not based on an assumed population distribution (Edgington, 1987; Good, 2000; Mielke and Berry, 2001). In most investigations, the population distribution will never be known and assuming an inappropriate distributional model can lead to weak or invalid statistical inferences. The normal distribution is an inappropriate model for many ecological data, which often are skewed, discontinuous, and multi-modal. When sample sizes are small, large sample (asymptotic) approximations often are questionable. Permutation procedures make efficient use of small samples, because probabilities can be calculated exactly by complete enumeration of all possible combinations under the null hypothesis. Of greater importance, the permutation testing framework allows us to use test statistics based on measures of variation other than squared deviations (variances). Test statistics based on variances are often derived from the distributional assumptions underlying the maximum likelihood approach. Other measures of variation may be more appropriate in a permutation test that does not require assumptions about the specific form of the error distribution.

The distance-functions that form the basis of the MRPP family of tests allow test statistics to be based on powers of Euclidean distances. The distance function between any two observations  $x_i$  and  $x_j$  with  $r$  response variables (dependent variables) in MRPP is defined by

$$\Delta_{i,j} = \left[ \sum_{h=1}^r (x_{h,i} - x_{h,j})^2 \right]^{v/2}$$

where  $v > 0$  (Mielke and Berry, 2001). We emphasize use of test statistics based on ordinary Euclidean distances  $v=1$  (`expon = 1` in the `mrpp` command), a metric measure of variation that is congruent with most data measurement scales (Mielke, 1986; Biondini and others, 1988). Euclidean distances are the common geometrical interpretation of distance applied to differences between replicate data values on their measurement scale. Most conventional parametric and nonparametric methods are based on squared Euclidean distances (squared deviations are squared Euclidean distances, that is, `expon = 2`). Statistics based on squared Euclidean distances (variances) are nonmetric measures (they violate the triangle inequality of a metric) that have no simple geometrical interpretation in an  $r$ -dimensional data space, where  $r$  is the number of response variables. In contrast to Euclidean distance statistics, geometrical interpretation of variance based statistics involves distances between vectors in an  $n$ -dimensional space, where  $n$  is sample size (Box and others, 1978, p. 197–203). An  $n$ -dimensional geometric interpretation is complex, does not coincide with the data space, and results in considerable loss of graphical information because distances between replicates vanish. It is impossible to graph individual data points in a nonmetric space to examine dissimilarities (Pielou, 1984, p. 41–46). Although we emphasize tests based on Euclidean distances, analyses based on powers other than 1 (Euclidean distance) are appropriate in some specific applications.

Euclidean distance based statistics have greater power (the probability of rejecting the null hypothesis when it is false) to detect location (central tendency) shifts among skewed distributions than do squared Euclidean distance statistics (Zimmerman and others, 1985; Biondini and others, 1988; Mielke and Berry, 2001). Power to detect location shifts in symmetric distributions with Euclidean distance statistics is greater than or equal to power with squared Euclidean distance statistics, depending on distributional form (Mielke and others, 1981; Mielke and Berry, 1982; Mielke, 1984; Mielke and Berry, 1994, 1999, 2001). Euclidean distance based

statistics have better power to detect location shifts across a greater variety of distributions than squared Euclidean distance (variance) statistics. Euclidean distance based statistics also are used to detect omnibus differences in distributions, sensitive to both dispersion (variation) and shifts in central tendency (median) (Biondini and others, 1988; Mielke and Berry, 1994). There is no reason to presume that shifts in central tendency of data distributions characterize the only effects of interest in ecological investigations.

The permutation procedures based on distance functions are readily extended to several novel applications, including truncation of values to detect multiple clustering, comparisons of circular distributions, assignment to an excess group, agreement of values, and first-order autoregressive analyses (Mielke, 1991; Mielke and Berry, 2001). Each of these applications will be discussed in appropriate examples.

Medians and other quantiles are estimates obtained by minimizing sums of absolute deviations and are appropriate descriptive statistics for permutation procedures based on Euclidean distance functions (Mielke and Berry, 2001). Functions are provided to estimate multivariate medians of grouped data and quantiles for distances between individual observations and their group median. This function can also be used to compute medians and any selected quantiles for univariate data distributions.

Permutation procedures for testing hypotheses in linear models are available for least absolute deviation (LAD) regression (Cade and Richards, 1996), a generalization for regression quantiles (Cade and others, 1999; Cade, 2003), and for ordinary least squares regression (Anderson and Legendre, 1999). LAD regression estimates rates of change in conditional medians, whereas the more familiar OLS regression estimates rates of change in conditional means. Regression quantiles estimate rates of change in any selected conditional quantile (Koenker and Bassett, 1978). The forms of the permutation test statistics are similar for all three of these estimation methods, and are based on a proportionate reduction in sums minimized when passing from a null, reduced parameter model to the alternative, full parameter model (Mielke and Berry, 2001; Cade, 2005). These tests are a drop in dispersion form. The observed test statistic,  $T_{obs}$ , equals the (sum of deviations for reduced parameter null model–sum of deviations for full parameter alternative model)/ sum of deviations for full parameter alternative model; where the deviations are squared residuals if OLS regression, absolute values of residuals if LAD regression, or weighted absolute values of residuals if regression quantiles. This test statistic is equivalent to the usual  $F$ -ratio used in OLS regression, except that the sums minimized are not divided by their degrees of freedom ( $df$ ) because they are invariant under the permutation arguments. Hypothesis testing for all three of these regression estimates are made either by permuting the dependent variables for full model tests that all slope parameters are zero (null model includes just an intercept) or by permuting residuals from reduced parameter null model for partial model tests (subhypotheses) that some specified subset of slope parameters are zero (null model includes more than just an intercept). Extensive simulation work has demonstrated the approximate validity of permuting residuals under the reduced parameter null model when making permutation tests involving nuisance parameters in linear models (Cade and Richards, 1996; Kennedy and Cade, 1996; Anderson and Legendre, 1999). Simulation research (Cade, 2003; Cade, 2005; Cade and others, 2005; Cade and others, 2006; Cade and Richards, 2006) has demonstrated that Type I error rates can be improved by using double permutation schemes when null models are constrained through the origin (no intercept) and by deleting all but a single zero residual when LAD and quantile regression null models include multiple independent variables.

All the tests described above for the linear model maintain validity of their type I error rates only if it is reasonable to assume independent and identically distributed (*i.i.d.*) errors. If

the errors are heterogeneous as happens when the variance changes as a function of the independent variables, other methods must be employed. One possibility is to estimate weighted versions of either LAD or OLS regression, where weights are selected to be inversely proportional to the square root of the variances. Permutation testing then is employed on the weighted transforms of the dependent ( $y$ ) and independent ( $X$ ) variables (Cade and Richards, 1996; Cade, 2005; Cade and others, 2005; Cade and others, 2006). Alternatively, for the regression quantile estimates, we provide a permutation test for the quantile rank score statistic (Koenker, 1994; Koenker and Machado, 1999), which is not as sensitive to heterogeneity of variances because it uses the signs of the residuals and not their magnitude. Statistical performance of the permutation test for the quantile rank score statistic was investigated by Cade, 2003; Cade and others, 2005; and Cade and others, 2006. Weighted estimates and rank score tests were required to maintain correct Type I error rates when heterogeneity exceeded a change in 2.5 standard deviations across the domain of the independent variable. Blossom also reports the asymptotic version of the quantile rank score statistic that is distributed as a Chi-square distribution with degrees of freedom equal to the difference in number of parameters ( $q$ ) between alternative (full  $p$  parameters) and null (reduced  $p - q$  parameters) models (Koenker, 1994; Koenker and Machado, 1999).

The empirical coverage tests included in Blossom are related to the Kolmogorov-Smirnov family of tests for equality of univariate cumulative distribution functions. One-sample goodness-of-fit and  $g$ -sample tests exist. The coverage test statistic is based on the spacings between the order statistics. These tests provide another permutation testing alternative to MRPP for univariate continuous data. Unlike MRPP, the coverage test are not appropriate when there are many tied values, as this violates the continuity assumption. Little can be said at this time about the power of the coverage tests relative to MRPP for data for which both tests are appropriate. Go forth and investigate!

## Statistical Commands

Blossom currently has nine statistical commands, `mrpp`, `mrbp`, `ptmp`, `sp`, `medq`, `lad`, `ols`, `hypothesis.test`, and `coverage`.

The distance-function procedures (MRPP, MRBP, and PTMP) are distribution-free techniques for making inferences about grouped data. Their advantages over many classical techniques include the ability to select an analysis space commensurate with the geometry of the data as perceived by the investigator. Several classical univariate and multivariate parametric and rank tests can be emulated with these procedures as well. The simplest MRPP analysis is for data consisting of two or more observations on objects in two or more groups. The MRBP and PTMP variants are for similar data that are blocked or paired.

The `medq` command calculates univariate or multivariate medians and distance quantiles either by groups specified by a grouping variable or for the entire dataset being used. Options allow specification of quantiles to report that differ from the default quantiles.

The `sp` command calculates the multiresponse sequence procedure to test for first-order autoregressive patterns (serial dependency). The default value produces an analysis in Euclidean space. A sequencing variable that determines the order of the data can be selected or Blossom assumes by default that the order in the dataset is the sequential order of interest.

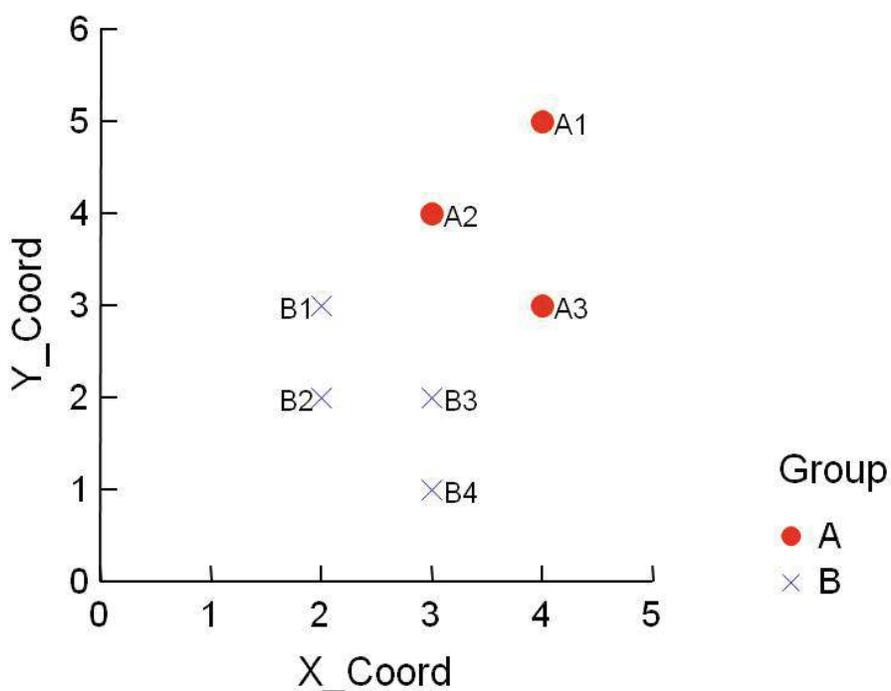
The `lad` command estimates a least absolute deviation regression or an optional quantile regression. The model specified in the `lad` command line is considered the full parameter alternative model for hypothesis tests. The associated command, `hypothesis.test`, can be used to test two models specified by the `lad` command.

The `ols` command estimates an ordinary least squares regression by setting the optional argument `OLS = TRUE` in the `lad` command. The `hypothesis.test` can be used for OLS regression just as with LAD regression.

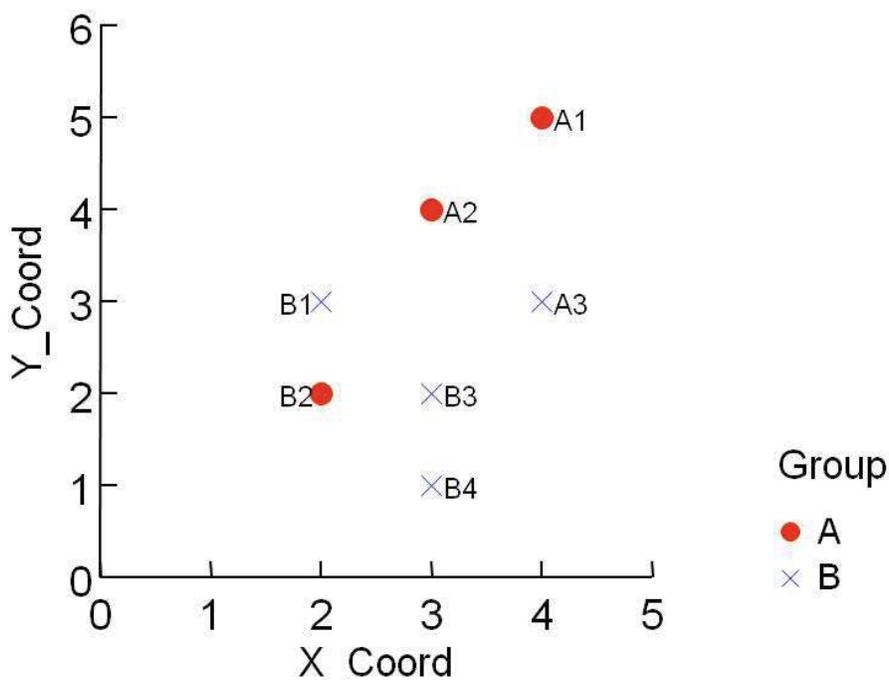
The `coverage` command provides for tests of g-sample empirical coverage tests it used with a grouping variable and related goodness-of-fit tests if specified without a grouping variable.

The MRPP variants, MRSP, LAD, OLS, and COV are discussed in turn. MEDQ is discussed with MRPP as it provides descriptive estimates that are useful for interpreting results of hypothesis tests with MRPP.

### Multiresponse Permutation Procedure (MRPP)



**Figure 1.** The observed sample for 2 groups with bivariate response  $Y_{Coord}$  and  $X_{Coord}$ .



**Figure 2.** One of the possible other 34 permutations of the data in figure 1.

MRPP is best introduced with an example. The following is a bivariate example adapted from Biondini and others (1985). A similar example is found in Zimmerman and others (1985), Biondini and others (1988), and a univariate example is given in Slauson (1988).

In figure 1, the values of two variables  $x$  and  $y$  are shown for seven observations in two groups,  $A$  and  $B$ . The objects in groups  $A$  and  $B$  seem to be clustered or concentrated in different parts of the  $x$ - $y$  plane representing the two response (measured) variables  $x$  and  $y$ . One way to determine if the two groups are so clustered is to measure or calculate the distances between all pairs of members of each group and calculate an average distance for each group ( $A = 1.609$ ,  $B = 1.344$ ). If group members are clustered together, then the intragroup average distances will be small compared to cases where the group members are spread out and overlap more with other groups. For example, figure 2 shows the same data except that the groups that observations  $A3$  and  $B2$  belong to are switched. In this case the intragroup average distances will be greater than for the case first shown above ( $A = 2.419$ ,  $B = 1.717$ ).

The strategy of MRPP is to compare the observed intragroup average distances with the average distances that would have resulted from all the other possible combinations of the data under the null hypothesis. The test statistic, usually symbolized with a lower case delta,  $\delta$ , is the average of the observed intragroup distances weighted by relative group size,  $3/7$  and  $4/7$  in this case. The observed delta  $\delta_{\text{obs}}$  is compared to the possible deltas  $\delta$  resulting from every permutation of the above seven points into two groups of three and four members. If the hypothesis that the distribution of the two groups are not different (the null hypothesis) is true, then each of the possible assignments (permutations) is equally likely. In this example there are 35 permutations possible, each with a  $1/35$  ( $1/35 = 0.0286$ ) chance of occurring. Here are the Blossom commands to compute the MRPP results on the dataset `example1`.

```
> out <- mrpp(example1, commens = FALSE, exact = TRUE)
```

`x_coord` and `y_coord` are the two response variables, `group` is the grouping variable, and the exact version of MRPP is chosen since this is such a small sample. `commens = FALSE` signifies that no multivariate commensuration is desired. Blossom by default will commensurate multiple variables by the average Euclidean distance for each variable ignoring group structure. Think of this as similar to the usual parametric approach of standardizing variables to unit variance (average squared Euclidean distance).

```
> summary(out)
      Exact Multi-Response Permutation Procedure (EMRPP)

Call:
mrpp(data = example1, commens = FALSE, exact = TRUE)

  Grouping Variable :  group
  Response Variables:  x_coord y_coord

Specification of Analysis:
  Number of Observations: 7
  Number of Groups       : 2
  Distance Exponent      : 1
  Weighting Factor       : n(I)/sum(n(I))=C(I) = 1

Group Summary:
  Group Value  Group Size
1             3
2             4

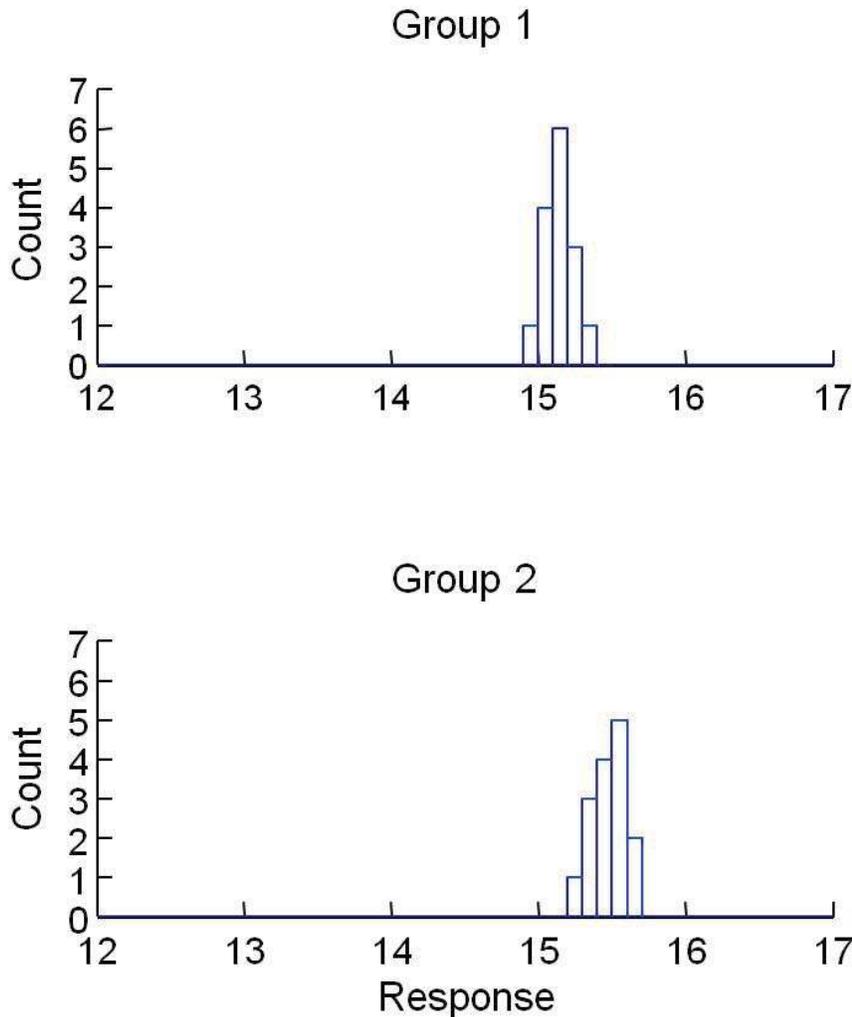
  Variables are not commensurated

Results:
  Delta Observed           : 1.458
  Probability (Exact)
  of a smaller or equal delta: 0.02857*
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The probability value (*P*-value) is 0.0286 which means that the observed delta was the smallest among the 35 possible deltas. Use the `exact = TRUE` option for `mrpp` with caution for it can take a long time if the sample sizes are greater than about 30, depending on the computer.

By default `mrpp` does not compute exact probabilities but uses an approximation of the exact distribution of the test statistic ( $\delta$ ) to estimate the *P*-value. The default approximation is based on the first three exact moments (mean, variance, and skewness) of the permutation distribution evaluated as a Pearson type III distribution (Berry and Mielke, 1983; Iyer and others, 1983; Mielke and Berry, 2001). The moments approximation avoids the simulation error associated with Monte Carlo resampling tests (Mielke and Berry, 1982; Berry and Mielke, 1985). However, we offer the option of approximating the permutation distribution of the test statistic with a Monte Carlo resampling procedure by setting `number.perms` equal a positive integer value. The user may specify any desirable number of resamples, for example, `number.perms = 10000`. Most examples we've encountered yield similar *P*-values from the Monte Carlo resampling and Pearson type III distribution approximations, but it is possible for the Monte Carlo resampling approximation to yield better estimates for some problems, for example, with a large number of discrete values clumped in some region of the data space or if interest is in upper

tail probabilities (for example,  $P > 0.90$ ) associated with detecting regularity of spatial data distributions. Further investigation of these properties is an open area for research.



**Figure 3.** Two-group example from Mielke, 1986 (no outliers).

The next example shows how to emulate a 2-sample  $t$ -test with the `mrpp` command. Consider the data for two groups in figure 3 (from Mielke, 1986). The single response variable is represented on the horizontal axis and the number of observations on the vertical.

Group 1 (median = 15.10, mean = 15.09 ) and 2 (median = 15.40, mean = 15.42 ) appear to differ slightly (0.3) in central tendency. To test for equality of means with the  $t$ -test, use the dataset `example3`, and enter the following `mrpp` command.

```
> out <- mrpp(example3, expon = 2, c.form = 2)
```

The `expon = 2` option causes `mrpp` to compute squared Euclidean distances (`expon = 1` is the default value and specifies Euclidean distance). The `c.form = a` option specifies how the intragroup distances are to be averaged. If `c.form = 2` is specified, then the analysis mimics the classical parametric  $t$ -test, where the group distances are weighted by the relative degrees of freedom. If `c.form = 1` then the intragroup distances are weighted by relative group

size, then averaged to arrive at delta. This is the default value. In this example, since the group sizes are equal, the choice of `c.form` does not matter. In general choose `c.form = 2` and `expon = 2` to calculate a test that mimics the classical parametric *t*- and *F*-tests for univariate data and Hotelling's *T*-square or MANOVA for multivariate data. Here are the results of the above `mrpp` command:

```
> summary(out)
      Multi-Response Permutation Procedure (MRPP)

Call:
mrpp(data = example3, expon = 2, c.form = 2)

  Grouping Variable : group
  Response Variables: response

Specification of Analysis:
  Number of Observations: 30
  Number of Groups       : 2
  Distance Exponent      : 2
  Weighting Factor: (n(I)-1)/sum(n(I)-1) = C(I) = 2

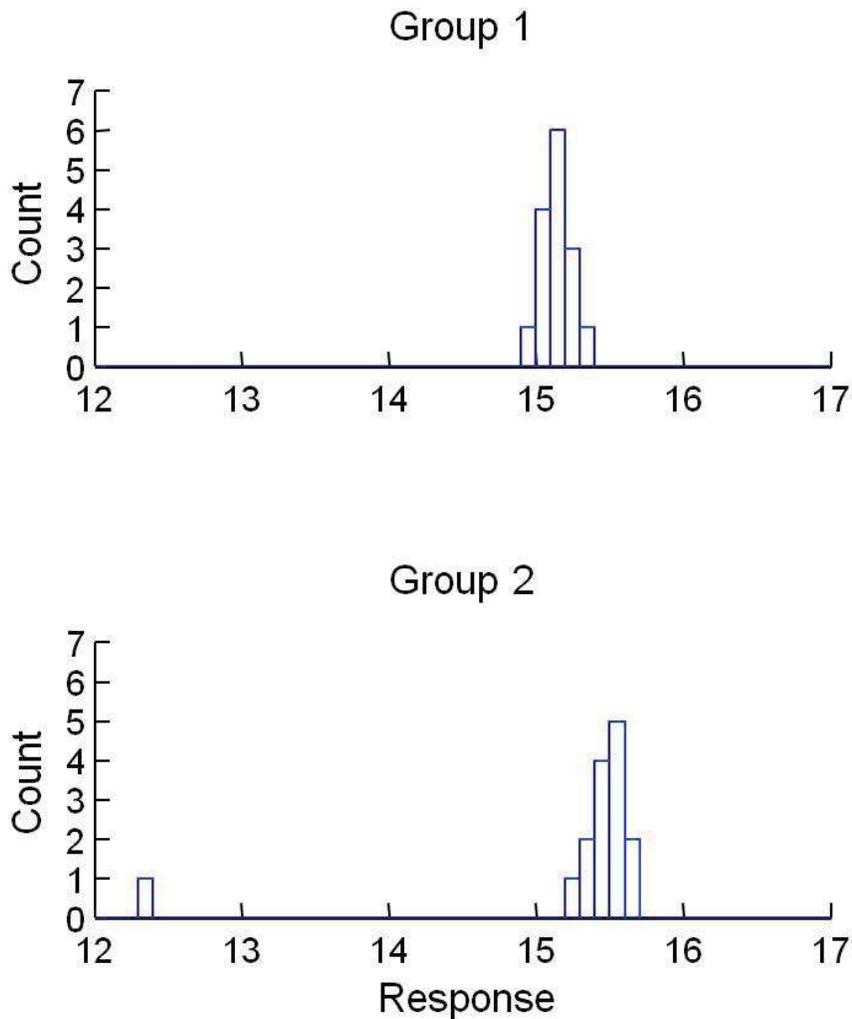
Group Summary:
  Group Value  Group Size  Group Distance
1             15           0.02133
2             15           0.02705

Results:
  Delta Observed      : 0.02419
  Delta Expected      : 0.08083
  Delta Variance      : 1.563e-05
  Delta Skewness      : -2.565

  Standardized test statistic : -14.32
  Probability (Pearson Type III)
of a smaller or equal delta : 1.926e-06***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The very small *P*-value (0.0000019) indicates that these two samples are unlikely to come from populations with the same mean, that is, they are different. The two sample *t*-test based on normal theory also gives a very low *P*-value for these data ( $P < 0.000001$ ).

Now consider the same data, but with one difference, a change in one of the 30 data values (fig. 4).



**Figure 4.** Two-group example from Mielke, 1986 (one outlier in Group 2).

To compare these samples use the dataset `example4` and issue the following `mrpp` command.

```
> out <- mrpp(example4, expon = 2, c.form = 2)
```

Here are the results:

```
> summary(out)
```

```
Multi-Response Permutation Procedure (MRPP)
```

```
Call:
```

```
mrpp(data = example4, expon = 2, c.form = 2)
```

```
Grouping Variable : group
```

```
Response Variables: response
```

```
Specification of Analysis:
```

```
Number of Observations: 30
```

```
Number of Groups      : 2
```

```
Distance Exponent     : 2
```

```
Weighting Factor: (n(I)-1)/sum(n(I)-1) = C(I) = 2
```

Group Summary:

Group Value	Group Size	Group Distance
1	15	0.02133
2	15	1.336

Results:

Delta Observed	:	0.6785
Delta Expected	:	0.6643
Delta Variance	:	0.0002558
Delta Skewness	:	-0.9893

Standardized test statistic	:	0.8879
Probability (Pearson Type III) of a smaller or equal delta	:	0.8144

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Now the  $P$ -value is quite large (0.81) indicating that it is likely that these samples come from the same population, that is, there is no difference between the means of the groups. The variances of the two groups differ considerably as evidenced by the average within group distance (when squared Euclidean distances are used this value is twice the variance). The medians are still 15.10 and 15.40, respectively, but the means now are 15.09 and 15.23, respectively. The parametric two-sample  $t$ -test also results in a large  $P$ -value (0.54). The reason for the discrepancy in results for data in which only one value is changed is the use of squared distance. In the squared Euclidean distance analysis space the distance of the outlier from the bulk of the data is exaggerated because it is squared. Now compare the results of analyzing the data of Example 4 in a space corresponding to the geometric space of the data itself. Issue the following command.

```
> out <- mrpp(example4, expon = 1, c.form = 1)
```

which, since these are the default values, is equivalent to

```
> out <- mrpp(example4)
```

Here are the results.

```
> summary(out)
```

Multi-Response Permutation Procedure (MRPP)

Call:

```
mrpp(data = example4)
```

Grouping Variable :	group
Response Variables:	response

Specification of Analysis:

Number of Observations:	30
Number of Groups	: 2
Distance Exponent	: 1
Weighting Factor: $n(I)/\sum(n(I))=C(I)=$	1

Group Summary:

Group Value	Group Size	Group Distance
1	15	0.1162
2	15	0.5314

Results:

Delta Observed	:	0.3238
----------------	---	--------

```

Delta Expected          : 0.4184
Delta Variance          : 6e-05
Delta Skewness          : -2.369

Standardized test statistic : -12.21
Probability (Pearson Type III)
of a smaller or equal delta : 6.262e-06***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Now the resulting  $P$ -value (0.0000063) is in line with the results obtained from the data without the single aberrant value. This is a demonstration of the sensitivity of variance- (squared Euclidean distance) based statistics and estimates of means to even a single outlying value. Estimates of medians and statistics based on absolute deviations (Euclidean distance) are far less sensitive to outlying data observations (Mielke and Berry, 2001).

Here is another example of how it is possible to get varying statistical results by methods that differ in their underlying geometry. The distance and elevation changes (in meters) for male and female blue grouse (*Dendragapus obscurus*) migrating from where they were marked on their breeding range to their winter range are given in the dataset `bgrouse` and are plotted in figure 5 (data from Cade and Hoffman, 1993). Generally the males seem to migrate farther and higher than the females and distance moved and elevation change are correlated ( $r = 0.71$ ). Here we will change `sex` in the `bgrouse` dataset to factor which is an option available in the R Blossom but not in the original Fortran program.

```

> BlueGrouse <- bgrouse
> BlueGrouse$sex <- as.factor(c("Female", "Male")[(BlueGrouse$sex ==
+ 3)+1])

```

To test gender differences in both distance and elevation, the multivariate parametric test is Hotelling's  $T^2$ , which gives  $P = 0.033$  for  $F = 4.145$  with  $df = (2, 18)$ , indicating some evidence of a difference in the bivariate means (*males* = 13388.9, 493.0; *females* = 5966.7, 231.66, distance and elevation respectively). To perform a permutation version of Hotelling's  $T^2$ , you would issue the following commands:

```

> out <- mrpp(BlueGrouse, hotelling = TRUE, exact = TRUE, expon =
+ 2, c.form = 2)

```

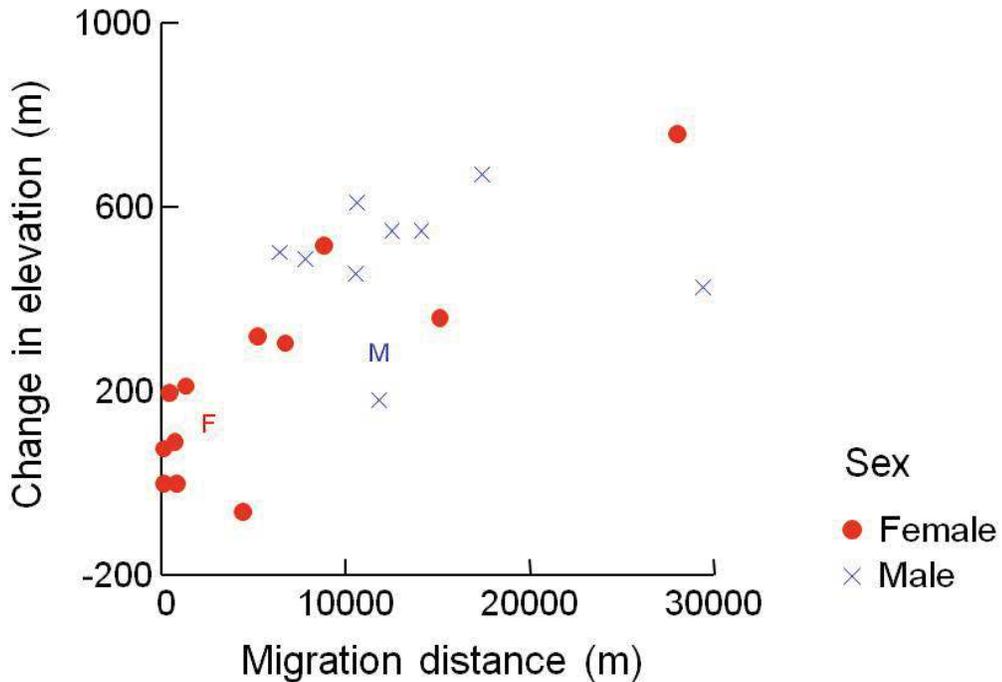
or equivalently by specifying the variables and grouping structure

```

> out <- mrpp(variables = c(distance, elev), group = sex, data = BlueGrouse,
+ hotelling = TRUE, exact = TRUE, expon = 2, c.form = 2)

```

where the options `hotelling = TRUE` indicated Hotelling's variance/covariance standardization of the multiple dependent variables, `expon = 2` requests squared Euclidean distances, and `c.form = 2` requests that groups be weighted by their relative degrees of freedom, and `exact = TRUE` requests a complete enumeration of all possible permutations for computing  $P$ -values.



**Figure 5.** Migration distance and elevation change for 9 male and 12 female blue grouse (from Cade and Hoffman, 1993). M and F denote bivariate medians for males and females, respectively.

Here are the results:

```
> summary(out)
Exact Multi-Response Permutation Procedure (EMRPP)

Call:
mrpp(variables = c(distance, elev), group = sex, data = BlueGrouse,
      expon = 2, c.form = 2, hotelling = TRUE, exact = TRUE)

Grouping Variable : sex
Response Variables: distance elev

Specification of Analysis:
Number of Observations: 21
Number of Groups      : 2
Distance Exponent     : 2
Weighting Factor      : (n(I)-1)/sum(n(I)-1) = C(I) = 2

Group Summary:
Group Value  Group Size
Female      12
Male        9

Hotelling's Commensuration Applied
Variance/Covariance Matrix:

Variable 1 : distance 1.412e+09 2.840e+07
Variable 2 : elev    2.840e+07 1.134e+06
```

```

Results:
  Delta Observed          : 0.1773
  Probability (Exact)
  of a smaller or equal delta: 0.02963*
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Notice that there is little difference between the  $P$ -values for the permutation (0.030) and parametric normal theory (0.033) versions of Hotelling's  $T^2$  for this data. Now if we want to analyze these data in the more natural Euclidean distance space, we can issue the following commands:

```
> out <- mrpp(variables = c(distance,elev),group = sex,data = BlueGrouse, +
exact = TRUE)
```

which uses the default average Euclidean distance of each variable, ignoring the group structure, to standardize the variables so that they have an average pairwise Euclidean distance  $(\Delta_{i,j}) = 1.0$ . Although distances and elevation changes are in the same units (meters) so that we might consider not commensurating the variables (`commens = FALSE`), there is some correlation between distance moved and elevation change so that it is possible that commensuration will provide more powerful hypothesis tests (Mielke and Berry, 1999, 2001). Here are the results:

```
> summary(out)
      Exact Multi-Response Permutation Procedure (EMRPP)

Call:
mrpp(variables = c(distance, elev), group = sex, data = BlueGrouse,
      exact = TRUE)

  Grouping Variable : sex
  Response Variables: distance elev

Specification of Analysis:
  Number of Observations: 21
  Number of Groups       : 2
  Distance Exponent      : 1
  Weighting Factor       : n(I)/sum(n(I))=C(I) = 1

Group Summary:
  Group Value  Group Size
  Female      12
  Male        9

Variable Commensuration Summary
  Variable Name  Average Distance (Euclidian if V=1)
  distance      9265
  elev          279.2

```

```

Results:
  Delta Observed          : 1.257
  Probability (Exact)
  of a smaller or equal delta: 0.003167**
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The same analysis but without any commensuration (`commens = FALSE`) produced a  $P = 0.008$ , over twice the size of the above analysis with average Euclidean distance

commensuration. Notice that the  $P$ -value with the MRPP statistic based on Euclidean distances (`expon = 1`) and average Euclidean distance commensuration is an order of magnitude smaller ( $P = 0.003$ ) than for the permutation version of Hotelling's  $T^2$  ( $P = 0.030$ ) based on squared Euclidean distances (`expon = 2`) and the variance/covariance commensuration. There are several contributing factors. Notice, that the bivariate medians for males and females in figure 5 indicated that the centroids of the groups were shifted in the same direction as the correlation between distance (`dist`) and elevation change (`elev`). Simulations conducted by Mielke and Berry (1999) demonstrated that the average Euclidean distance commensuration of bivariate variables provided greater power than the variance/covariance standardization when the group structure was shifted parallel to the covariance structure of the two variables. Furthermore, since the MRPP comparisons with `expon = 1` focus on shifts in the bivariate medians which were separated by 9,271.6 m rather than shifts in the bivariate means which were only separated by 7,426.8 m, there was a larger estimated effect size for the Euclidean distance compared to the squared Euclidean distance analysis. For these data, the analysis based on Euclidean distances and bivariate medians was more powerful with greater estimated effect sizes (shift in bivariate medians). When the groups are shifted orthogonal to the covariance structure of the dependent variables, then MRPP analyses with Hotelling's variance/covariance standardization (option `hotelling = TRUE`) and `expon = 1` can be more powerful. The bivariate medians for the blue grouse movements in figure 5 were estimated by giving the following command:

```
> out <- medq(data = BlueGrouse)
```

where the `save.test = TRUE` option stores the distance between each observation and its group bivariate median these can be later extracted using the command `Dist2mvm(output)`. This can be used for additional analysis or graphing. The output are:

```
> summary(out)
      2-Dimensional Median and Distance Quantiles

Call:
medq(data = BlueGrouse)

Specification of Analysis:

      Grouping Variable      : sex
      Number of Report Variables: 2
      Report Variables       : distance, elev

      Number of Observations  : 21
      Number of Groups       : 2
-----

Results for Group Value: Female

      Observations in Group      : 12
      Iterations to Solution     : 500
      Solution Tolerance         : 1.6e-11

      Within Group Median Coordinates for Variables
      Variable Name  Multivariate Median Coordinate
      distance      2526.8
      elev          139.4

      2-Dimensional Distance From Median Quantiles
      Average Distance to Multivariate Median: 5405
      Quantile                               Distance from Median
```

0	(Minimum)	1229
0.05		1229
0.1		1732
0.25		1884
0.5	(Median)	2429
0.75		6285
0.9		12575
0.95		25481
1	(Maximum)	25481

---

Results for Group Value: Male

```
Observations in Group      : 9
Iterations to Solution     : 91
Solution Tolerance        : 1.6e-11
```

```
Within Group Median Coordinates for Variables
Variable Name  Multivariate Median Coordinate
distance      11797.2
elev          292.2
```

```
2-Dimensional Distance From Median Quantiles
Average Distance to Multivariate Median: 4260
Quantile      Distance from Median
0             (Minimum)    109.2
0.05          109.2
0.1           109.2
0.25         1238.6
0.5          (Median)    2317.1
0.75         5401.3
0.9          17603.3
0.95        17603.3
1           (Maximum)    17603.3
```

The bivariate median coordinates are given for the two variables (`dist` and `elev`), and summary quantiles are provided for the distances between observations and the bivariate median for each group. The average distances to the bivariate median differ for males (4,260.3) and females (5,404.6), suggesting that there may be dispersion differences being detected by the MRPP analysis as well as shifts in bivariate medians. It is possible to test for equality of multivariate dispersions using a permutation version of a modification of Van Valen's (1978) test; the effect of the shift in group centroids removed are made with the multivariate medians rather than the multivariate means. This is accomplished for the blue grouse movements by performing a permutation version of the 2-sample *t*-test on the distances from the bivariate medians (variable `Dist2MVM`) by sex in the output from the previous command:

```
> BgDist2mvm <- Dist2mvm(out)

> out <- mrpp(variables = Dist2MVM, group = sex, data = BgDist2mvm, exact =
+ TRUE, expon = 2, c.form = 2)
```

The output below suggests there is little statistical support for dispersion differences.

```
> summary(out)
Exact Multi-Response Permutation Procedure (EMRPP)
```

Call:

```

mrpp(variables = Dist2MVM, group = sex, data = BgDist2mvm, expon = 2,
      c.form = 2, exact = TRUE)

  Grouping Variable : sex
  Response Variables: Dist2MVM

Specification of Analysis:
  Number of Observations: 21
  Number of Groups       : 2
  Distance Exponent      : 2
  Weighting Factor       : (n(I)-1)/sum(n(I)-1) = C(I) = 2

Group Summary:
  Group Value  Group Size
  Female      12
  Male         9

Results:
  Delta Observed          : 82227846
  Probability (Exact)
  of a smaller or equal delta: 0.7082
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Note that tests for equality of univariate dispersions based on the median modification of Levene's test (Good, 2000) can also be performed by requesting the univariate medians be calculated for each group with `medq`, saving the distances from the group medians, and then comparing those distances (`Dist2mvm(output)`) with the permutation version of the *t*-test implemented in `mrpp` by using the `expon = 2, c.form = 2` options. Testing for equality of dispersions after removing the effect of the estimated medians is one of those special cases where tests based on squared deviations (`expon = 2`) have better statistical performance than using Euclidean distances (`expon = 1`).

Because the sample size is only 21 for the blue grouse data, all the examples used the optional `exact = TRUE` enumeration of all permutations to compute probabilities. This is not practical to do with larger sample sizes and by default `mrpp` would use the Pearson Type III moments approximation. The following command yields the default approximation:

```

> out <- mrpp(c(distance,elev),sex,data = BlueGrouse)
The output is:
> summary(out)
  Multi-Response Permutation Procedure (MRPP)

```

```

Call:
mrpp(variables = c(distance, elev), group = sex, data = BlueGrouse)

  Grouping Variable : sex
  Response Variables: distance elev

Specification of Analysis:
  Number of Observations: 21
  Number of Groups       : 2
  Distance Exponent      : 1
  Weighting Factor       : n(I)/sum(n(I))=C(I)= 1

```

```

Group Summary:

```

Group Value	Group Size	Group Distance
Female	12	1.396
Male	9	1.072

Variable Commensuration Summary

Variable Name	Average Distance (Euclidian if V=1)
distance	9265
elev	279.2

Results:

Delta Observed	:	1.257
Delta Expected	:	1.513
Delta Variance	:	0.002706
Delta Skewness	:	-2.098
Standardized test statistic	:	-4.904
Probability (Pearson Type III) of a smaller or equal delta	:	0.002983**

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Alternatively, we can approximate the probabilities by Monte Carlo resampling with the command:

```
> out <- mrpp(c(distance,elev),sex,data = BlueGrouse,number.perms = 10000)
```

where the option `number.perms = 10000` specifies that 9,999 random samples + the 1 observed test statistic are to be used to approximate the probabilities. The output is:

```
> summary(out)
```

Multi-Response Permutation Procedure (MRPP)

Call:

```
mrpp(variables = c(distance, elev), group = sex, data = BlueGrouse,
      number.perms = 10000)
```

Grouping Variable : sex  
Response Variables: distance elev

Specification of Analysis:

Number of Observations:	21
Number of Groups	: 2
Distance Exponent	: 1
Weighting Factor: $n(I)/\sum(n(I))=C(I)=$	1

Group Summary:

Group Value	Group Size	Group Distance
Female	12	1.396
Male	9	1.072

Variable Commensuration Summary

Variable Name	Average Distance (Euclidian if V=1)
distance	9265
elev	279.2

Results:

Delta Observed	:	1.257
----------------	---	-------

```

      Probability (Resample)
of a smaller or equal delta : 0.0033**
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Notice that with these data that the exact, Pearson Type III approximation, and Monte Carlo resampling approximation all yield very similar  $P$ -values even though sample sizes were only  $n = 9$  and  $n = 12$ .

If the data given to Blossom have been rank transformed (substituting the original values by their rank order), then MRPP can be used to emulate some well known nonparametric rank tests. Using ranks combined with the selection of `expon = 2` and `c.form = 2` produces these analyses. Analyze the dataset `example4`, which has been rank transformed in the dataset `ex4rank`, with a permutation version of the Mann-Whitney-Wilcoxon test as follows.

```

> out <- mrpp(variables = rank, group = group, data = ex4rank, expon = 2, c.form
+ = 2)
> summary(out)

```

```

      Multi-Response Permutation Procedure (MRPP)

Call:
mrpp(variables = rank, group = group, data = ex4rank, expon = 2,
      c.form = 2)

      Grouping Variable : group
      Response Variables: rank

Specification of Analysis:
      Number of Observations: 30
      Number of Groups      : 2
      Distance Exponent     : 2
      Weighting Factor: (n(I)-1)/sum(n(I)-1) = C(I) = 2

Group Summary:
      Group Value  Group Size  Group Distance
1                15          43.2
2                15          103.1

Results:
      Delta Observed      : 73.17
      Delta Expected     : 151.9
      Delta Variance     : 55.34
      Delta Skewness     : -2.572

      Standardized test statistic : -10.58
      Probability (Pearson Type III)
of a smaller or equal delta : 4.006e-05***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

If there are more than three groups the test is analogous to the Kruskal-Wallis one-way analysis of variance by ranks. Note that both these tests are for univariate data (one response variable), but MRPP also is able to analyze multivariate data (ranked or unranked) as well, offering a generalization of these tests. Further, the approximation used by MRPP is more accurate than the normal approximation used by the classical rank tests, since it uses the

skewness of the probability distribution in the Pearson Type III approximation. Of course, it is also possible to approximate the probabilities with the Monte Carlo resampling option. Since these tests use `expon = 2` and `c.form = 2`, they are not congruent with the data space. Use the default values of `expon` and `c.form` to produce a congruent analysis. Thus besides generalizing some standard nonparametric tests to multiple dependent variables, MRPP adds congruent Euclidean distance variants to the statistical repertoire.

The `max.dist = num` (truncation) option, if given in the `mrpp` command, causes the MRPP analysis to replace interobject distances  $\Delta_{i,j}$  greater than the truncation value (call it  $B$ ) with the truncation value ( $\Delta_{i,j} = \Delta_{i,j} : \Delta_{i,j} < B; \Delta_{i,j} = B : \Delta_{i,j} \geq B$ ). For example,

```
mrpp(variables = c(var1, var2), group = group, max.dist = 55)
```

will replace distances greater than 55 with 55 in the permutation calculations. This is useful for detecting pattern and group clustering where one (or more) of the groups itself clusters in more than one region of the analysis space and another group is distributed uniformly or randomly in the same space. The truncation value (for example, 55) specified is the average diameter of the sub-clusters. Data plotting and experimentation with truncation values are advised. Examples where truncation is useful include: One kind of archeological artifact may be found in two distinct areas of a site while another artifact type is found scattered throughout the site. Clumping of plants in a homogeneous site or pattern of habitat types within a landscape are detectable with a truncated MRPP analysis (Reich and others, 1991). For further information see Mielke (1991).

The `has.excess` or `excess.value` option allows for several comparisons not possible with other statistical procedures. MRPP takes data that, before analysis, are classified into groups. In the usual case the groups represent comparable levels of classification (for example, male-female; treatments  $a$ ,  $b$ , and  $c$ ; or before and after observations). But in some cases one of the groups may not be comparable to the other groups of interest. This happens for example when one group is considered miscellaneous or otherwise contains unclassifiable objects. When such a group exists it may, in MRPP, be treated as an excess group. Since the concept of an excess group is not dealt with by most familiar statistical methods, a few examples will help clarify the idea.

In a study of the spatial distribution of artifacts in an archeological site Berry and others (1983) note that many times artifacts cannot readily be classified. A particular artifact may be anomalous, lack sufficient defining characteristics, or be broken or too worn to be classifiable. Such objects are definitely artifacts and may contain information, yet treating such a class on equal footing with other well defined artifact classes seems inappropriate. Investigators usually have the choice of excluding such miscellaneous classes from analysis or including them and risking bias in results or interpretation. MRPP gives the additional choice of including the excess group, but without elevating its status to that of the other groups. The observations of the excess group are treated as background noise, against which the observations on the other groups are analyzed.

Another example of the use of an excess group concerns the presence of higher lead concentration in soils near the center of a city (Mielke and others, 1983). The locations ( $x$  and  $y$  spatial coordinates) of high concentration soil samples (greater than or equal to median) were compared with the locations of all samples, low and high concentration, to determine whether higher concentrations of lead are associated with the city center.

In the excess group MRPP with a group of size  $n$  and an excess group of size  $m$  an intragroup average distance is computed for each possible combination of  $n$  observations out of the  $n + m$  possible observations. These values comprise the distribution of the test statistic, delta, to which is compared the actual intragroup distance.

The excess group can be implemented in comparisons of used versus available resources for a particular organism in a design where a random sample of resources is obtained and then presence (used) and absence (unused) observed. The used habitats are alike in that they all share the features necessary for the organism's survival. But the unused habitats may not form such a unitary group, some may be suitable for the organism and just happen not to be used, others may not be suitable at all, and among these some may not be suitable for lack of one requirement and others for lack of another requirement.

Here is an example comparing used versus available blue grouse habitat described by the basal area measurements of four kinds of trees present in stands on winter range (data from Cade and Hoffman, 1990). Note that the  $n = 16$  forest stands measured are an exhaustive and exclusive partitioning of the finite population of habitats studied (that is no random sampling assumptions apply).

```
> out <- mrpp(habitat, expon = 1, c.form = 1, interv = 0.0, has.excess = TRUE,
+ excess.value = 2, commens = FALSE)
```

Here are the results:

```
> summary(out)
Multi-Response Permutation Procedure (MRPP)

Call:
mrpp(data = habitat, expon = 1, c.form = 1, commens = FALSE, interv =
0, has.excess = TRUE, excess.value = 2)

Grouping Variable : use
Response Variables: dfir junip aspen other

Specification of Analysis:
Number of Observations: 16
Number of Groups      : 1
Distance Exponent     : 1
Weighting Factor: n(I)/sum(n(I))=C(I)= 1

Group Summary:
Group Value  Group Size  Group Distance
1            12         9.168
2*           4*

Variables are not commensurated
* Excess Group

Results:
Delta Observed      : 9.168
Delta Expected      : 10.08
Delta Variance      : 1.268
Delta Skewness      : -0.5313

Standardized test statistic : -0.8079
Probability (Pearson Type III)
of a smaller or equal delta : 0.1994
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In this example the used habitats do not seem to differ ( $P = 0.200$ ) in tree basal area from the available (that is, used plus unused) habitats. `commens = FALSE` was selected for no variable commensuration because tree basal areas were all in the same units (square meters/ha) and occurred at the same scale (tens of square meters/ha). However, there is some covariation

among the basal areas, so commensurating them with the average Euclidean distance may be desirable. Use of the average Euclidean distance here leads to even less difference with an exact  $P = 0.896$ .

The `interv = num` option allows an analysis to be conducted on univariate circular data such as time or compass orientation. This analysis recognizes that there are no endpoints to the measurement scale. Distances between replicates used in the ARC analyses are the shorter of the 2 possible distances around the circular distribution, that is  $\min(|x_i - x_j| \text{ and } \text{ARC} - |x_i - x_j|)$ . The `interv = num` specifies the number of units in the circular distribution so that input data can be standardized to values on a unit circle. The `interv = num` command submits the standardized data to an MRPP program configured for circular distributions.

As an example, consider an analysis of the orientation of movements of striped newts (*Notophthalmus peristriatus*) immigrating to and emigrating from Breezeway Pond, Florida in 1985–1990 (Dodd and Cade, 1998). Figure 6 presents the angular orientation of 585 females immigrating to and 564 emigrating from the pond that were captured in pitfall buckets inside and outside of a drift fence surrounding the pond.

Implement the arc-distance analysis with the following commands:

```
> out <- mrpp(variables = angle, group = ei, data = npof, interv = 360)
```

The grouping variable `ei` has 1's for emigrating and 2's for immigrating females. Here are the results of this analysis:

```
> summary(out)
```

```
Multi-Response Permutation Procedure (MRPP)
```

```
Call:
```

```
mrpp(variables = angle, group = ei, data = npof, interv = 360)
```

```
Grouping Variable : ei
Response Variables: angle
```

```
Specification of Analysis:
```

```
Number of Observations: 1149
Number of Groups       : 2
Distance Exponent     : 1
Weighting Factor: n(I)/sum(n(I))=C(I)= 1
ARC distances used    : 360
Intervals in unit circle
```

```
Group Summary:
```

Group Value	Group Size	Group Distance
1	585	89.17
2	564	89.53

```
Results:
```

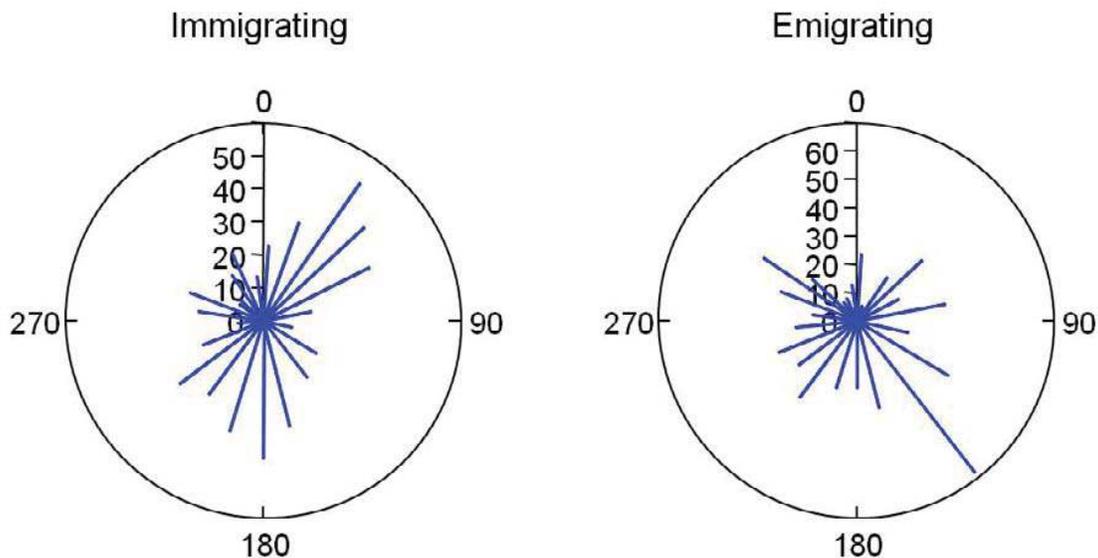
```
Delta Observed          : 89.35
Delta Expected          : 89.74
Delta Variance          : 0.004177
Delta Skewness          : -1.969
```

```
Standardized test statistic : -6.092
Probability (Pearson Type III)
of a smaller or equal delta : 0.0007966***
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The ARC (*interv*) analyses indicated that immigration and emigration orientation of the striped newts differed ( $P = 0.008$ ). More females immigrated to the northeast and southwest, whereas more emigrated from the southeast and northwest. The arc-distance analyses with *mrpp* are likely to be better than the more conventional Watson's test, especially useful when comparing circular distributions that have unequal angular variation or that are multimodal (Mielke and Berry, 2001). The *interv* option in Blossom is intended to be used with any univariate cyclical data (angular orientation, days of the year, hour of the day); more complicated transformations are possible for spherical data and combinations of scalar and circular data (see Mielke, 1986; Mielke and Berry, 2001).



**Figure 6.** Pattern of immigration and emigration for female striped newts at Breezeway Pond, Florida, 1985–1990. Length of the lines indicates number of newts counted in pitfalls (data from Dodd and Cade, 1997).

### Multiresponse Randomized Block Procedure (MRBP)

Data from a complete randomized block design or data that can be construed in a treatment by block manner can be analyzed by specifying a blocking variable in the *mrbp* command. The following data (Mielke and Iyer, 1982) are from a mine reclamation study comparing oven-dried biomass (gm) of three species of shrubs in six treatments (1 = no fertilizer, 2 = low fertilizer, 3 = high fertilizer, 4 = mulch and no fertilizer, 5 = mulch and low fertilizer, and 6 = mulch and high fertilizer) by three blocks (different plots). A complete randomized block analysis is done with the following commands which can be specified equivalently in four different ways:

```
> a <- mrbp(data = mrbp1)
> a <- mrbp(mrbp1)
> a <- mrbp(variables = c(spp1, spp2, spp3), group = trtmt, block = block, data
+ = mrbp1)
> a <- mrbp(c(spp1, spp2, spp3), trtmt, block, data = mrbp1)
```

Here are the results of the MRBP analysis with the default multivariable commensuration and block alignment. Note, the original analysis by Mielke and Iyer (1982) did not

commensurate or align the data and you can duplicate their analysis by using the options align = FALSE, commens = FALSE.

```
> summary(a)
```

Multiresponse Randomized Block Procedure (MRBP)

Call:

```
mrbp(variables = c(spp1, spp2, spp3), group = trtmt, block = block,  
      data = mrbp1)
```

Specification of Analysis:

```
Response Variables      : spp1 spp2 spp3  
Number of Observations : 18  
Grouping Variable      : trtmt  
  Number of Groups     : 6  
Blocking Variable      : block  
  Number of Blocks     : 3  
Distance Exponent      : 1  
Data Aligned  
Hotelling's Commensuration Applied
```

Group Summary:

Group Value	Group Size
1	3
2	3
3	3
4	3
5	3
6	3

Block Alignment Summary:

Block Value	Variable Name	Alignment Value
1	spp1	6.500
	spp2	3.165
	spp3	2.170
2	spp1	9.915
	spp2	1.165
	spp3	2.665
3	spp1	6.250
	spp2	1.915
	spp3	2.415

Variable Commensuration Summary

Variable Name	Average Euclidean Distance
spp1	7.602
spp2	3.107
spp3	0.9006

Results:

```
Delta Observed          : 1.785  
Delta Expected          : 1.98  
Delta Variance          : 0.02093  
Delta Skewness          : -0.3897  
Agreement measure among blocks: 0.09861  
Standardized test statistic : -1.35  
Probability (Pearson Type III)
```

```

of a smaller or equal delta : 0.09499.
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The  $P$ -value is 0.095, indicating weak evidence to reject the null hypothesis of no treatment effect. The original analysis without commensurating and aligning variables gave  $P = 0.067$ . Because of the small number of blocks and treatments it is possible to conduct this analysis by complete enumeration of the permutation distribution by using the option `exact = TRUE`. This yields  $P = 0.099$ . The Monte Carlo resampling approximation also is available for problems with large block and treatment structure.

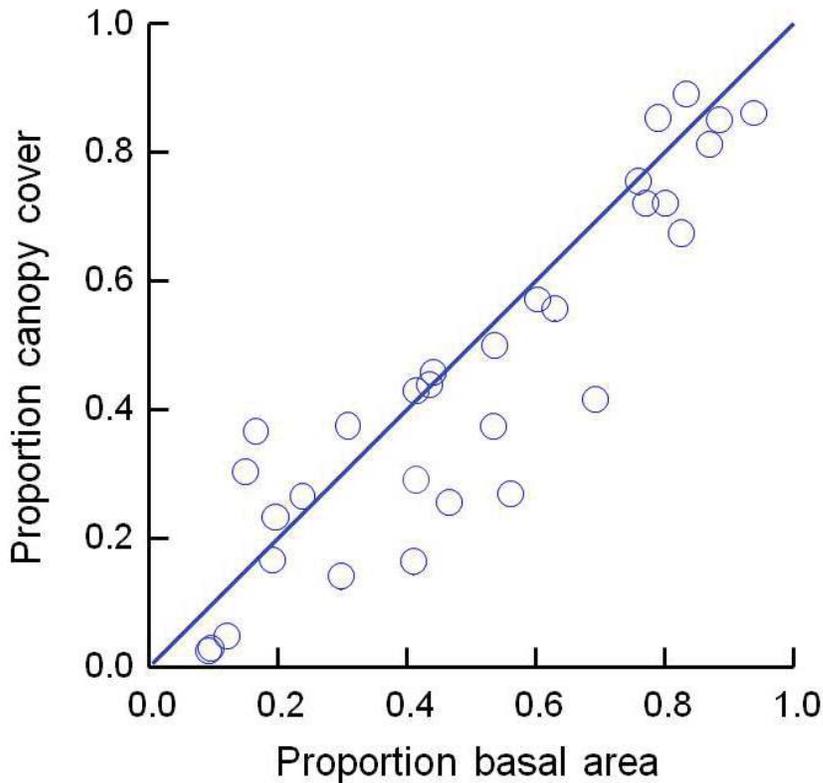
The data used in the MRBP test have been aligned so that the medians of the blocks are all equal. The value chosen to align each block is selected to make the block medians all equal to zero. If there is more than one response variable then Blossom adjusts or commensurates variables by their average Euclidean distance by default as in `mrpp`. The block alignment values and variable commensuration values are reported.

It is possible to turn off one or both of the alignment and variable commensuration options. The `align = FALSE` in the `mrbp` command produces an analysis without data alignment. The `commens = FALSE` option in `mrbp` produces an analysis without multivariate commensuration. These options can be important for special applications of MRBP. Here is an example command:

```
mrbp(variables = length, group = trtmt, block = block, align = FALSE)
```

Of course since only one variable, `length`, was specified, no variable commensuration is done. This option is especially useful when the blocked design is used not so much to detect treatment effects but to get a measure of the agreement among blocks. One use for this option is numerical model verification. Here blocks contain the predictions of one or more models and one block contains measured results. See Tucker and others (1989) for details. Agreement measures (1–observed delta/expected delta) based on Euclidean distances are generalizations of Cohen's kappa extended to multiple groups, multiple variables, and interval data (Berry and Mielke, 1988). The agreement measure based on squared Euclidean distances (`expon = 2`) applied to interval data is a linear transform of Pearson's correlation coefficient, that is, a probability value for a correlation coefficient based on a permutation argument can be obtained.

Here is an example analysis comparing measures of the proportion of basal area to the proportion of canopy cover of lodgepole pine (*Pinus contorta*) in 31 stands of subalpine forest in Colorado (fig. 7) (Cade, 1997).



**Figure 7.** Proportion of basal area and canopy cover for lodgepole pine in 31 stands (data from Cade, 1997). Solid line corresponds to perfect agreement.

The 31 sample plots are specified by the grouping variable `stand` and the proportion of either basal area or canopy cover are specified by the blocking variable `method`. The response variable for proportion lodgepole pine is `pctlcc`.

```
> out <- mrbp(agree2[,c(5,6,3)],align = FALSE)
```

Here are the results of the analysis:

```
> summary(out)
```

```
Multiresponse Randomized Block Procedure (MRBP)
```

```
Call:
```

```
mrbp(align = FALSE)
```

```
Specification of Analysis:
```

```
Response Variables      : pctlcc
Number of Observations  : 62
Grouping Variable       : stand
  Number of Groups      : 31
Blocking Variable       : method
  Number of Blocks      : 2
Distance Exponent       : 1
```

```
Group Summary:
```

```
Group Value  Group Size
      1         2
      2         2
```

```

3          2
4          2
5          2
6          2
7          2
8          2
9          2
10         2
11         2
12         2
13         2
14         2
15         2
16         2
17         2
18         2
19         2
20         2
21         2
22         2
23         2
24         2
25         2
26         2
27         2
28         2
29         2
30         2
31         2

```

Data are not aligned within blocks  
Variables are not commensurated

Results:

```

Delta Observed          : 0.09431
Delta Expected          : 0.3062
Delta Variance          : 0.001219
Delta Skewness          : -0.08266
Agreement measure among blocks: 0.692
Standardized test statistic : -6.067
Probability (Pearson Type III)
of a smaller or equal delta : 8.653e-09***

```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The agreement measure in this analysis (0.692) indicates that there is an average reduction in Euclidean distance between the proportions of basal area and canopy cover that is 69% greater than expected by chance and this differs from zero with  $P < 0.0001$ . The observed delta = 0.094 which indicates that the two proportionate measures of lodgepole pine (`pctlcc`) differed on average by 0.094 across all 31 stands (fig. 7). There was good but not perfect agreement between measures of the proportion of basal area and the proportion of canopy cover for characterizing the lodgepole pine contribution to the forest composition. Additional univariate agreement comparisons for subalpine fir (*Abies lasiocarpa*, `pctfcc`) and Engelmann spruce (*Picea engelmannii*, `pctsc`) are given in Cade (1997). A multivariate measure of agreement that considers all three species simultaneously given in Cade (1997) is performed with the command:

```
> agree <- agree2[,c(5,6,1,2,3)]
> out <- mrbp(agree,align = FALSE,commens = FALSE)
```

The results indicate that the average deviation between proportionate measures of basal area and canopy cover is 0.168 (observed delta) across the 31 stands for the three conifer species and the agreement measure indicates a 62% reduction in the observed deviation over that expected by chance.

```
> summary(out)
```

Multiresponse Randomized Block Procedure (MRBP)

Call:

```
mrbp(commens = FALSE, align = FALSE)
```

Specification of Analysis:

```
Response Variables      : pccscc pccfcc pccfcc
Number of Observations : 62
Grouping Variable       : stand
Number of Groups        : 31
Blocking Variable       : method
Number of Blocks        : 2
Distance Exponent       : 1
```

Group Summary:

Group Value	Group Size
1	2
2	2
3	2
4	2
5	2
6	2
7	2
8	2
9	2
10	2
11	2
12	2
13	2
14	2
15	2
16	2
17	2
18	2
19	2
20	2
21	2
22	2
23	2
24	2
25	2
26	2
27	2
28	2
29	2
30	2
31	2

Data are not aligned within blocks  
Variables are not commensurated

```

Results:
  Delta Observed          : 0.1681
  Delta Expected          : 0.4409
  Delta Variance          : 0.001553
  Delta Skewness          : -0.08762
  Agreement measure among blocks: 0.6186
  Standardized test statistic : -6.921
  Probability (Pearson Type III)
  of a smaller or equal delta : 1.161e-10***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

For information on other ways to align data useful for analyzing incomplete block and Latin square designs with MRBP see Fawcett (1990), Mielke and Iyer (1982), and Hodges and Lehmann (1962).

If `expon = 2` is chosen, then the univariate version of this test is a permutation version of analysis of variance for complete randomized blocks. Note that when `expon = 2` is used in an MRBP analysis that the blocks are self-aligning to a common mean and no alignment is required; analyses made with `mrbp` and `expon = 2` and the option `align = FALSE` should result in identical test statistics and *P*-values as when alignment is not turned off. Specification of the `c.form` (group averaging method) parameter has no effect, since group sizes have to be the same. Also the `excess.value` option is not supported for `mrbp` and is ignored. The `exact = TRUE` option is available only for some small block (<10) and group combinations. The Monte Carlo resampling approximation of *P*-values is available with the option `number.perms = num`.

If ranked data are used and `expon = 2` is specified, then the test (with one response variable) is functionally related to Friedman's nonparametric randomized block analysis.

## Permutation Tests for Matched Pairs (PTMP)

Matched pair tests can be performed by the MRBP command. Essentially the matched pairs test is a special case of the randomized block version MRBP with one or more response variables, two groups, and a blocking variable identifying pairs. Data of this sort can be analyzed by an MRBP command specified just like that for performing an MRBP. For example the sample dataset `paired1` contains one response (`response`), for two groups (`groups`), and with the paired members of each group indicated by a blocking variable (`pair`). Use this dataset and perform a matched pairs test by issuing the following command:

```

> a <- mrbp(paired1)

> ###Or equivalently
> a <- mrbp(variables = response, group = group, block = pair, data = paired1)

> summary(a)

```

```

  Multiresponse Randomized Block Procedure (MRBP)

```

```

Call:
mrbp(variables = response, group = group, block = pair, data =
paired1)

```

```

Specification of Analysis:
  Response Variables      : response
  Number of Observations  : 20
  Grouping Variable       : group

```

```

      Number of Groups      : 2
      Blocking Variable     : pair
      Number of Blocks     : 10
      Distance Exponent    : 1
      Data Aligned

```

Group Summary:

```

      Group Value  Group Size
           1         10
           2         10

```

Block Alignment Summary:

```

      Block Value  Variable Name  Alignment Value
           1      response      4.275
           2      response      3.340
           3      response      6.545
           4      response      3.070
           5      response      2.880
           6      response      8.190
           7      response      6.105
           8      response      5.065
           9      response      2.695
          10      response      0.790

```

Variables are not commensurated

Results:

```

      Delta Observed      : 1.211
      Delta Expected     : 1.887
      Delta Variance     : 0.02099
      Delta Skewness     : -1.984
      Agreement measure among blocks: 0.3581
      Standardized test statistic : -4.664
      Probability (Pearson Type III)
      of a smaller or equal delta : 0.003421**

```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

With one response variable and `expon = 2` specified on the command line, then this test mimics the *t*-test for matched pairs.

Sometimes it is convenient to structure paired data such that the values for each pair are given on a single line in the dataset with a separately named variable for the response of the first and of the second members of each pair. Blossom allows for this different data structure. Use the example dataset `paired2` and simply issue the following command:

```
> a <- ptmp(paired1, number.perms = 10000, save.test = TRUE)
```

```
> summary(a)
```

```
Permutation Tests for Matched Pairs (PTMP)
```

Call:

```
ptmp(data = paired1, number.perms = 10000, save.test = TRUE)
```

```

      Grouping Variable: group
      Blocking Variable: pair
      Response Variable: response

```

Specification of Analysis:

```
Number of Observations      : 20
Number of Groups            : 2
Number of Pairs             : 10
Number of Non-Zero Differences: 10
Distance Exponent          : 1
```

Group Summary:

```
Group Value  Group Size
1            10
2            10
```

Results:

```
Delta Observed              : 2.423
Probability (Resample)
of a smaller or equal delta : 0.0044**
```

---

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> ## or alternatively
> a <- ptmp(variables = c("first", "second"), data = paired2,
+ number.perms = 10000, save.test = TRUE)
```

The `ptmp` command is used when the observations are paired (next to each other) in the dataset. Thus the pairing is indicated by position not by a blocking variable. Also no grouping variable is specified because in PTMP there can only be two groups. The univariate observations for each group correspond to the columns named `first` and `second`. Note, this is a special dataset format useful only for PTMP, which is a univariate, two group, paired comparison, where the number of blocks equals the number of pairs.

Because the number of pairs in this dataset is less than 20 the  $P$ -value reported was obtained by exact enumeration of the permutation distribution (and thus differs slightly from the  $P$ -value given in the previous example). With more than 20 pairs an approximation with the Pearson Type III distribution is used by default or the Monte Carlo resampling option can be invoked with the option `number.perms = num`. Notice that the different test statistic structures produce an observed delta in PTMP that is exactly twice the observed test statistic for the same problem in MRBP. Also, data in PTMP are aligned to a median of 0 by the structure of the test statistic. It is possible to do a 1-sample comparison of data with a hypothesized parameter for central tendency (either median or mean) with PTMP by making one of the column variables equal to the hypothesized parameter and the other the observed data vector (Mielke and Berry, 2001). If the hypothesized parameter is a median and `ptmp` is implemented with `expon = 1` then this test is for a null hypothesis that the sample comes from a population with median equal to the specified value. If the hypothesized parameter is a mean and `ptmp` is implemented with `expon = 2` then this test is for a null hypothesis that the sample comes from a population with mean equal to the specified value.

Multivariate extensions of the 1-sample comparison are made by using the `mrbp` command and specifying the vector of hypothesized parameters for the multivariate median (mean) as one group, the observed vector as the second group, for each of  $n$  blocks comprising the sample. As an example, consider the data on ring-necked pheasant (*Phasianus colchicus*) habitat selection from Aebischer and others (1993: Appendix 1), where the percentage of home ranges in 5 habitat types—scrub (`scrub`), broadleaf woodlands (`broad`), conifer woodlands

(conifer), grasslands (grass), and crops (crop)—for 13 radio-marked birds was compared to the available percentages of these habitat types. Because these data are compositions with a unit sum constraint, Aebischer and others (1993) chose to analyze these data with log ratios in a MANOVA. We can perform a similar 1-sample analysis comparing the observed percentages of the habitat types for the 13 birds with the hypothesized available percentages in MRBP without resorting to log ratios (which are problematic when you have some zero proportions). The dataset `prefer` has 13 blocks (`bird`) for the grouping `use = 1` corresponding to the observations for the 13 birds, and the same 13 block values for the grouping `use = 0` corresponding to 13 replications of the hypothesized available percentage of the habitat types. Issue the commands:

```
> out <- mrbp(variables = c("scrub", "broad", "conifer", "grass", "crop"),
+ group = use, block = bird, data = prefer)
```

The following output indicated that the 13 pheasants were not using habitat types in proportion to their availability. Of course, it is possible to do a permutation version of the 1-sample MANOVA analysis on log ratios as done by Aebischer and others (1993), but the Euclidean distance statistics of MRPP avoid concerns about singular matrices with dependent variables having the unit sum constraint and ad hoc procedures needed to deal with zero proportions when transforming to log ratios.

We can compute the multivariate median for the percentage of the habitat types used by the 13 pheasants to compare with the hypothesized percentages by issuing the command:

```
> summary(out)
```

```
      Multiresponse Randomized Block Procedure (MRBP)
```

```
Call:
```

```
mrpb(variables = c("scrub", "broad", "conifer", "grass", "crop"),
      group = use, block = bird, data = prefer)
```

```
Specification of Analysis:
```

```
Response Variables      : scrub broad conifer grass crop
Number of Observations  : 26
Grouping Variable       : use
  Number of Groups      : 2
Blocking Variable       : bird
  Number of Blocks      : 13
Distance Exponent       : 1
Data Aligned
Hotelling's Commensuration Applied
```

```
Group Summary:
```

Group Value	Group Size
1	13
0	13

```
Block Alignment Summary:
```

Block Value	Variable Name	Alignment Value
1	scrub	11.410
	broad	5.600
	conifer	0.365
	grass	26.415
	crop	56.190
2	scrub	11.900
	broad	11.965
	conifer	0.365
	grass	26.615

	crop	49.145
3	scrub	5.770
	broad	7.480
	conifer	0.365
	grass	55.865
	crop	30.505
4	scrub	6.000
	broad	16.545
	conifer	0.365
	grass	32.535
	crop	44.540
5	scrub	3.815
	broad	19.760
	conifer	5.525
	grass	53.905
	crop	16.990
6	scrub	4.325
	broad	19.875
	conifer	5.420
	grass	53.385
	crop	16.990
7	scrub	3.780
	broad	20.235
	conifer	5.875
	grass	53.110
	crop	16.990
8	scrub	5.940
	broad	23.970
	conifer	0.365
	grass	52.720
	crop	16.990
9	scrub	6.430
	broad	31.195
	conifer	0.365
	grass	45.000
	crop	16.990
10	scrub	7.470
	broad	9.025
	conifer	0.365
	grass	66.135
	crop	16.990
11	scrub	8.790
	broad	20.895
	conifer	0.365
	grass	52.940
	crop	16.990
12	scrub	6.460
	broad	10.090
	conifer	0.365
	grass	66.080
	crop	16.990
13	scrub	4.375
	broad	14.655
	conifer	2.420
	grass	61.555
	crop	16.990

Variable Commensuration Summary

Variable Name	Average Euclidean Distance
scrub	4.993
broad	11.69
conifer	2.456
grass	15.07
crop	18.21

Results:

Delta Observed	:	2.238
Delta Expected	:	2.66
Delta Variance	:	0.008276
Delta Skewness	:	-1.165
Agreement measure among blocks:	:	0.1585
Standardized test statistic	:	-4.634
Probability (Pearson Type III) of a smaller or equal delta	:	0.001209**

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
> out2 <- medq(variables = c("scrub","broad","conifer","grass","crop"),
+ group=use,data = prefer)
> summary(out2)
```

5-Dimensional Median and Distance Quantiles

Call:

```
medq(variables = c("scrub", "broad", "conifer", "grass", "crop"),
      group = use, data = prefer)
```

Specification of Analysis:

Grouping Variable	:	use
Number of Report Variables:	:	5
Report Variables	:	scrub, broad, conifer, grass, crop
Number of Observations	:	26
Number of Groups	:	2

---

Results for Group Value: 0

Observations in Group	:	13
Iterations to Solution	:	1
Solution Tolerance	:	1.6e-13

Within Group Median Coordinates for Variables

Variable Name	Multivariate Median Coordinate
scrub	3.22
broad	9.23
conifer	0.73
grass	52.83
crop	33.98

5-Dimensional Distance From Median Quantiles

Average Distance to Multivariate Median: 0		
Quantile		Distance from Median
0	(Minimum)	0
0.05		0
0.1		0

0.25		0
0.5	(Median)	0
0.75		0
0.9		0
0.95		0
1	(Maximum)	0

---

Results for Group Value: 1

Observations in Group	:	13
Iterations to Solution	:	55
Solution Tolerance	:	1.6e-13

Within Group Median Coordinates for Variables

Variable Name	Multivariate Median Coordinate
scrub	7.616
broad	28.677
conifer	5.666
grass	54.042
crop	3.988

5-Dimensional Distance From Median Quantiles

Average Distance to Multivariate Median: 33.74

Quantile		Distance from Median
0	(Minimum)	6.621
0.05		6.621
0.1		7.164
0.25		10.466
0.5	(Median)	30.597
0.75		33.372
0.9		83.137
0.95		96.678
1	(Maximum)	96.678

The output indicated that the multivariate median vector for the proportions of habitats used is shifted towards a much higher proportion of broadleaf woodlands, moderately higher proportions of scrub and conifer woodlands, much lower proportions of crops, with little difference in the proportion of grasslands compared to available habitat types. Note that this summary doesn't recognize the blocked by animal nature of the design and could be made more appropriate by first taking differences between components of used and available habitat types by animal and then taking the multivariate medians of those differences.

### Multiresponse Sequence Procedure (MRSP)

The multiresponse sequence procedure `sp` command initiates a test of first-order serial dependency on univariate or multivariate response variables (Mielke, 1991). In this analysis of ungrouped data, the agreement measure (1- average Euclidean distance between ordered observations/average Euclidean distance among all possible pairs of observations) is a statistic describing first-order serial dependency. Significance of the null hypothesis of no first-order serial dependency is provided by the Pearson Type III approximation on the first three exact moments of the permutation distribution by default, optionally by exact enumeration for small samples, or by a Monte Carlo resampling procedure. In a univariate test, MRSP is analogous to the Durbin-Watson test. A permutation version of the Durbin-Watson test can be initiated by

selecting the option `expon = 2` for squared Euclidean distances. In a bivariate test of animal locations, where latitude and longitude coordinates are the two response variables, MRSP provides a Euclidean distance analogue of Schoener's  $t^2/r^2$  statistic (Solow, 1989), which is a nonmetric measure based on squared Euclidean distances. A permutation version of Schoener's  $t^2/r^2$  test can be implemented on bivariate data by specifying the `expon = 2` and `commens = FALSE`. MRSP obviously provides the possibility of evaluating first-order serial dependency of greater than two response variables. Blossom will commensurate (standardize) multiple variables to unit average Euclidean distance by default.

Example data of biweekly grouse locations during November through March (Cade and Hoffman, 1993) are in the dataset `blue162` and graphed below (fig. 8), where the numbers correspond to the temporal order of observations (variable `date`) for the response variables `lat` and `long`. Multiple observations at the same location are indicated by ordered values next to their location.

To implement an analysis of first-order serial dependency on these bivariate locational data, the following commands are issued:

```
> out <- sp(as.matrix(blue162[,4:5]), commens = FALSE)

> ##or equivalently as
> out <- sp(variables = c("lat", "long"), sequence = seq(1:12), data =
+ blue162, commens = FALSE)
```

Note that the same analysis on this dataset can be implemented without specifying the sequencing variable `date` because the dataset is ordered by `date` already and Blossom by default assumes the order in the dataset is the sequencing variable if none is specified.

Here are the results of this analysis:

```
> summary(out)
      Multiresponse Sequence Procedure (MRSP)

Call:
sp(data = blue162, commens = FALSE, sequence = seq(1:12), variables =
  c("lat", "long"))

Specification of Analysis:
  Number of Observations: 12
  Distance Exponent      : 1

  Variables are not commensurated

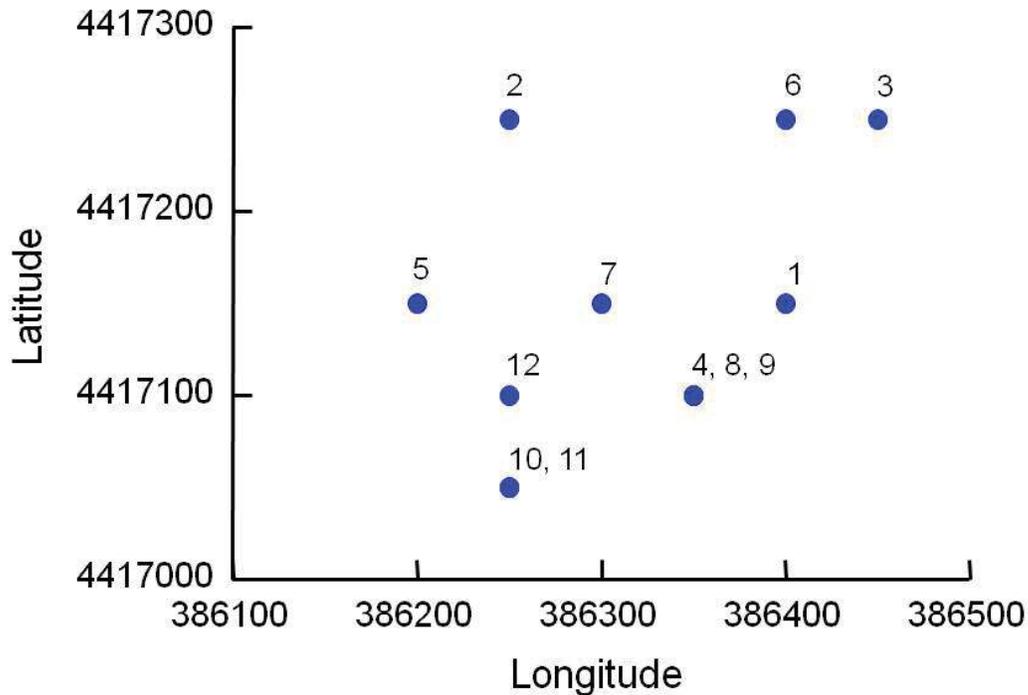
Results:
  Delta Observed: 119.7
  Delta Expected: 135.5
  Delta Variance: 252
  Delta Skewness: -0.2363

  Standardized test statistic : -0.9959

  Agreement measure           : 0.1167
  Probability (Pearson Type III)
  of a smaller or equal delta : 0.159
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The agreement measure (0.117) is interpreted as the percent reduction in average Euclidean distance between sequentially ordered values (observed delta) over that expected

without any order (expected delta). In this analysis there is little evidence of first-order serial pattern and the null hypothesis of no serial dependency has  $P = 0.159$ . The expected delta is the average Euclidean distance among the locations ignoring any serial dependence (135.5 m) and the observed delta is the average Euclidean distance between sequentially ordered locations (119.7 m); both these measures are useful summary statistics describing animal home ranges (Cade and Hoffman, 1993). See Mielke and Berry (2001) for descriptions of how to extend the sequence procedure to higher orders of serial dependence.



**Figure 8.** Latitude and longitude coordinates (m) for blue grouse (no. 162) locations on 12 dates November through March. Numbers next to locations correspond to order of dates for locations (note multiple dates at same locations).

### Least Absolute Deviation (LAD) and Quantile Regression

LAD regression differs from least squares (OLS) regression in that the sum of the absolute, not squared, deviations of the fit from the observed values is minimized to obtain estimates. LAD regression estimates the conditional median (0.5 regression quantile) of the dependent variable ( $y$ ) given independent variables ( $X$ ), and its generalization, regression quantiles, estimates the conditional quantile ( $\tau$ , where  $0 \leq \tau \leq 1$ ) of  $y$  given  $X$ . Since LAD does not use squared distances, it is an obvious companion to the MRPP which emphasizes Euclidean distances. Both LAD and MRPP satisfy the congruence principle (Mielke and Berry, 2001). Asymptotic distributional theory for testing procedures for LAD regression are found in Dodge (1987) and a concise, readable implementation is provided by Birkes and Dodge (1993). Cade and Noon (2003) is a primer on quantile regression for ecologists.

The `lad` command is used to compute a fit of one dependent response variable by one or more independent predictor variables. The parameters in a LAD regression are tested by using a test statistic that compares the proportionate reduction in sums of absolute deviations when

passing from a reduced to full parameter model (that is, a test statistic very similar to general F-tests in OLS regression). The drop in dispersion test statistic,  $T_{obs}$ , equals (sum of absolute deviations for reduced model–sum of absolute deviations for full model) / sum of absolute deviations for full model (Cade and Richards, 1996; Cade 2003, 2005). Large values of  $T_{obs}$  are evidence against the null hypothesis that the parameter(s) equal(s) zero. If all slope parameters are tested simultaneously against a reduced parameter model that includes only the intercept, then the reference permutation distribution for the test statistic  $T_{obs}$  is obtained by randomly sampling the  $n!$  permutations of the dependent variable to the matrix of independent variables as described by Manly (1991) and calculating  $T$  for each permutation. However, if only a subset of parameters is being tested (partial model tests), then the reference permutation distribution for the test statistic  $T_{obs}$  is obtained by randomly sampling the  $n!$  permutations of residuals from the reduced model to the matrix of independent variables and calculating  $T$  for each permutation, following Freedman and Lane (1983). Probabilities under the null hypothesis are given by (number of  $T \geq T_{obs} + 1$ )/number of permutations sampled. Extensive power simulations demonstrated that these procedures maintained nominal error rates under the null hypothesis well across a variety of error distributions and design configurations (correlated and uncorrelated independent variables) provided the error distributions are independent and identically distributed (Cade and Richards, 1996). Similar conclusions were reached for the same form of the test statistic used with OLS regression (Kennedy and Cade, 1996; Anderson and Legendre, 1999). The LAD permutation test is extended to any selected regression quantile (LAD is just 0.5 regression quantile) by replacing sums of absolute deviations in the test statistic computation with the appropriate sums of weighted absolute deviations used in regression quantile estimation (Cade and Richards, 2006).

Cade (2003, 2005) and Cade and Richards (2006) found that Type I error rates were improved when testing subsets of parameters in quantile regression models by deleting all but a single zero residual associated with the fit to  $p - q$  parameters under the null hypothesis, where  $p$  is the number of parameters in the full model and  $q$  is the number of parameters being tested. As this reduces the length of the residual vector so that it no longer conforms to the  $n \times p$  matrix  $\mathbf{X}$  of predictors, the corresponding number of rows of  $\mathbf{X}$  are randomly deleted at each permutation. This deletion of zero residuals and random deletion of rows of  $\mathbf{X}$  are done by default for this drop in dispersion permutation test. In addition, Cade (2003, 2005) and Cade and Richards (2006) found that anytime the null, reduced ( $p - q$ ) parameter model was constrained through the origin (no intercept), Type I error rates were improved by randomly recentering the residual vector since the residuals from the null model will no longer have zero associated with the specified quantile (or mean zero for OLS). This is implemented as a double permutation procedure where the first step at each iteration is to randomly recenter the selected quantile of the residual vector by a quantity generated as a random binomial for the specified quantile (for example, 0.90). A similar operation is done for OLS regression where the quantile = 0.50 is always used to generate random binomials. The second step at each iteration then (the doubling of permutations) permutes these randomly recentered residuals to the matrix  $\mathbf{X}$ . Because it is not always obvious when a model is constrained through the origin (for example, some weighted model tests will require this and some won't), we elected to make the double permutation scheme selected by an option of the hypothesis testing command `double.permutation = TRUE`.

If error distributions are not identical (heteroscedastic) then they must be transformed or weighted to be made approximately identical (homogeneous) (Cade and Richards, 1996; Cade, 2003, 2005; Cade and Richards, 2006). Cade and Noon (2003) and Cade and others (2006) discuss two weighting schemes, one where all quantiles have the same weights in a location-scale form of heterogeneity, and one where the weights must be estimated separately for the

selected quantiles in more general models of heterogeneity. When the weights are based on a function of the independent variables ( $\mathbf{X}$ ) many of the permutation hypothesis tests will implicitly constrain the null model through the origin and the double permutation procedure will be required to maintain correct Type I error rates (Cade, 2005; Cade and others, 2005; Cade and Richards, 2006).

As an alternative test for LAD and its generalization to regression quantiles, we provide a quantile rank score statistic that is less sensitive to heterogeneous error distributions (Koenker, 1994; Cade and others, 1999; Koenker and Machado, 1994). The permutation version of the quantile rank score test (Cade, 2003; Cade and others, 2006) maintains Type I error rates better than the asymptotic Chi-square distributional approximation (Koenker, 1994) at smaller  $n$  and more extreme quantiles. It is important to note that the rank score test is not immune to the effects of heterogeneity and maintaining correct Type I error rates with this test often requires weighted estimates and test statistics just as the drop in dispersion test does (Cade, 2003; Cade and others, 2005; Cade and others, 2006).

We will demonstrate the procedures with an example from Cade (1997), where lodgepole pine canopy cover (`lcc`) was modeled as a function of basal area (`apico`) and density of the trees (`picopha`). Use the dataset `fraserf`. Issue the following command for the simple regression of canopy cover (`lcc`) as a linear function of basal area (`apico`):

```
> Out <- lad(lcc~apico,data = fraserf,test = TRUE)
```

The model to be computed is written out algebraically where the dependent variable is `lcc` (lodgepole pine canopy cover) and the single independent variable is `apico` (basal area of lodgepole pine adjusted for slope of terrain). By default, `lad` will estimate an intercept. If `-1` is included in the formula specification, the intercept is left out, and the fit is forced through the origin. The `test = TRUE` option indicates that the model is to be compared to a reduced model that is a straight line parallel to the  $X$  axis going through the median  $y$  value (`lcc`). Thus, the reduced model has just one parameter, the constant. In this test Blossom uses a default sample size of 5,000 permutations (including the observed value) to approximate the permutation distribution.

Here are the results of the above `lad` command:

```
> summary(Out)
Least Absolute Deviation Regression (LAD)

Call:
lad(formula = lcc ~ apico, data = fraserf, test = TRUE)

Specification of Analysis:
  Number of Observations: 31
  Response Variable      : lcc

Independent variables      Regression coefficients
(Intercept)                8.789
apico                      1.054

Number of iterations: 3

Sum of absolute values of the residuals: 252.9
Solution: Successful

Regression Evaluation:
LAD Model: lcc ~(Intercept)+apico
```

```

Test Summary:
  Number of permutations      : 5000
  P-value of Full Model      : 2e-04***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Because canopy cover must be zero when basal area is zero, Cade (1997) used LAD regression models without an intercept term. Here the following command estimates the model above without an intercept:

```
> Out <- lad(lcc~-1+apico,data = fraserf)
```

The output is given below:

```
> summary(Out)
Least Absolute Deviation Regression (LAD)

Call:
lad(formula = lcc ~ -1 + apico, data = fraserf)

```

```

Specification of Analysis:
  Number of Observations: 31
  Response Variable      : lcc

  Independent variables      Regression coefficients
  apico                      1.314

  Number of iterations: 1

  Sum of absolute values of the residuals: 267.4
  Solution: Successful

```

A multiple independent variable LAD regression is specified by adding the appropriate independent variable names to the `lad` command. Here we consider the model used by Cade (1997) with lodgepole pine density (`picopha`) as an additional explanatory variable:

```
> Out <- lad(lcc~-1+apico+picopha,data = fraserf)
```

The added variables are assumed to be in the dataset in use. The commands `predict(Out)` and `residuals(Out)` can be used to obtain the predicted  $y$  values and the residuals respectively. Here are the results:

```
> summary(Out)
Least Absolute Deviation Regression (LAD)

Call:
lad(formula = lcc ~ -1 + apico + picopha, data = fraserf)

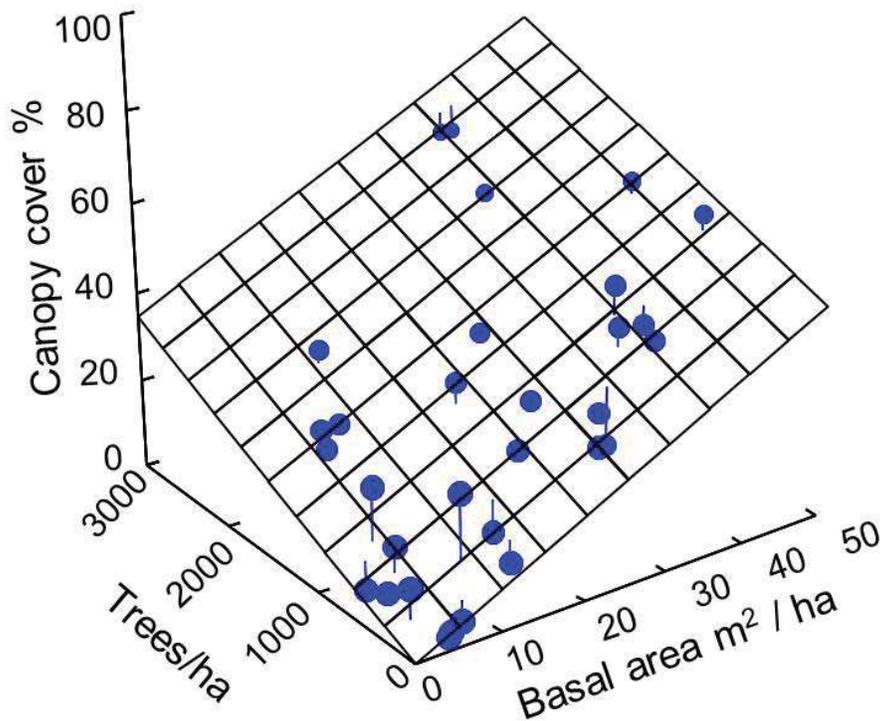
Specification of Analysis:
  Number of Observations: 31
  Response Variable      : lcc

  Independent variables      Regression coefficients
  apico                      0.93470
  picopha                    0.01157

  Number of iterations: 3

  Sum of absolute values of the residuals: 127.5
  Solution: Successful

```



**Figure 9.** Lodgepole pine canopy cover as a linear function of basal area and tree density estimated with LAD regression for 31 sample stands.

The regression function, observed values and residuals are plotted in figure 9.

A polynomial regression on a single independent variable, its square, its cube, and so on can be performed by including in the data a column containing the square, cube, and so on of the independent variable as well as the original independent and dependent variable. Use the dataset *fraserf* and enter the following *lad* command:

```
> Out <- lad(scc~-1+apien+pienpha+I(apien^2),data = fraserf)
```

to estimate the model used in Cade (1997), where canopy cover of Engelmann spruce (*scc*) is predicted as a function of basal area (*apien*), basal area<sup>2</sup> (*I(apien^2)*), and stem density (*pienpha*). The results are below and the regression surface is plotted in figure 10:

```
> summary(Out)
```

Least Absolute Deviation Regression (LAD)

Call:

```
lad(formula = scc ~ -1 + apien + pienpha + I(apien^2), data =
fraserf)
```

Specification of Analysis:

```
Number of Observations: 31
Response Variable      : scc
```

Independent variables	Regression coefficients
apien	1.582000
pienpha	0.008423
I(apien^2)	-0.030190

```
Number of iterations: 5
```

```
Sum of absolute values of the residuals: 85.84
```

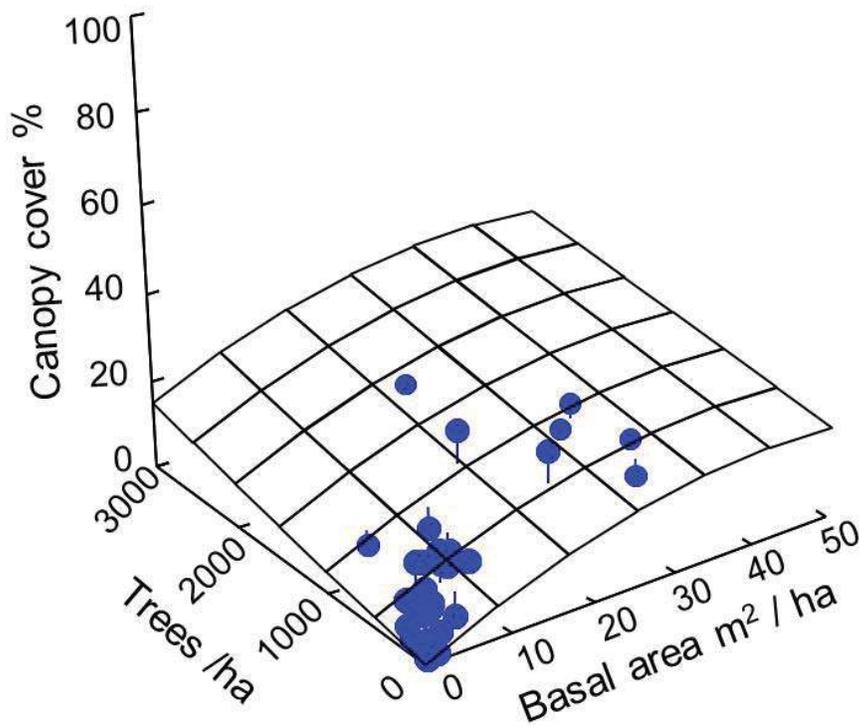
```
Solution: Successful
```

Here the quadratic curvature implied by use of basal area<sup>2</sup> can be tested with the `hypothesis.test` command to test whether the addition of the squared term yielded an improvement in fit. This is equivalent to testing the full model specified above against a reduced model that doesn't include the term  $(I(\text{apien}^2))$  for basal area<sup>2</sup>. This is done by fitting the reduced parameter null model using the `lad` command and then using the `hypothesis.test` with the full and reduced parameter LAD objects:

```
> Out2 <- lad(scc~-1+apien+pienpha,data = fraserf)
```

```
> set.seed(9)
```

```
> Test.Out <- hypothesis.test(Out,Out2,number.perms = 10000,  
+ double.permutation = TRUE,save.test = TRUE)
```



**Figure 10.** Engelmann spruce canopy cover as a quadratic function of basal area and linear function of tree density estimated with LAD regression for 31 sample stands.

Here are the results for the `hypothesis.test` command where we optionally have selected the double permutation scheme because our null hypothesized model is constrained through the origin:

```
> summary(Test.Out)
```

```
Least Absolute Deviation Regression (LAD)
```

```
Hypothesis Test, drop p-q-1 zero residuals with double
```

```

permutation

Call:
hypothesis.test(object1 = Out, object2 = Out2, number.perms = 10000,
  save.test = TRUE, double.permutation = TRUE)

Specification of Analysis:
Number of Observations: 31
Response Variable      : scc

Independent variables      Regression coefficients
      apien                0.61350
      pienpha              0.01353

Number of iterations: 3

Sum of absolute values of the residuals: 99.38
Solution: Successful

Regression Evaluation:
LAD Model:
scc ~apien+pienpha+I(apien^2)
Versus Hypothesis Model:
scc ~apien+pienpha

Test Summary:
Number of permutations      : 10000
Observed Test Statistic    : 0.1577
P-value of variables in full model but
not reduced model          : 0.0144*
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The results indicate that the coefficient for the quadratic basal area term differs from zero with  $P = 0.014$ . Here both double permutation and dropping of all but one of the zero residuals under the null model were implemented because the null model includes two parameters but no intercept. If we had not used the double permutation option (`double.permutation = FALSE`) and not deleted one of the zero residuals associated with the two parameters fit under the null model, then the  $P$ -value would be slightly smaller (0.0091) as in Cade (1997). The double permutation and dropping of zero residuals usually will increase the size of  $P$ -values slightly.

A goodness-of-fit measure for regression models is often a useful summary statistic. It is possible to compute an LAD coefficient of determination for the full model with reference to some reduced model (usually that specifies just an intercept term) by estimating the full model and obtaining the sums of absolute deviation (call it  $SAF$ ), then estimating the reduced parameter model and obtaining its sum of absolute deviations (call it  $SAR$ ), and computing the coefficient of determination  $R^1 = 1 - (SAF/SAR)$  (Cade and Richards, 1996; Cade, 1997). This can be extended to any selected regression quantile by replacing the sums of absolute deviations in the formula above with the sum of weighted absolute deviations minimized by regression quantiles (Koenker and Machado, 1999). We've already obtained the sums for the full parameter model, `scc = apien + pienpha + I(apien^2)` as  $SAF = 85.839$ , so to obtain them for the reduced parameter model:

```

> Out <- lad(scc~1,data = fraserf)
> summary(Out)

```

### Least Absolute Deviation Regression (LAD)

```
Call:
lad(formula = scc ~ 1, data = fraserf)
```

#### Specification of Analysis:

```
Number of Observations: 31
Response Variable      : scc
```

```
Independent variables      Regression coefficients
(Intercept)                10
```

```
Number of iterations: 1
```

```
Sum of absolute values of the residuals: 230
Solution: Successful
```

yields a sum,  $SAR = 230.0$  and, thus, the coefficient of determination  $R^1 = 1 - (85.839 / 230.000) = 0.627$ . This is interpreted as the model with variables `apien`, `pienpha`, and `I(apien^2)` yield estimates of conditional medians of `lcc` with a 63% reduction in sum of absolute deviations compared to the model that is just a simple estimate of the median of `lcc`.

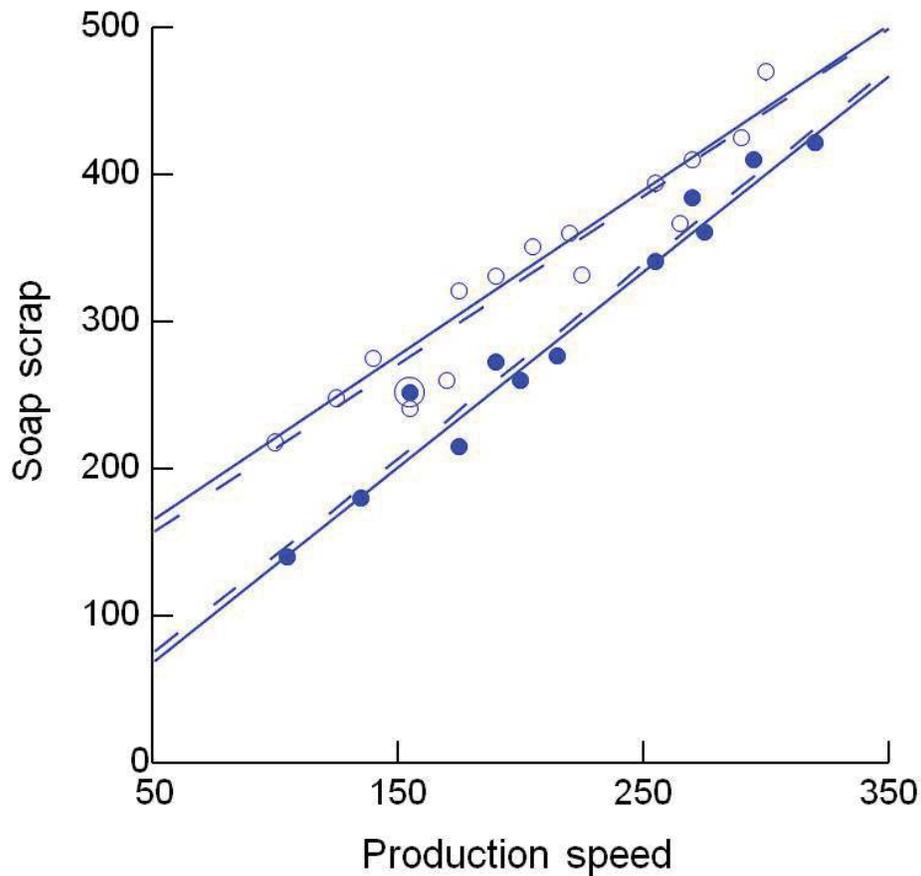
It is possible to specify greater or fewer permutations for calculating probabilities by specifying number of permutations as an option after either the `test` option for `lad` command or as an option in the `hypothesis.test` command. For example:

```
> Out <- lad(lcc~-1+apico+picopha,data = fraserf,number.perms = 10000,
+ test = TRUE)
```

will test all slope parameters equal to zero using 10,000 permutations of `y`. Manly (1991) summarizes recommendations on number of permutations to use in Monte Carlo sampling procedures. More is better but comes at increased computational cost. Specifying the random number seed is done with the standard `set.seed(num)` command in R before any function call.

It is important to recognize that the LAD regression model (and generalization to regression quantiles discussed below) can be extended to any linear model design that might be estimated with OLS regression, including various variable transformations, and mixtures of continuous independent variables with indicator variables for categorical predictors. Extensive examples are in Mielke and Berry (2001). Indeed it is possible to use LAD regression for linear model analyses of multifactorial experimental designs, where the focus is on estimating changes in conditional medians rather than estimating changes in conditional means as typically done with OLS regression (Cade and Richards, 1996; Mielke and Berry, 2001).

As an example, consider the soap production example from Cade and Richards (1996), where soap scrap (`y` is `soap`) is modeled as a linear function of production line speed ( $X_1$  is `speed`) and an indicator variable  $X_2$  is `line.n = 1` for production line 1 and  $X_2$  is `line.n = 0` for production line 2 (fig. 11). We are interested in testing whether the rates of change in soap scrap (`y`) as a function of line speed ( $X_1$ ) differs by production line ( $X_2$ ), which requires that we estimate a model with an interaction term (`lxs` is  $X_1 X_2$ ). Here, unlike in the original Fortran, we have the ability to utilize categorical terms and specify interactions as well as other functions of the predictors and response using the standard `formula` specification in R. We will convert line number (`line.n`) to a factor in order to demonstrate this capability.



**Figure 11.** Soap scrap as a linear function of production speed for line 1 (open circles) and line 2 (solid circles). Circled solid circle is an outlying value. Solid lines are LAD estimates and dashed lines are OLS estimates.

Use the dataset `neter365` and estimate the full parameter model with the interaction term specified

```
> neter365$line.n <- as.factor(c("Line1", "Line0")[(neter365$line.n == 0)+1])
> Out <- lad(soap~speed+line.n+lx, data = neter365)
```

Because `lx` is the interaction between `line.n` and `speed` we could specify an identical model using

```
> lad(soap~speed*line.n, data = neter365)
Least Absolute Deviation Regression (LAD)
```

```
Call:
lad(formula = soap ~ speed * line.n, data = neter365)
```

```
Specification of Analysis:
Number of Observations: 27
Response Variable       : soap
```

```
Independent variables      Regression coefficients
```

```

              (Intercept)          -3.197e-14
                speed            1.333e+00
            line.nLine1          1.076e+02
    speed:line.nLine1          -2.103e-01

```

Number of iterations: 5

Sum of absolute values of the residuals: 389.4  
 Solution: Successful

where the `lxs` is a column variable created by multiplying `speed` times `line.n` across all observations. Here are the results:

```
> Out <- lad(soap~speed+line.n+lxs,data = neter365)
```

Interpretation of the parameter estimates is identical to the interpretation for linear models estimated by OLS regression: the constant term is the intercept and the `speed` term ( $X_1$ ), is the slope for the regression of soap scrap on line speed for line 2, the `line.n` term ( $X_2$ ) is the difference between intercepts for the regressions for line 1 and line 2, and the `lxs` interaction term ( $X_1 X_2$ ) is the difference between slopes for the regressions for lines 1 and 2. We want to test the null hypothesis that the estimated interaction term is equal to zero, that is, differences in slopes equals zero, by specifying the reduced parameter null model in the `hypothesis.test` command:

```
> Out2 <- lad(soap~line.n+speed,data = neter365)
```

```
> Test.Out <- hypothesis.test(Out2,Out,number.perms = 10000,save.test = TRUE)
```

The results below indicated that there was moderate evidence ( $P = 0.046$ ) that the estimated difference in slopes of  $-0.21$  for the interaction term `lxs` was not equal to zero. Note that without dropping 2 of the 3 zero residuals in the null hypothesized model the  $P$ -value would be slightly smaller at  $P = 0.031$ .

```
> summary(Test.Out)
```

```

Least Absolute Deviation Regression (LAD)
Hypothesis Test, drop p-q-1 zero residuals

```

Call:

```
hypothesis.test(object1 = Out2, object2 = Out, number.perms = 10000,
  save.test = TRUE)
```

Specification of Analysis:

```

Number of Observations: 27
Response Variable      : soap

```

Independent variables	Regression coefficients
(Intercept)	39.250
line.nLine1	1.183
speed	60.420

Number of iterations: 6

Sum of absolute values of the residuals: 451.8  
 Solution: Non-Unique

Regression Evaluation:

```

LAD Model:
soap ~ (Intercept)+speed+line.nLine1+lxs
Versus Hypothesis Model:
soap ~ (Intercept)+line.nLine1+speed

```

```

Test Summary:
  Number of permutations           : 10000
  Observed Test Statistic         : 0.16
  P-value of variables in full model but
  not reduced model               : 0.0413*
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Confidence intervals on parameters in LAD regression models can be constructed by inverting the hypothesis testing process in an iterative fashion. This is accomplished by recognizing that testing for nonzero values of parameters in null hypotheses only requires a linear transformation of the dependent variable,  $y$ . For example, for the  $H_0: \beta_1 = \lambda$ , where  $\lambda$  is some hypothesized value of the parameter, you transform  $y$  to, say  $z$ , by  $z = y - \lambda X_1$ . The transformed values of the dependent variable,  $z$ , are then substituted for  $y$  in the regression model and estimation and hypothesis testing of the null  $H_0: \beta_1 = 0$  proceed as before. Cade and Richards (1996) describe in more general matrix notation how you accomplish this linear transformation for multiple parameters. Note that the formula defaults to what is done automatically when we test null hypotheses that parameters,  $\lambda$ , equal zero. The complication that arises in implementing this procedure for a  $(1 - \alpha)\%$  confidence interval is that you must iterate through many possible values of  $\lambda$  to define the bounds on the set of values of  $\lambda$  with  $P \geq \alpha$  for  $H_0: \beta_1 = \lambda$ . This can require many transformations of  $y$ , estimation with LAD, and testing the null hypothesis with the `hypothesis.test` command.

As an example of constructing confidence intervals, return to the model of lodgepole pine canopy cover as a function of pine basal area and stem density (Cade, 1997). Endpoints of the 95% confidence interval for the basal area parameter ( $b_1 = 0.935$ ) were given as 0.81–1.05 in Cade (1997). This means that the transformations `lcc-0.81(apico)`, call it `z81`, and `lcc-1.05(apico)`, call it `z105`, should have approximate  $P = 0.05$  when `z81` and `z105` are substituted for `lcc` in the regression model that includes `apico` (basal area) and `picopha` (tree density) as predictors for the partial model hypothesis of `apico`. Any transformation of `lcc` by values between 0.81 and 1.05 ought to yield  $P > 0.05$  and any outside of this interval ought to yield  $P \leq 0.05$ . Minor discrepancies can occur, of course, because of the resampling variation inherent in Monte Carlo procedures and because of discreteness in the permutation distribution. We try and make the resampling error as small as possible by using a large number of permutations (`number.perms=10,000`). The dataset `fraserf` includes the transformations `z81` and `z105`, as well as `z90 = lcc - 0.90(apico)` and `z50 = lcc - 0.50(apico)`. Here we know that the interval presented in Cade (1997) is slightly narrower than expected when the more recently developed double permutation scheme (Cade, 2005; Cade and Richards, 2006) is used because the null model is constrained through the origin. To run the hypothesis test corresponding to the null model that the parameter for `apico` equals 0.81, issue the commands:

```

> Out <- lad(z81~-1+apico+picopha,data = fraserf,number.perms = 10000)
> Out2 <- lad(z81~-1+picopha,data = fraserf,number.perms = 10000)
> set.seed(111)
> Test.Out <- hypothesis.test(Out2,Out,number.perms = 10000,
+   double.permutation = TRUE)

```

The output indicates  $P = 0.0592$ , well within the Monte Carlo resampling variation of 0.05 as it should be and only slightly larger than  $P = 0.0523$  obtained without the double permutation scheme.

```

> summary(Out)

```

Least Absolute Deviation Regression (LAD)

```
Call:
lad(formula = z81 ~ -1 + apico + picopha, data = fraserf,
     number.perms = 10000)
```

Specification of Analysis:

```
Number of Observations: 31
Response Variable      : z81
```

Independent variables	Regression coefficients
apico	0.12470
picopha	0.01157

```
Number of iterations: 4
```

```
Sum of absolute values of the residuals: 127.5
Solution: Successful
```

```
> summary(Test.Out)
```

```
Least Absolute Deviation Regression (LAD)
Hypothesis Test with double permutation
```

```
Call:
hypothesis.test(object1 = Out2, object2 = Out, number.perms = 10000,
                 double.permutation = TRUE)
```

Specification of Analysis:

```
Number of Observations: 31
Response Variable      : z81
```

Independent variables	Regression coefficients
picopha	0.01282

```
Number of iterations: 1
```

```
Sum of absolute values of the residuals: 140
Solution: Successful
```

Regression Evaluation:

```
LAD Model:
z81 ~apico+picopha
Versus Hypothesis Model:
z81 ~picopha
```

Test Summary:

```
Number of permutations      : 10000
Observed Test Statistic    : 0.09781
P-value of variables in full model but
not reduced model          : 0.0592.
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Similarly, we can run the hypothesis test corresponding to the null that the parameter for `apico` equals 0.50 by issuing the commands:

```
> Out <- lad(z50~-1+apico+picopha,data = fraserf,number.perms = 10000)
> Out2 <- lad(z50~-1+picopha,data =fraserf,number.perms =10000)
> set.seed(111)
```

```
> Test.Out <- hypothesis.test(Out2,Out,number.perms = 10000,  
+   double.permutation = TRUE)
```

The output here indicates the null hypothesis that the parameter equals 0.50 has  $P = 0.0001$ , much smaller than 0.05 so that this hypothesized parameter value must be outside the 95% confidence interval.

```
> summary(Out)
```

Least Absolute Deviation Regression (LAD)

Call:

```
lad(formula = z50 ~ -1 + apico + picopha, data = fraserf,  
    number.perms = 10000)
```

Specification of Analysis:

```
Number of Observations: 31  
Response Variable      : z50
```

Independent variables	Regression coefficients
apico	0.43470
picopha	0.01157

```
Number of iterations: 4
```

```
Sum of absolute values of the residuals: 127.5  
Solution: Successful
```

```
> summary(Test.Out)
```

Least Absolute Deviation Regression (LAD)  
Hypothesis Test with double permutation

Call:

```
hypothesis.test(object1 = Out2, object2 = Out, number.perms = 10000,  
    double.permutation = TRUE)
```

Specification of Analysis:

```
Number of Observations: 31  
Response Variable      : z50
```

Independent variables	Regression coefficients
picopha	0.01562

```
Number of iterations: 1
```

```
Sum of absolute values of the residuals: 211.2  
Solution: Successful
```

Regression Evaluation:

```
LAD Model:  
z50 ~apico+picopha  
Versus Hypothesis Model:  
z50 ~picopha
```

Test Summary:

```
Number of permutations      : 10000  
Observed Test Statistic    : 0.6561  
P-value of variables in full model but  
not reduced model          : 1e-04***
```

```
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Presently, hypothesized values of the parameter and their transformations must be made iteratively by successive approximation, that is, guess at values, compute the  $P$ -values, and then based on the size of the  $P$ -value successively move toward larger or lower values until you have values with  $P = \alpha$ , which define the confidence interval endpoints. This can require 20 or more iterations depending on how close your initial choice of hypothesized parameter values are to the final values. It is possible to use asymptotic procedures described in Birkes and Dodge (1993) to help pick initial values for confidence interval endpoints that might be close to those obtained by the iterative permutation testing process.

## Regression Quantiles

The `quant = num` option of the LAD regression command fits any specified conditional quantile as a linear regression model. LAD regression is the 0.50 (50th percentile) regression quantile. Various regression quantiles, e.g. 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95 (that is, 5<sup>th</sup>, 10<sup>th</sup>, 25<sup>th</sup>, 75<sup>th</sup>, 90<sup>th</sup> and 95<sup>th</sup> percentiles), can be estimated to examine linear trends in a dependent variable ( $y$ ) as a function of one or more independent variables ( $X$ ). Selecting `all.quant = TRUE` will yield all possible quantile regression estimates. If there is little variation in the errors across the independent variables (homogeneous errors), the regression quantiles will have similar slopes but different intercepts. However, if the errors are heterogeneous across the independent variables, then slopes and intercepts can differ greatly (Cade and Richards, 1996; Terrell and others, 1996; Cade and others, 1999; Koenker and Machado, 1999). Regression quantiles, thus, provide a way of modeling rates of change associated with heterogeneous variation in linear models without having to specify a functional link between conditional measures of means and variances. Regression quantiles are especially useful when the consequences of over and under prediction differ in a linear model. Cade and Noon (2003) present a primer on quantile regression for ecologists.

In studies of ecological limiting factors it is often expected that important measured processes operate as constraints on the response distribution ( $y$ ) and, thus, we may focus on estimating regression quantiles associated with the upper percentiles (for example, 90–99<sup>th</sup>) of the dependent variable, that is, rates of change estimated are along the upper boundary of the distribution as it changes across the independent variables (Terrell and others, 1996; Cade and others, 1999; Haire and others, 2000; Cade and Guo, 2000). Rates of change in the responses below the boundary constraint may be lower because of the impact of unmeasured processes (Cade and others, 1999). Many ecological processes can be considered constraints on responses, where rates of change estimated with regression quantiles for upper percentiles might yield new insights. Examples include animal responses to habitat, self-thinning in plants, algal productivity as a function of limiting nutrients, animal abundance and body size relations in macroecology, comparisons of local and regional species diversity, plant productivity as a function of species diversity, and competition field experiments. Estimating rates of change for endpoints of some interval of quantiles (for example, 10<sup>th</sup> and 90<sup>th</sup> percentiles) also provides a flexible way to estimate prediction intervals for responses without resorting to untenable distributional assumptions.

Returning to the soap production example, after using the dataset `neter365`, issue the following command:

```
> Out <- lad(soap~speed+line.n+lbs,quant = .5,data = neter365)
```

The output indicates that the coefficients estimated are identical to those above without the `quant = 0.5` option, because the 0.5 quantile is LAD regression. Notice also that both the

sum of absolute deviations minimized in LAD regression and the sum of weighted absolute deviations minimized in regression quantiles are reported. The weights used when minimizing sums of absolute deviations in regression quantiles are  $\tau$  for positive residuals and  $1-\tau$  for zero and negative residuals, where  $0 \leq \tau \leq 1$  is the selected quantile with `quant = num`. Thus, in this example the sum of weighted absolute deviations is exactly half the sum of absolute deviations.

```
> summary(Out)
      Quantile Regression

Call:
lad(formula = soap ~ speed + line.n + lxs, data = neter365, quant =
    0.5)

Specification of Analysis:
  Number of Observations: 27
  Response Variable      : soap
  For Quantile           : 0.5

      Independent variables      Regression coefficients
      (Intercept)                -3.197e-14
      speed                      1.333e+00
      line.nLine1                1.076e+02
      lxs                        -2.103e-01

Number of iterations: 5

Sum of absolute values of the residuals: 389.4
Weighted sum of the absolute deviations: 194.7
Solution: Successful
```

It is possible to test a full versus a reduced parameter regression quantile model with either the `lad` command with the option `test` set equal to `TRUE` or the `hypothesis.test`, where the test statistic is identical in computation as for LAD except that the simple sum of absolute deviations are replaced with the sum of weighted absolute deviations (Cade, 2005; Cade and Richards, 2006). Validity of hypothesis tests for regression quantiles using this test statistic requires the same assumption of independent, identical error distributions as for LAD regression. However, we expect most applications of regression quantiles to be made when it is unreasonable to assume homogeneous variation across the independent variables, that is, the identical error distribution assumption is violated. Therefore, we have included the regression quantile rank score test (Koenker, 1994; Koenker and Machado, 1999), its asymptotic  $P$ -value approximation with a Chi-square distribution, and a permutation approximation that makes use of the permutation test for OLS regression. Type I errors of the regression quantile rank score test are less sensitive to heterogeneous error distributions because the statistic is based on the sign of the residuals from the reduced parameter null model and not their size. However, as Cade (2003) and Cade and others (2006) make abundantly clear, valid Type I error rates often will require appropriate weighted estimates and test statistics. This quantile rank score test is implemented with the option `rank.score = TRUE` given with the `hypothesis.test` command.

As an example, consider the acorn production data as related to oak (*Quercus* spp.) forest characteristics (Schroeder and Vangilder, 1997) as analyzed with regression quantiles by Cade and others (1999). We will estimate 0.10 and 0.90 (10<sup>th</sup> and 90<sup>th</sup> percentiles) regression quantiles of annual acorn biomass (kg/ha, `wtperha`) as a function of a forest suitability index (`oakccsi`)

based on canopy cover and number of oak species (Schroeder and Vangilder, 1997) using the dataset `acorn` and will issue the command for a 0.10 regression quantile:

```
> Out <- lad(wtperha~oakccsi,quant = .1,data = acorn)
```

The command then is issued to test the hypothesis that the slope for the 0.10 quantile equals zero with the rank score test:

```
> Out2 <- lad(wtperha~1,quant = .1,data = acorn)
> Test.Out <- hypothesis.test(Out2,Out,number.perms = 10000,rank.score = TRUE)
```

The output indicates that the estimated slope for the 0.10 regression quantile (21.8) likely differs from zero ( $P = 0.012$ ).

```
> summary(Out)
```

#### Quantile Regression

Call:

```
lad(formula = wtperha ~ oakccsi, data = acorn, quant = 0.1)
```

Specification of Analysis:

```
Number of Observations: 43
Response Variable       : wtperha
For Quantile           : 0.1
```

Independent variables	Regression coefficients
(Intercept)	2.44
oakccsi	21.77

Number of iterations: 2

```
Sum of absolute values of the residuals: 1526
Weighted sum of the absolute deviations: 173.7
Solution: Successful
```

```
> summary(Test.Out)
```

#### Quantile Regression Hypothesis Test of Rank Score

Call:

```
hypothesis.test(object1 = Out2, object2 = Out, number.perms = 10000,
rank.score = TRUE)
```

Specification of Analysis:

```
Number of Observations: 43
Response Variable       : wtperha
For Quantile           : 0.1
```

Independent variables	Regression coefficients
(Intercept)	12.82

Number of iterations: 1

```
Sum of absolute values of the residuals: 1737
Weighted sum of the absolute deviations: 194.2
Solution: Successful
```

Regression Evaluation:

```
0.1 Quantile Regression Model:
```

```

wtperha ~ (Intercept) + oakccsi
Versus Hypothesis Model at Quantile 0.1:
wtperha ~ (Intercept)

Test Summary:
Number of permutations                : 10000

Observed Rank Score Test Statistic    : 0.1842
P-value of Rank Score Test            : 0.01**

Asymptotic Rank Score Statistic      : 6.326
(Distributed as Chi-square with degrees of
freedom equal to difference in number of
parameters between full and reduced models)
P-Value of Asymptotic RS Stat        : 0.0119*

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Similarly, we can estimate the 0.90 regression quantile for the same functional relation by issuing the command:

```
> Out <- lad(wtperha~oakccsi, quant = 0.9, data = acorn)
```

followed by the command:

```
> Out2 <- lad(wtperha~1, quant = 0.9, data = acorn)
> Test.Out <- hypothesis.test(Out2, Out, number.perms = 10000, rank.score = TRUE)
```

The output for the 0.90 regression quantile indicates that the rate of change of acorn biomass with the suitability index is five times greater (102.3) at the 90<sup>th</sup> percentile of the distribution compared to the 10<sup>th</sup> percentile of the distribution (fig. 12). Clearly, there is heterogeneous variation in the acorn biomass changes across the acorn suitability index, with only larger biomass occurring at higher values of the suitability index. The estimated slope of the 0.90 regression quantile also likely differs from zero ( $P = 0.035$ ). Here because of the heterogeneity, improved Type I error rates could be obtained by using weighted estimates with the rank score tests.

```
> summary(Out)
```

#### Quantile Regression

```
Call:
lad(formula = wtperha ~ oakccsi, data = acorn, quant = 0.9)
```

#### Specification of Analysis:

```
Number of Observations: 43
Response Variable      : wtperha
For Quantile          : 0.9
```

Independent variables	Regression coefficients
(Intercept)	14.45
oakccsi	102.30

```
Number of iterations: 3
```

```
Sum of absolute values of the residuals: 1722
Weighted sum of the absolute deviations: 268.5
Solution: Successful
```

```
> summary(Test.Out)
```

Quantile Regression  
Hypothesis Test of Rank Score

Call:  
hypothesis.test(object1 = Out2, object2 = Out, number.perms = 10000,  
rank.score = TRUE)

Specification of Analysis:

Number of Observations: 43  
Response Variable : wtperha  
For Quantile : 0.9

Independent variables	Regression coefficients
(Intercept)	89.92

Number of iterations: 1

Sum of absolute values of the residuals: 1951  
Weighted sum of the absolute deviations: 324.1  
Solution: Successful

Regression Evaluation:

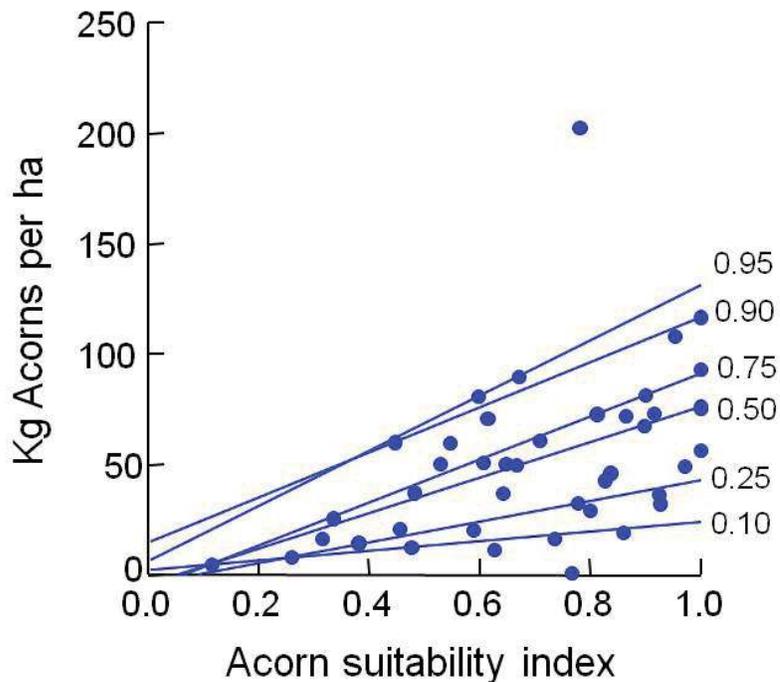
0.9 Quantile Regression Model:  
wtperha ~ (Intercept) + oakccsi  
Versus Hypothesis Model at Quantile 0.9:  
wtperha ~ (Intercept)

Test Summary:

Number of permutations	:	10000
Observed Rank Score Test Statistic	:	0.1151
P-value of Rank Score Test	:	0.0354*
Asymptotic Rank Score Statistic	:	4.198
(Distributed as Chi-square with degrees of freedom equal to difference in number of parameters between full and reduced models)		
P-Value of Asymptotic RS Stat	:	0.04048*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



**Figure 12.** Average annual biomass of acorns and acorn suitability indices based on oak forest characteristics in 43 0.2-ha sample plots in Missouri. Solid lines are estimates for 6 selected regression quantiles.

Estimates for other regression quantiles can be obtained by changing the value used in the option `quant = num`. Note that the  $P$ -values approximated by the permutation evaluation of the rank score tests are similar to those produced by the asymptotic Chi-square distributional approximation (uses a Chi-square distribution with degrees of freedom equal to difference in number of parameters in full versus reduced models). Although the permutation  $P$ -values are slightly smaller than those for the asymptotic Chi-square approximation, the differences may be attributable just to the resampling error associated with the Monte Carlo approximation. Simulation research in Cade (2003) and Cade and others (2006) established that the permutation version of the rank score test maintains valid Type I error rates at more extreme quantiles ( $\tau$ ) with smaller  $n$  than does the Chi-square distributional approximation.

Confidence intervals based on the regression quantile rank score statistic can be formed by a process identical to that described above for LAD regression. However, if you want to use the asymptotic Chi-square approximation of  $P$ -values for computing confidence intervals, there are fast implementations in linear programming algorithms available for S-Plus, R, and SAS (Koenker, 1994; Cade and others, 1999; Koenker and Machado, 1999).

A multiple regression quantile example is provided by Cade and others (1999), where glacier lily (*Erythronium grandiflorum*) seedlings (`seedlings`) are linearly related to the number of flowers (`flowers`) and an index of rockiness (`rockiness`) in  $n = 256$  contiguous  $2 \times 2$  m quadrats (fig. 13).

To estimate the 95<sup>th</sup> regression quantile model issue the following commands and obtain the following output:

```
> Out <- lad(seedlings~flowers+rockiness,quant = 0.95,data = lily)
> summary(Out)
```

### Quantile Regression

```
Call:
lad(formula = seedlings ~ flowers + rockiness, data = lily, quant =
0.95)
```

#### Specification of Analysis:

```
Number of Observations: 256
Response Variable      : seedlings
For Quantile          : 0.95
```

Independent variables	Regression coefficients
(Intercept)	20.30000
flowers	0.08504
rockiness	-0.08987

```
Number of iterations: 5
```

```
Sum of absolute values of the residuals: 3800
Weighted sum of the absolute deviations: 272.8
Solution: Successful
```

The estimates indicate a 0.085 increase in seedling numbers with each increase in flower numbers at a given level of rockiness, and a decrease of 0.090 of seedling numbers with each increase in unit of the rockiness index. We can test that these parameters jointly are equal to zero by comparing the full parameter model above with the reduced parameter model having just an intercept by the command:

```
> Out2 <- lad(seedlings~1,quant = .95,data = lily)
> Test.Out <- hypothesis.test(Out2,Out,number.perms = 10000,rank.score=TRUE)
> summary(Test.Out)
```

#### Quantile Regression Hypothesis Test of Rank Score

```
Call:
hypothesis.test(object1 = Out2, object2 = Out, number.perms = 10000,
rank.score = TRUE)
```

#### Specification of Analysis:

```
Number of Observations: 256
Response Variable      : seedlings
For Quantile          : 0.95
```

Independent variables	Regression coefficients
(Intercept)	16

```
Number of iterations: 1
```

```
Sum of absolute values of the residuals: 3377
Weighted sum of the absolute deviations: 301.2
Solution: Successful
```

#### Regression Evaluation:

```
0.95 Quantile Regression Model:
seedlings ~(Intercept)+flowers+rockiness
Versus Hypothesis Model at Quantile 0.95:
```

```

seedlings ~ (Intercept)

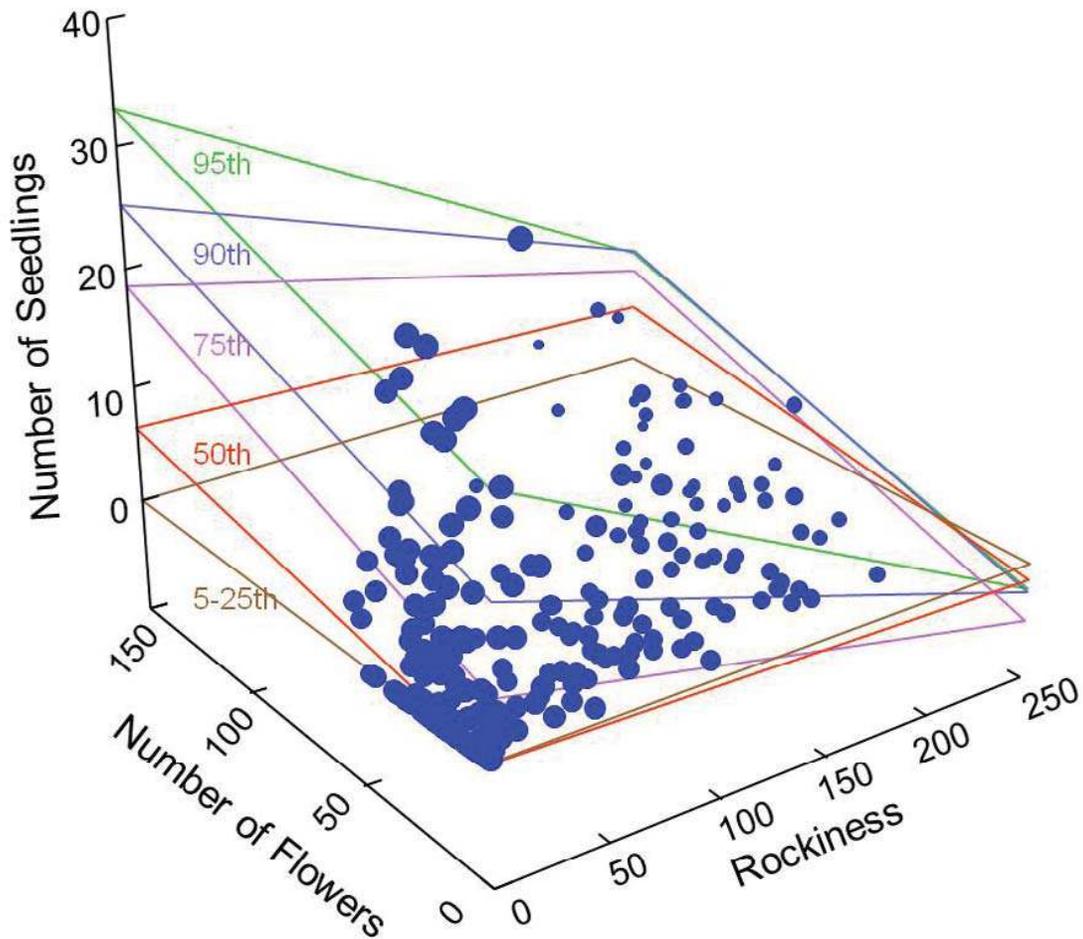
Test Summary:
Number of permutations           : 10000

Observed Rank Score Test Statistic : 0.02857
P-value of Rank Score Test         : 0.029*

Asymptotic Rank Score Statistic   : 7.017
(Distributed as Chi-square with degrees of
freedom equal to difference in number of
parameters between full and reduced models)
P-Value of Asymptotic RS Stat     : 0.02995*

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```



**Figure 13.** Glacier lily seedling counts, lily flower numbers, and rockiness index for 256 2 x 2 m quadrats in subalpine meadow of western Colorado. Surfaces are for selected regression quantile estimates (1 outlying count of 72 seedlings is not plotted).

The output indicates some evidence that at least one of the parameters is unlikely to equal zero ( $P = 0.030$  for asymptotic approximation and  $P = 0.029$  for permutation approximation). We can test each of the parameters individually by issuing the series of commands:

```

> Out3 <- lad(seedlings~flowers,quant = 0.95,data = lily)
> Out4 <- lad(seedlings~rockiness,quant = 0.95,data = lily)
> test.3 <- hypothesis.test(Out3,Out,number.perms = 10000,rank.score = TRUE)
> test.4 <- hypothesis.test(Out4,Out,number.perms = 10000,rank.score = TRUE)

```

The output indicates stronger evidence that the parameter for rockiness does not equal zero ( $P = 0.036$ ) than for the parameter for flowers ( $P = 0.073$ ).

```
> summary(test.3)
```

```

      Quantile Regression
Hypothesis Test of Rank Score

```

```
Call:
```

```
hypothesis.test(object1 = Out3, object2 = Out, number.perms = 10000,
rank.score = TRUE)
```

```
Specification of Analysis:
```

```

Number of Observations: 256
Response Variable      : seedlings
For Quantile           : 0.95

```

Independent variables	Regression coefficients
(Intercept)	18.58000
flowers	-0.07273

```
Number of iterations: 4
```

```

Sum of absolute values of the residuals: 3460
Weighted sum of the absolute deviations: 296.1
Solution: Successful

```

```
Regression Evaluation:
```

```

0.95 Quantile Regression Model:
seedlings ~ (Intercept)+flowers+rockiness
Versus Hypothesis Model at Quantile 0.95:
seedlings ~ (Intercept)+flowers

```

```
Test Summary:
```

```

Number of permutations           : 10000

Observed Rank Score Test Statistic : 0.01715
P-value of Rank Score Test        : 0.0372*

Asymptotic Rank Score Statistic   : 4.17
(Distributed as Chi-square with degrees of
freedom equal to difference in number of
parameters between full and reduced models)
P-Value of Asymptotic RS Stat     : 0.04115*

```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> summary(test.4)
```

```

      Quantile Regression
Hypothesis Test of Rank Score

```

```
Call:
```

```
hypothesis.test(object1 = Out4, object2 = Out, number.perms = 10000,
rank.score = TRUE)
```

```

Specification of Analysis:
  Number of Observations: 256
  Response Variable      : seedlings
  For Quantile          : 0.95

      Independent variables      Regression coefficients
      (Intercept)                22.00000
      rockiness                  -0.06522

Number of iterations: 4

Sum of absolute values of the residuals: 3918
Weighted sum of the absolute deviations: 278.6
Solution: Successful

Regression Evaluation:
  0.95 Quantile Regression Model:
  seedlings ~ (Intercept)+flowers+rockiness
  Versus Hypothesis Model at Quantile 0.95:
  seedlings ~ (Intercept)+rockiness

Test Summary:
  Number of permutations                : 10000

  Observed Rank Score Test Statistic   : 0.01254
  P-value of Rank Score Test           : 0.0779

  Asymptotic Rank Score Statistic      : 3.094
  (Distributed as Chi-square with degrees of
  freedom equal to difference in number of
  parameters between full and reduced models)
  P-Value of Asymptotic RS Stat       : 0.07856

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Both  $P$ -values are consistent with the 90% confidence intervals given in Cade and others (1999) that did not overlap zero for either variable. Note that the permutation  $P$ -values are slightly smaller than the Chi-square distribution approximation. The confidence intervals in Cade and others (1999) were based on inverting the asymptotic Chi-square distribution approximation of the rank score statistic as part of the linear programming solution for regression quantiles that are available in the "quantreg" package for R. Because of the heterogeneity evident in this model, confidence intervals and rank score testing would be better based on weighted estimates (Cade and others, 2006)

The use of all quantile regression estimates and weighting is provided for an example relating the logarithm of Lahontan cutthroat trout (*Oncorhynchus clarki henshawi*) numbers per meter of stream ( $\ln lctm$ ) to stream width:depth ratio ( $widrat$ ) for  $n = 71$  observations of streams across years in Nevada (Dunham and others, 2002; Cade, 2005; Cade and others, 2006). The scatter plot in figure 14 (A) indicate moderate heterogeneity and some nonlinearity in the relationship. Dunham and others (2002) chose to use a nonlinear model  $y = \exp(\beta_0 + \beta_1 X_1 + \varepsilon)$  estimated in the linear scale by taking natural logarithms of both sides of the equation. Cade (2005), Cade and others (2006), and Cade and Richards (2006) also used weighted estimates, where the coefficients of the weight function  $w = (1.310 - 0.0017 X_1)^{-1}$  were estimated from the average pairwise differences (by using expected value obtained from multiresponse sequence

procedure) between all possible quantile regression estimates for  $\beta_0$  and for  $\beta_1$  obtained by using the `all.quantis = TRUE` option:

```
> Out <- lad(lnlctm~widrat,data = lahontan,all.quantis = TRUE)
> summary(Out)
All Quantile Regression
```

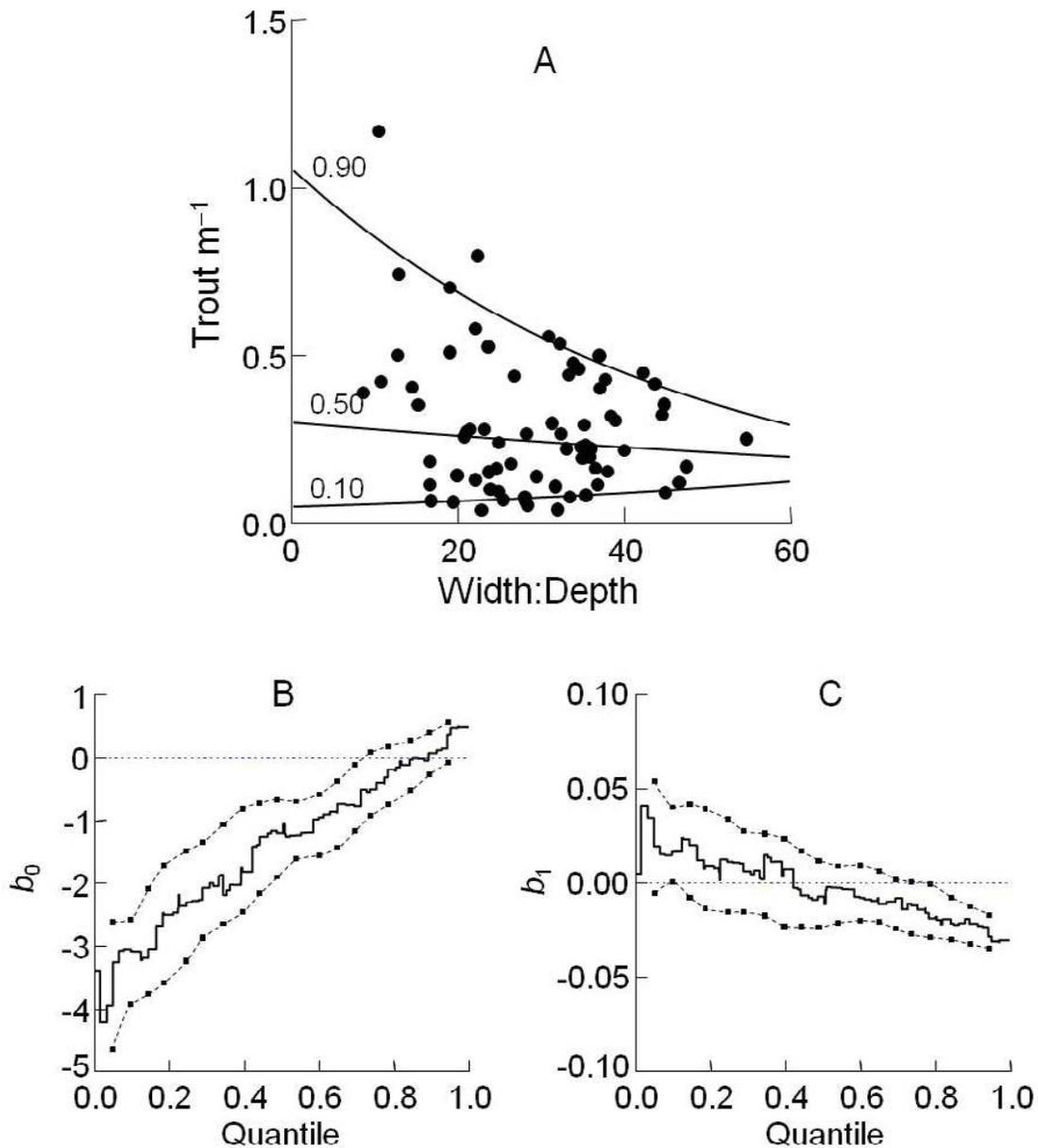
```
Call:
lad
  lnlctm ~ widrat
 lahontan
 TRUE
```

```
Specification of Analysis:
  Number of Observations: 71
  Dependent Variable: lnlctm
  Number of Parameters: 2
  Number of Solutions: 77
  Solution Result Was: Successful
```

Output can be obtained using `QuantValues()`

Columns are `Quantile`, `Objective Function Solution`, `Predicted Y at X-Bar`, followed by estimates of coefficients for independent variables. Each row describes an estimated conditional quantile function.

The output object contains a row for each unique interval of quantiles, with column variables specifying the upper endpoint of the quantile interval (`Quantile`), the objective function minimized (`ObjFuncSol` is weighted sum of absolute deviations), the predicted value for that quantile at the mean of the independent variables (`PredY_Xbar`), and the parameter estimates (here, `b_(Intercept)` and `b_widrat`). Plots of the parameter estimates by quantile suggested the linear location-scale (in log scale) form of heterogeneity was a reasonable approximation so that a single weight function could reasonably be applied to all quantiles. The empirical distribution plots for each parameter estimate by quantile in figure 14 (B and C) were made from the weighted estimates by connecting the point estimates with an appropriate step function (fig. 14 B and C). The weighted estimates were made by multiplying all variables (`lnlctm`, a column of 1's for the intercept, and `widrat`) by the weights (`wt`) to form the variables `wtlnlctm`, `wt`, and `wtwidrat`. The model was estimated as:



**Figure 14.** Lahontan cutthroat trout m<sup>-1</sup> and width:depth ratios for small streams sampled 1993 to 1999 ( $n = 71$ ); exponentiated estimates for 0.90, 0.50, and 0.10 regression quantiles for the weighted model  $w(\ln y) = w(\beta_0 + \beta_1 X_1 + (\gamma_0 + \gamma_1 X_1)\varepsilon)$ ,  $w = (1.310 - 1.017 X_1)^{-1}$ . Solid lines in (B) and (C) are step functions for estimates of  $\beta_0$  and  $\beta_1$  by  $\tau \in [0, 1]$  and dashed lines connect pointwise 90% confidence intervals for  $\tau \in \{0.05, 0.10, 0.15, \dots, 0.95\}$  based on inverting the double permutation test.

```
> Out <- lad(wtlnlctm~-1+wt+wtwidrat,data = lahontan,all.quant = TRUE)
> summary(Out)
```

```
All Quantile Regression
```

```
Call:
```

```
lad
```

```
wtlnlctm ~ -1 + wt + wtwidrat
```

```
lahontan
TRUE
```

```
Specification of Analysis:
  Number of Observations: 71
    Dependent Variable: wtlnlctm
  Number of Parameters: 2
  Number of Solutions: 79
  Solution Result Was: Successful
```

```
Output can be obtained using QuantValues()
```

```
Columns are Quantile, Objective Function Solution, Predicted Y at X-Bar,
  followed by estimates of coefficients for independent variables.
Each row describes an estimated conditional quantile function.
```

Note that the variable  $wt$  ( $= 1 \times wt$ ) replaces the usual constant term because the weighted model requires that weights are multiplied by all independent variables including the column of 1's for the constant. The confidence intervals formed around the parameter estimates by quantiles in figure 14 (B and C) were made by using the drop in dispersion permutation test with double permutation (because null models for weighted estimates were constrained through the origin). Cade and Richards (2006) formed 90% confidence intervals at quantiles = 0.05, 0.10, 0.15 , ..., 0.90, 0.95 by successive iteration of hypothesized values as explained for LAD regression starting on page 79. These intervals were only slightly narrower than intervals formed by inverting the permutation version or Chi-square distributional approximation of the rank score test (Cade and others, 2006). Here, we provide an example of the hypothesis tests for the weighted 0.90 quantile regression estimates:

```
> Out <- lad(wtlnlctm~-1+wt+wtwidrat,data = lahontan,quant=.9)
> summary(Out)
```

```
Quantile Regression
```

```
Call:
lad(formula = wtlnlctm ~ -1 + wt + wtwidrat, data = lahontan, quant =
  0.9)
```

```
Specification of Analysis:
  Number of Observations: 71
  Response Variable      : wtlnlctm
  For Quantile           : 0.9
```

Independent variables	Regression coefficients
wt	0.05762
wtwidrat	-0.02154

```
Number of iterations: 2
```

```
Sum of absolute values of the residuals: 82.41
Weighted sum of the absolute deviations: 8.797
Solution: Successful
```

```
> Out1 <- lad(wtlnlctm~-1+wt,data = lahontan,quant = .9)
```

```
> Out2 <- lad(wtlnlctm~-1+wtwidrat,data = lahontan,quant = .9)
```

```

> Test.Out <- hypothesis.test(Out,Out1,double.permutation = TRUE)

> summary(Test.Out)
      Quantile Regression
      Hypothesis Test with double permutation

Call:
hypothesis.test(object1 = Out, object2 = Out1, double.permutation =
TRUE)

Specification of Analysis:
  Number of Observations: 71
  Response Variable      : wtlnlctm
  For Quantile          : 0.9

      Independent variables      Regression coefficients
              wt                    -0.6764

Number of iterations: 1

Sum of absolute values of the residuals: 80.3
Weighted sum of the absolute deviations: 9.831
Solution: Successful

Regression Evaluation:
  0.9 Quantile Regression Model:
  wtlnlctm ~wt+wtwidrat
  Versus Hypothesis Model at Quantile 0.9:
  wtlnlctm ~wt

Test Summary:
  Number of permutations          : 5000
  Observed Test Statistic        : 0.1176
  P-value of variables in full model but
  not reduced model              : 0.0018**
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

> Test.Out <- hypothesis.test(Out,Out2,double.permutation = TRUE)
> summary(Test.Out)
      Quantile Regression
      Hypothesis Test with double permutation

Call:
hypothesis.test(object1 = Out, object2 = Out2, double.permutation =
TRUE)

Specification of Analysis:
  Number of Observations: 71
  Response Variable      : wtlnlctm
  For Quantile          : 0.9

      Independent variables      Regression coefficients
              wtwidrat          -0.02022

```

```

Number of iterations: 1

Sum of absolute values of the residuals: 81.26
Weighted sum of the absolute deviations: 8.808
Solution: Successful

Regression Evaluation:
0.9 Quantile Regression Model:
wtlnlctm ~wt+wtwidrat
Versus Hypothesis Model at Quantile 0.9:
wtlnlctm ~wtwidrat

Test Summary:
Number of permutations                : 5000
Observed Test Statistic               : 0.001285
P-value of variables in full model but
not reduced model                    : 0.7912
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Note that both null hypothesized models above do not include a constant for a column of 1's because of the weighting scheme, so that the double permutation option `double.permutation` was used to provide better Type I error rates. The output indicates a strong, nonzero slope but an intercept that doesn't differ from zero (in the log scale) for the 0.90 regression quantile. Notice that these results are consistent with the 90% CI which indicate nonzero slopes for quantiles  $\geq 0.80$  and nonzero intercepts for quantiles  $\leq 0.70$ .

There are several alternative approaches for estimating weights discussed in Cade and others (2005, 2006), Koenker (2005), and Cade and Richards (2006).

## Ordinary Least Squares Regression (OLS)

Estimation and permutation testing alternatives for the familiar ordinary least squares regression are available with the `OLS` command that feeds the argument `OLS = TRUE` option to the `lad` command. OLS regression estimates rates of change in conditional means. The permutation testing approaches are identical to those used for LAD regression, and are described in Kennedy and Cade (1996) and Anderson and Legendre (1999). The test statistic is similar in structure to that for LAD regression, except for  $T_{obs}$  equals (sum of squared residuals for reduced parameter model–sum of squared residuals for full parameter model) / sum of squared residuals for full model. Large values of  $T_{obs}$  are evidence against the null hypothesis that the parameter(s) equal(s) zero. Our test statistic is equivalent to an  $F$  statistic without the degrees of freedom ( $df$ ), which are not necessary because they are invariant under permutation.  $T_{obs} \times (df \text{ full model} / (df \text{ reduced} - df \text{ full model})) = F$  statistic with numerator  $df$  equal to  $df \text{ reduced} - df \text{ full model}$  and denominator  $df$  equal to  $df \text{ full model}$ . For testing all slope parameters equal zero, the dependent variable is permuted against the matrix of independent variables, and for testing partial models (subhypotheses) involving some subset of parameters, residuals from the reduced parameter, null model are permuted to the matrix of independent variables. The benefits and validity of these permutation schemes are described in Kennedy and Cade (1996) and Anderson and Legendre (1999).

We will demonstrate the OLS permutation procedure by returning to the soap scrap example previously analyzed with LAD regression (fig. 11). Issue the following commands:

```
> Out <- ols(soap~speed+line.n+lx, data = neter365)
```

We see in the output below that the estimate for the parameters differs slightly from those estimated with LAD regression. In particular, notice that the estimate for the interaction term of line speed and production line number is -0.176 for OLS compared to -0.210 for LAD regression.

```
> summary(Out)
      Ordinary Least Squares Regression

Call:
ols(formula = soap ~ speed + line.n + lxs, data = netter365)

Specification of Analysis:
  Number of Observations: 27
  Response Variable      : soap

      Independent variables      Regression coefficients
      (Intercept)                7.5740
      speed                      1.3220
      line.nLine1                90.3900
      lxs                       -0.1767

Sum of squares of the residuals: 9904
```

We can test that the interaction coefficient for `lxs` equals zero by using the `hypothesis.test` command after the OLS command similar to what was done for LAD regression:

```
> set.seed(39)
> Out2 <- ols(soap~speed+line.n,data = netter365)
> Test.Out <- hypothesis.test(Out,Out2,number.perms = 10000,save.test = TRUE)
> summary(Test.Out)
      Ordinary Least Squares Regression
      Hypothesis Test

Call:
hypothesis.test(object1 = Out, object2 = Out2, number.perms = 10000,
  save.test = TRUE)

Specification of Analysis:
  Number of Observations: 27
  Response Variable      : soap

      Independent variables      Regression coefficients
      (Intercept)                27.280
      speed                      1.231
      line.nLine1                53.130

Sum of squares of the residuals: 10714

Regression Evaluation:
  Ordinary Least Squares:
  soap ~ (Intercept)+speed+line.nLine1+lxs
  Versus Hypothesis Model:
  soap ~ (Intercept)+speed+line.nLine1
```

```

Test Summary:
  Number of permutations           : 10000
  Observed Test Statistic         : 0.08175
  P-value of variables in full model but
  not reduced model               : 0.1831
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The output indicates that there is little evidence to believe that the interaction term ( $1_{xs}$ ) differs from zero with  $P = 0.1831$ .

The normal theory test of the same estimate yields  $F_{1,23} = 1.88$  with  $P = 0.184$ , indicating the similarity of the permutation and normal theory probabilities for this example. Remember, that the LAD regression estimate (-0.21) and permutation test suggested that there was some evidence that the interaction term differed from zero with  $P = 0.046$ . As explained in Cade and Richards (1996), the one outlying value for line 2 (circled value in fig. 11) has a studentized residual of 3.18, and though not a large outlier, has enough impact on the OLS regression estimate to reduce the magnitude of the  $1_{xs}$  interaction term and increase the standard error of the estimate. If this outlier is deleted, the OLS estimate for the interaction term  $1_{xs}$  becomes -0.23 with  $F_{1,23} = 3.70$  and  $P = 0.068$ , much more similar to the LAD estimate and test results. Minor outliers such as this one (fig. 11) are likely to be missed or ignored in many analyses made with OLS regression and, thus, not detect nonzero effects with as much power as possible. LAD regression estimates and their permutation test are far less sensitive to the impacts of one or a few outliers (Cade and Richards, 1996; Mielke and Berry, 2001).

## G-sample and 1-sample Goodness-of-fit Coverage Tests

The  $g$ -sample and 1-sample goodness-of-fit variants of the empirical coverages tests (Mielke and Yao, 1988, 1990; Mielke and Berry, 2001) are alternatives to the Kolmogorov-Smirnov family of tests for comparing cumulative distribution functions of continuous variates. If  $x_{1/i} < \dots < x_{n_i/i}$  are the  $n_i$  order statistics of the  $i^{\text{th}}$  sample ( $i = 1, \dots, g$ ),  $N = \text{sum of the } n_i \text{ from } i = 1 \text{ to } g$ , and  $F_N(x) = (\text{number of observed values among the } N \text{ pooled values which are } \leq x) / (N + 1)$ , the  $n_i + 1$  coverages associated with the  $n_i$  observed values of the  $i^{\text{th}}$  sample ( $i = 1, \dots, g$ ) are denoted by  $C_{j/i} = F_N(x_{j/i}) - F_N(x_{j-1/i})$ . Consider an example with two groups of four and three observations with order statistics  $x_{1/1} < x_{2/1} < x_{3/1} < x_{1/2} < x_{4/1} < x_{2/2} < x_{3/2}$ . The  $N + 1 = 8$  empirical coverages are  $C_{1/1} = C_{2/1} = C_{3/1} = C_{3/2} = 1/8$ ,  $C_{4/1} = C_{2/2} = 2/8$ ,  $C_{5/1} = 3/8$ , and  $C_{1/2} = 4/8$ , for example,  $C_{4/1} = 5/8$  ((the number of observations  $\leq x_{4/1}$ )/8) - 3/8 ((the number of observations  $\leq x_{3/1}$ )/8) = 2/8. The coverage test statistic is a function of the absolute value of the difference between the observed coverages ( $C_{j/i}$ ) and their expected value  $(n_i + 1)^{-1}$  raised to some exponent  $v$ . The 1-sample goodness-of-fit coverage test implemented here is based on raising the absolute value of the coverages to an exponent of 1, and is equivalent to the test described by Sherman (1950). A special variant of this goodness-of-fit test for circular distributions is equivalent to Rao's (1976) spacing test for a uniform circular distribution. Probabilities under the null hypothesis are provided by a Pearson type III approximation based on the exact mean, variance, and skewness for the 1-sample goodness-of-fit test and based on Monte Carlo resampling approximations for  $g$ -sample tests.

We will examine a two group example of the coverage test by returning to the blue grouse migration example given for `mrpp`. Here we will examine just the migration distances

(dist) since the coverage test is limited to univariate comparisons. Quantile plots of the distances are given in figure 15.

The commands to compare these data are:

```
> set.seed(24)
> Out <- coverage(variable = bgrouse$distance, group = bgrouse$sex,
+   number.perms = 100000)
```

and the output is:

```
> summary(Out)
      Univariate G-Sample Empirical Coverage Test

Call:
coverage(variable = bgrouse$distance, group = bgrouse$sex,
  number.perms = 1e+05)

Specification of Analysis:
  Number of Observations: 21
  Number of Groups       : 2
  Distance Exponent      : 1
  Number of Permutations: 100000

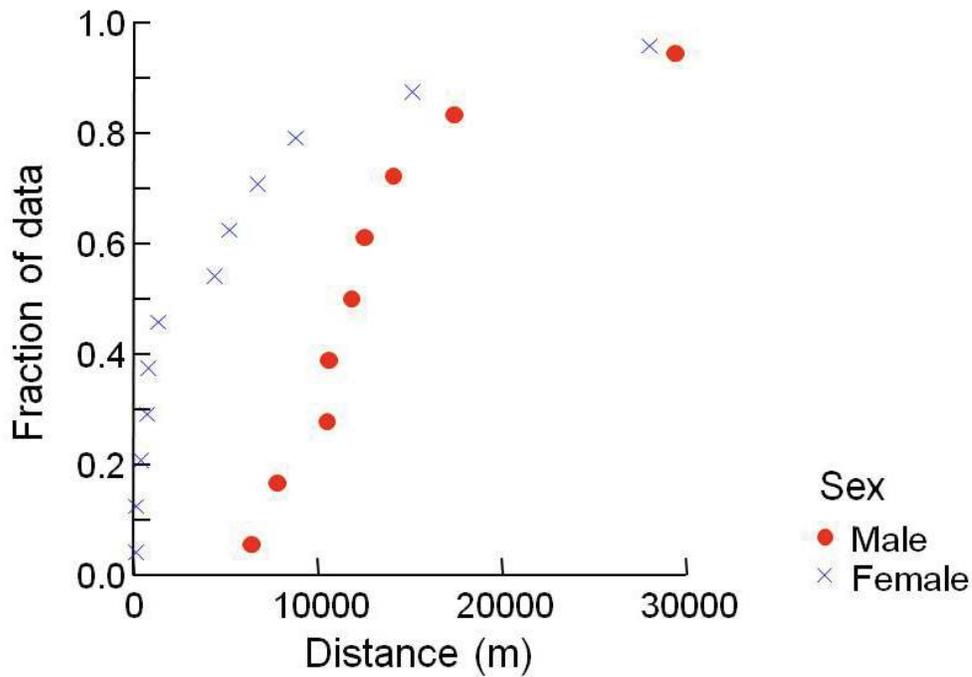
Group Summary:
  Group Value  Group Size
           3           9
           4          12

Results:
  Observed coverage statistic           : 1.122
  Mean of coverage statistic           : 0.9818
  Estimated variance of coverage statistic: 0.03034
  Standard deviation of the variance
  of the coverage statistic           : 0.0001314

  Observed standardized coverage statistic: 0.803
  Skewness of observed coverage statistic : 0.05628
  Probability (Pearson Type III)
  of a larger or equal coverage statistic : 0.21
  Probability (Resampled)
  of a larger or equal coverage statistic : 0.1963
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

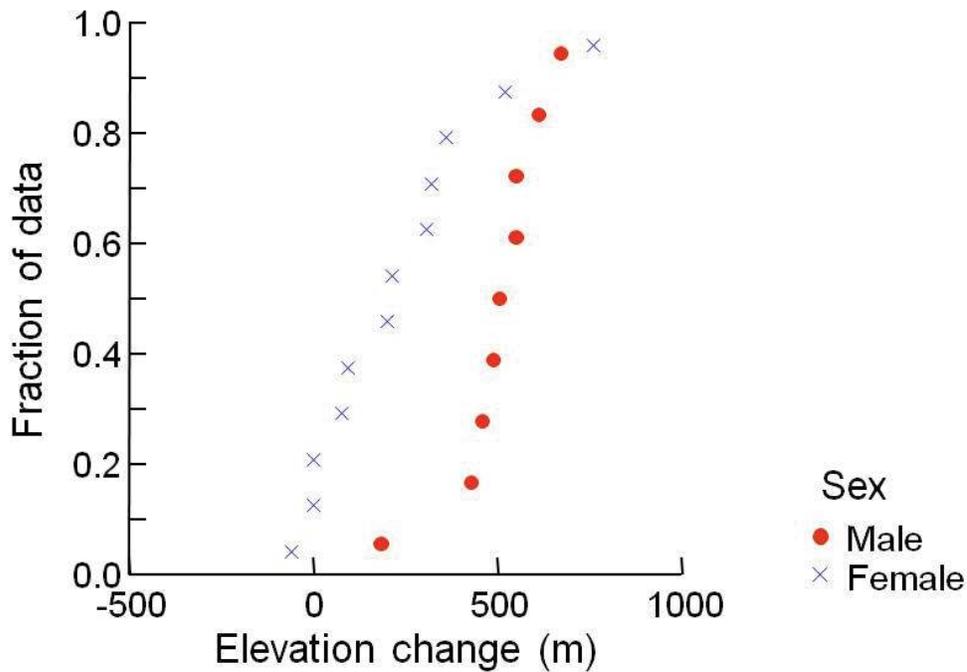
The option `number.perms = 100000` requested that we use 100,000 resamples for the Monte Carlo approximation of the  $P$ -values. Two probabilities are reported from this approximation, one which is the standard Monte Carlo approach of referencing the observed test statistic to those generated by the resampling, and a second which uses the resampled statistics to estimate the variance and skewness of the sampling distribution to be evaluated with the Pearson type III curve. Note their similarity here. We obtained  $P = 0.196$  (Monte Carlo resampling approximation), which suggests that there is little evidence to conclude that the cumulative distributions of male and female blue grouse migration distances differ. It also is possible to get an exact enumeration of all possible permutations of the data for the coverage test statistic for small sample sizes ( $n < 24$ ) by using the `exact = TRUE` option, which gives a  $P = 0.204$  for

the data in figure 15. Interestingly, comparisons made with `mrpp` with `expon = 1` yield  $P = 0.008$  (exact enumeration) and with `mrpp` and `expon = 2`, `c.form = 2` (permutation version of  $t$ -test) yield  $P = 0.040$  (exact enumeration), both which suggest greater evidence that male and female migration distances differ.



**Figure 15.** Quantile plots of migration distances for 9 male and 12 female blue grouse (data from Cade and Hoffman, 1993).

Consider now the univariate comparisons of elevation changes (`elev`) made by male and female blue grouse in migration (fig. 16).



**Figure 16.** Quantile plots of elevation changes (m) made by 9 male and 12 female blue grouse when migrating from breeding to winter areas (data from Cade and Hoffman, 1993).

We can compare these distributions with the coverage statistic by issuing the commands:

```
> Out <- coverage(variable = bgrouse$elev, group = bgrouse$sex, exact = TRUE)
```

Here the  $P = 0.019$  yields similar evidence of differences in movements as does `mrpp` with `expon = 2`, `c.form = 2` ( $P = 0.010$ , exact), whereas `mrpp` with `expon = 1` ( $P = 0.004$ , exact) yields slightly stronger evidence of differences.

```
> summary(Out)
```

```
Exact Univariate G-Sample Empirical Coverage Test
```

```
Call:
```

```
coverage(variable = bgrouse$elev, group = bgrouse$sex, exact = TRUE)
```

```
Specification of Analysis:
```

```
Number of Observations: 21
Number of Groups       : 2
Distance Exponent      : 1
```

```
Group Summary:
```

```
Group Value  Group Size
          3           9
          4          12
```

```
Results:
```

```
Observed coverage statistic           : 1.348
Probability (Exact)
of a larger or equal coverage statistic : 0.01884*
```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The power characteristics to detect nonzero differences between the groups clearly differ among these various test statistics. Mielke and Berry (2001) give other examples where conclusions differ greatly for coverages tests compared to MRPP comparisons. More extensive simulation research is needed to better characterize the types of distributional differences better detected by the coverage tests relative to MRPP. Simple location shifts (change in medians) appear to be detected with greater power by MRPP than by the coverage tests.

The 1-sample goodness-of-fit coverage tests are of the Kendall-Sherman type (Mielke and Berry, 2001). Given an observed set of univariate data with order statistics  $x_1 < x_2 < \dots < x_n$ , for  $i = 1$  to  $n$  these values must be transformed to the cumulative probability of the distribution function  $F(x)$  specified under the null hypothesis,  $U(i) = F(x_i)$  for  $i = 1$  to  $n$ . These  $U(i)$  are then the probability integral transformed values used in the coverage test. For example, consider the five values given by Bradley (1968, p. 301–302) that were hypothesized to come from a normal distribution with mean = 3 and standard deviation = 2; -0.311, -0.078, 0.555, 1.462, and 5.711. The data `brad302` has these five values ( $X$ ) and the five transformed cumulative probabilities from a normal distribution with a mean = 3 and standard deviation = 2 for these values ( $FX$ ). The 1-sample goodness-of-fit test for these data are implemented by issuing the commands:

```
> Out <- coverage(variable = fx, data = brad302)
```

The output below yield  $P = 0.031$ , which suggests that there is some evidence to support the belief that these five observations did not come from a population with a normal distribution having a mean = 3 and standard deviation = 2, similar to conclusions reached with the Kolmogorov-Smirnov goodness-of-fit test (Bradley, 1968, p. 301–302).

```
> summary(Out)
      Kendall-Sherman Goodness of Fit Test
```

```
Call:
coverage(variable = fx, data = brad302)
```

```
Specification of Analysis:
  Number of Observations      : 5
  Number of Intervals         : 5
```

```
Results:
  Observed Statistic T        : 1.05
  Expected Statistic T        : 0.6698
  Variance of Statistic T     : 0.03755
  Standardized statistic T    : 1.96
  Skewness of statistic T     : 0.2467
```

```
  P-value of observed statistic,
  P(Expected ≥ Observed)     : 0.03135*
```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Because there are so many possible distributions that might be hypothesized and used with the 1-sample goodness-of-fit coverage test, we have not implemented any specific cumulative distribution function transformations in Blossom. We expect the user to make such transformations on the data prior to conducting an analysis in Blossom. There is one special cumulative distribution function transformation offered in Blossom because it is not commonly

available in other statistical packages. That is we transform the data to a cumulative uniform random distribution on the unit circle to test the null hypothesis that the sample came from a population with a uniform random circular distribution. This is done with the `coverage` test option `interv = num`, where the number provided tells the test how many units describe the circular units of measure recorded (for example, `interv = 360` would be used for angular orientations recorded in degrees). When these transformed values are tested with the coverage test, it performs a goodness-of-fit test equivalent to Rao's (1976) spacing test for uniformity of circular distributions. We will consider the example given by Rao (1976) for the compass orientation at which 10 homing pigeons departed when released 25 km west of their loft: 20, 35, 350, 120, 85, 345, 80, 320, 280, and 85 degrees. Use the dataset `rao`.

```
> Out <- coverage(variable = angl, interv = 360, data = rao)
```

The output below has  $P = 0.328$  for an observed test statistic = 0.7611. This observed test statistic is related to Rao's  $U$  by  $(360/2) T_{obs} = U$ . Rao obtained  $U_{10} = 137$  degrees.

```
> summary(Out)
```

```

Kendall-Sherman Goodness of Fit Test

Call:
coverage(variable = angl, interv = 360, data = rao)

Specification of Analysis:
  Number of Observations      : 10
  Number of Intervals         : 9
  Arc distance used           : 360

Results:
  Observed Statistic T        : 0.7611
  Expected Statistic T        : 0.6974
  Variance of Statistic T     : 0.02304
  Standardized statistic T    : 0.42
  Skewness of statistic T     : 0.1888

  P-value of observed statistic,
  P(Expected ≥ Observed)     : 0.3278
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It is important to point out that the coverage tests assume continuous data with no tied values. The grouse examples above for the  $g$ -sample coverage tests had two tied values for both distances and elevations. This is only a minor violation of the assumption of continuity that likely has minimal impact on the analysis. At this point in time it is difficult to say what proportion of a sample comprised of tied observations constitutes a serious violation of the continuity assumption of the coverage tests. Beware of tied values.

## Appendix 1. Common Statistical Tests Embraced by the MRPP Command

Multiresponse permutation procedures (MRPP) can duplicate many common statistical tests (the parametric tests listed below are all permutation versions).

- Two-sample  $t$ -test
- One-way analysis of variance
- Multivariate analysis of variance
- Hotelling's  $T^2$
- Median test (2 and  $k$ -sample)
- Wilcoxon-Mann-Whitney test
- Kruskal-Wallis test
- Goodman and Kruskal contingency table tests of association  
( $\tau$ -a,  $\tau$ -b)
- Generalized runs tests (including Wald-Walfowitz runs test)
- Durbin-Watson for univariate first-order autoregression
- Schoener's  $t^2/r^2$  for bivariate first-order autoregression

The multiresponse permutation procedures for randomized block data (MRBP) and permutation tests for matched pairs (PTMP) can duplicate the following tests (the parametric tests listed below are all permutation versions).

- Matched pairs and 1-sample  $t$ -test
- Analysis of variance for complete randomized blocks
- Sign test
- Wilcoxon signed rank test
- Pearson correlation coefficient
- Spearman rank correlation
- Kendall tau (correlation)
- Friedman's test for randomized blocks
- Spearman's footrule and multi-block extension
- Cochran's  $Q$  and McNemar's tests
- Cohen's kappa

Other less familiar tests are also known to be special cases of MRPP and MRBP. Be aware, too, that many of the above tests are strictly univariate or bivariate, but MRPP and MRBP often generalize to the multivariate case as well. Further, most of the above listed tests use the square of Euclidean distance in the definition of the test statistic, whereas MRPP and MRBP have the option of choosing a distance measure commensurate with the data space. The generalized distance function in MRPP yields alternative, often more powerful versions of these tests.

## Appendix 2. Compilation of the internal Blossom code

The code that performs the statistical analysis for all functions in the Blossom package is based on the code used in the Blossom statistical software developed by Brian Cade and Jon Richards at the U.S. Geological Survey. The Fortran subroutines required to run the package are compiled into either a .so (shared object on Linux) or a .dll (dynamic-linked library on Windows). The Fortran code has no dependencies on external libraries or subroutines with the exception of some subroutines available in R which are used for printing from Fortran, generating random numbers (done indirectly by calling C functions which in turn use R), and checking if the user has tried to escape the current function call. Given the dependencies on the

subroutines available through R, the .dll/.so should be compiled using the compiler available through Rtools which will find these subroutines automatically for example using R CMD SHLIB InternalFcts.f90 myHelp.c -o Blossom.dll will build the .dll correctly. No flags are required and a Makevars file has not been written for this package. Building the .dll from source code on windows requires the appropriate version of Rtools and building the entire package from source will require other tools as described in "R Installation and Administration" which is available on the CRAN website (<http://cran.r-project.org/manuals.html>). Once the proper tools are installed and paths are correctly set the command `install.packages("Blossom.tar.gz", type = "source")` can be used to build the package from source. Building R packages from source code under Windows is nontrivial given the list of tools and intricate setup required and should not be undertaken without good reason.

## Appendix 3. Blossom Statistics Program Installation, Configuration, Requirements

### Blossom Requirements and Program Limits

Development and testing was done originally under Windows 98 and NT but for the last few years we have used only Windows 2000, XP, and 7.

Blossom allocates virtual memory space dynamically at runtime. This means that the amount of memory required by Blossom depends on the program analyses being run and the size of the data and associated internal storage required for the analysis. If this exceeds the physical Random Access Memory (RAM) available, Blossom uses Windows virtual memory management. This runtime memory required cannot exceed the paging file size (swap space) available to Windows. Documentation to Windows virtual memory management should be consulted.

Internally, Blossom has some limits on the amount of data it can support. The total amount of memory the program and dynamically allocated array space can occupy is about 2 gigabytes. Here are some limits within Blossom:

- Number of elements in a Blossom command: 1024 (command plus all variables and options and delimiters).
- Maximum single command element length in ASCII representation: 25 bytes
- Maximum size of Blossom command: 8192 bytes
- Number of variables: 1024
- Number of observations: about 2 billion (depends on number of variables, total memory limited to about 2GB)
- Number of quantiles in a Median and Quantile (MEDQ) analysis: 250,000,000
- Number of observations in a MEDQ: 250,000,000
- Number of groups in one MEDQ: 250,000,000
- Maximum group size in MEDQ: 250,000,000
- Maximum number of variables in a MEDQ analysis: 255
- Maximum number of blocks in an Exact Multiresponse Randomize Block Procedure: 9
- Dataset size: Limit of Windows Virtual Memory (depends on associated memory required for an analysis)
- Variable name size: 25 bytes
- Maximum significant digits of double precision numbers: about 15

- Missing value internal representation:  $0.10 \times 10^{-37}$  (no datum within a dataset should have this value)

Note that there may be other, smaller limits depending on the combination of all factors considered.

Some specific Blossom statistical analyses require minimum numbers of elements:

- Minimum number of observations for Multi-Response Permutation Procedure (MRPP): 6
- Minimum number of groups for MRPP: 2
- Minimum group size for MRPP: 2
- Minimum number of observations for Exact MRPP: 3
- Minimum number of groups for an Exact MRPP: 2
- Minimum group size for Exact MRPP: 2
- Minimum number of groups for Multiresponse Randomized Block Procedure (MRBP): 2
- Minimum number of groups for an MRBP: 2
- Minimum number of blocks for Exact MRBP: 2
- Minimum number of cases for Permutation Test for Matched Pairs (PTMP): 3
- Minimum number of observations for Multi-Response Sequence Procedure (MRSP): 6
- Minimum number of observations for Exact MRSP: 2

## Appendix 4. Blossom Development and Testing

The primary computing routines in the Blossom package were written in Fortran 95. Hundreds of hours of testing have gone into this (and previous) version of Blossom since development was initiated in 1989. We have done everything possible to ensure that our modifications of the principal computing routines for the MRPP family of statistics obtained from Paul. W. Mielke, Jr., yield numerical results identical to the original routines. We made comparisons with published and other known results, and had Dr. Mielke compare some of his original analyses with our program. Similar comparisons of numerical output were made for the regression quantile and rank score tests made with programs provided by Roger Koenker. No doubt, some errors remain undetected and we urge you to report any obvious or suspicious errors to us. We recommend running analyses on the datasets provided with the Blossom software to see if your computer duplicates the output in the User Manual and output files.

## Appendix 5. Acknowledgments

We thank B. J. Biggerstaff (Centers for Disease Control and Prevention) and G. M. Zimmerman (Lake Superior State University) for providing reviews of this user manual and the R software implementation. We would like to acknowledge the continual interest and support provided by Drs. Mielke and Koenker, and the testing work done on previous versions by Dr. David R. Smith. We thank Dr. William L. Slauson for initiating the first version of the Blossom User's Manual, and Jon D. Richards for his programming efforts over the many years.

## References Cited

- Aebischer, N.J., Robertson, P.A., and Kenward, R.E., 1993, Compositional analysis of habitat use from animal radio-tracking data: *Ecology*, v. 74, p. 1313–1323.
- Anderson, M.J., and Legendre, P., 1999, An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model: *Journal Statistical Computation and Simulation*, v. 62, p. 271–303.
- Barrodale, I., and Roberts, F.D.K., 1973, An improved algorithm for discrete L1 linear approximation: *SIAM Journal of Numerical Analysis*, v. 10, no. 5, p. 839–848.
- Barrodale, I., and Roberts, F.D.K., 1974, Algorithm 478: solution of an overdetermined system of equations in the L1 norm: *Comm. Assoc. Comp. Math.*, v. 17, p. 319–320.
- Berry, K.J., and Mielke, P.W., 1983, Computation of finite population parameters and approximate probability values for multi-response permutation procedures (MRPP): *Communications in Statistics—Simulation and Computation*, v. 12, p. 83–107.
- Berry, K.J., and Mielke, P.W., 1985, Goodman and Kruskal's tau-b statistic—A nonasymptotic test of significance: *Sociological Methods and Research*, v. 13, p. 543–550.
- Berry, K.J., and Mielke, P.W., 1988, A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters: *Educational and Psychological Measurement*, v. 48, p. 921–933.
- Berry, K.J., and Mielke, P.W., 1992, A family of multivariate measures of association for nominal independent variables: *Educational and Psychological Measurement*, v. 52, p. 41–55.
- Berry, K.J., Kvamme, K.L., and Mielke, P.W., 1983, Improvements in the permutation test for the spatial analysis of the distribution of artifacts into classes: *American Antiquity*, v. 48, p. 547–553.

- Biondini, M.E., Bonham, C.D., and Redente, E.F., 1985, Secondary successional patterns in a sagebrush (*Artemisia tridentata*) community as they relate to soil disturbance and soil biological activity: *Plant Ecology*, v. 60, p. 25–36.
- Biondini, M.E., Mielke, P.W., and Redente, E.F., 1988, Permutation techniques based on Euclidean analysis spaces—A new and powerful statistical method for ecological research: *Coenoses* v. 3, n. 3, p. 155–174.
- Birkes, D., and Dodge, Y., 1993, *Alternative methods of regression*: John Wiley and Sons, Inc., New York, 228 p.
- Box, G.E.P., Hunter, W.G., and Hunter, J.S., 1978, *Statistics for experimenters*: New York, John Wiley and Sons, Inc., 653 p.
- Bradley, J.V., 1968, *Distribution-free statistical tests*: Prentice-Hall, Inc., 388 p.
- Cade, B.S., 1997, Comparison of tree basal area and canopy cover in habitat models—Subalpine forest: *Journal of Wildlife Management*, v. 61, p. 326–335.
- Cade, B.S., 2003, *Quantile regression models of animal habitat relationships*: Fort Collins, Colo., Colorado State University, Ph.D. dissertation, 186 p.
- Cade, B.S., 2005, Linear models—Permutation methods, *in* Everitt, B., and Howell, D., eds., *Encyclopedia of statistics in the behavioral science*: John Wiley and Sons, v. 2, p. 1049–1054.
- Cade, B.S., and Guo, Q., 2000, Estimating effects of constraints on plant performance with regression quantiles: *Oikos*, v. 91, p. 245–254.
- Cade, B.S., and Hoffman, R.W., 1990, Winter use of Douglas-fir forests by blue grouse in Colorado: *Journal of Wildlife Management*, v. 54, p. 471–479.
- Cade, B.S., and Hoffman, R.W., 1993, Differential migration of blue grouse in Colorado: *Auk*, v. 110, p. 70–77.
- Cade, B.S., and Noon, B.R., 2003, A gentle introduction to quantile regression for ecologists: *Frontiers in Ecology and the Environment*, v. 1, p. 412–420.
- Cade, B.S., and Richards, J.D., 1996, Permutation tests for least absolute deviation regression: *Biometrics*, v. 52, p. 886–902.
- Cade, B.S., and Richards, J.D., 2005, *User manual for Blossom statistical software*: U.S. Geological Survey, Biological Resources Discipline, Open-File Report 2005–1353, 124p.
- Cade, B.S., and Richards, J.D., 2006, A permutation test for quantile regression: *Journal of Agricultural, Biological, and Environmental Statistics*, v. 11, p. 106–126.
- Cade, B.S., Noon, B.R., and Flather, C.H., 2005, Quantile regression reveals hidden bias and uncertainty in habitat models: *Ecology*, v. 86, p. 786–800.
- Cade, B.S., Richards, J.D., and Mielke, P.W. Jr., 2006, Rank score and permutation testing alternatives for regression quantile estimates: *Journal of Statistical Computation and Simulation*, v. 76, p. 331–355.
- Cade, B.S., Terrell, J.W., and Schroeder, R.L., 1999, Estimating effects of limiting factors with regression quantiles: *Ecology*, v. 80, p. 311–323.
- Dodd, C.K., and Cade, B.S., 1998, Movement patterns and the conservation of amphibians breeding in small, temporary wetlands: *Conservation Biology*, v. 12, p. 331–339.
- Dodge, Y., ed., 1987, *Statistical data analysis based on the L1-norm and related methods*: Elsevier Science Publishers, B. V. (North-Holland), 464 p.
- Dunham, J.B., Cade, B.S., and Terrell, J.W., 2002, Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles: *Transactions of the American Fisheries Society*, v. 131, p. 86–98.
- Dyadkin, I.G. and Hamilton, K.G., 1997, A study of 64-bit multipliers for Lehmer Pseudorandom Number Generators: *Computer Physics Communications*, v. 103, p. 103–130.

- Freedman, D., and Lane, D., 1983, A nonstochastic interpretation of reported significance levels: *Journal Business and Economic Statistics*, v. 1, p. 292–298.
- Edgington, E.S., 1987, *Randomization tests*: New York, Marcel Dekker, Inc., 341 p.
- Fawcett, R.F., 1990, The alignment method for displaying and analyzing treatments in blocking designs: *American Statistician*, v. 44, p. 204–209.
- Gentle, J.E., 1977, Least absolute values estimation: an introduction: *Communications in Statistics—Simulation and Computation*, v. B6, n. 4, p. 313–328. [See also the rest of this volume]
- Good, P., 2000, *Permutation tests—A practical guide to resampling methods for testing hypotheses* (2d ed.): Springer, 270p.
- Haire, S.L., Bock, C.E., Cade, B.S., and Bennett, B.C., 2000, The role of landscape and habitat characteristics in limiting abundance of grassland nesting songbirds in an urban open space: *Landscape and Urban Planning*, v. 48, n. 1–2, p. 65–82.
- Hodges, J.L. and Lehmann, E.L., 1962, Rank methods for combination of independent experiments in analysis of variance: *Annals of Mathematical Statistics*, v. 33, p. 482–497.
- Iyer, H.K., Berry, K.J., and Mielke, P.W., 1983, Computation of finite population parameters and approximate probability values for multi-response randomized block permutation procedures (MRBP): *Communications in Statistics—Simulation and Computation*, v. 12, n. 4, p. 479–499.
- Kennedy, P.E., and Cade, B.S., 1996, Randomization tests for multiple regression: *Communications in Statistics—Simulation and Computation*, v. 25, p. 923–936.
- Koenker, R., 1994, Confidence intervals for regression quantiles, *in* P. Mandl, P., and Hušková, M., ed., *Asymptotic statistics:Physica-Verlag, Proceedings of the 5th Prague Symposium*, p. 349–359.
- Koenker, R., 2005, *Quantile regression*: Cambridge University Press, New York, *Econometric Society Monographs*, 349 p.
- Koenker, R., and Bassett, G., 1978, Regression quantiles: *Econometrica*, v. 46, p. 33–50.
- Koenker, R., and Bassett, G., 1982, Robust tests for heteroscedasticity based on regression quantiles: *Econometrica*, v. 50, p. 43–61.
- Koenker, R., and Machado, J.A.F., 1999, Goodness of fit and related inference processes for quantile regression: *Journal of the American Statistical Association*, v. 94, p. 1296–1310.
- Koenker, R., and Portnoy, S., 1996, *Quantile regression*: University of Illinois at Urbana-Champaign, College of Commerce and Business Administration, Office of Research Working Paper, 97–0100, 77p.
- Manly, B.F.J., 1991, *Randomization and Monte Carlo Methods in Biology*: Chapman and Hall, New York, 281 p.
- Matsumoto, M., and Nishimura, T., 1998, Mersenne Twister—A 623-dimensionally equidistributed uniform pseudo-random number generator: *ACM Transactions on Modeling and Computer Simulation*, v. 8, p. 3–30.
- Mielke, P.W., 1984, Meteorological applications of permutation techniques based on distance functions, *in* Krishnaiah, P.R., and Sen, P.K., eds., *Handbook of statistics*: Elsevier Science Publishers, Amsterdam, North-Holland, v. 4, p. 813–830.
- Mielke, P.W., 1985, Geometric concerns pertaining to applications of statistical tests in the atmospheric sciences: *Journal of Atmospheric Sciences*, v. 42, p. 1209–1212.
- Mielke, P.W., 1986, Non-metric statistical analyses—Some metric alternatives: *Journal of Statistical Planning and Inference*, v. 13, p. 377–387.
- Mielke, P.W., Jr., 1991, The application of multivariate permutation methods based on distance functions in the earth sciences: *Earth-Science Reviews*, v. 31, p. 55–71.

- Mielke, P.W., and Berry, K.J., 1982, An extended class of permutation techniques for matched pairs: *Communications in Statistics—Theory and Methods*, v. A11, p. 1197–1207.
- Mielke, P.W., and Berry, K.J., 1983, Asymptotic clarifications, generalizations, and concerns regarding an extended class of matched pairs tests based on powers of ranks: *Psychometrika*, v. 48, p. 483–485.
- Mielke, P.W., and Berry, K.J., 1994, Permutation tests for common locations among samples with unequal variances: *Journal of Educational and Behavioral Statistics*, v. 19, p. 217–236.
- Mielke, P.W., Jr., and Berry, K.J., 1999, Multivariate tests for correlated data in completely randomized designs: *Journal of Educational and Behavioral Statistics*, v. 24, p. 109–131.
- Mielke, P.W., Jr., and Berry, K.J., 2001, *Permutation methods—A distance function approach*: Springer-Verlag, 352p.
- Mielke, P.W. and Iyer, H.K., 1982, Permutation techniques for analyzing multi-response data from randomized block experiments: *Communications in Statistics—Theory and Methods*, v. 11, p. 1427–1437.
- Mielke, P.W., and Yao, Y.C., 1988, A class of multiple sample tests based on empirical coverages: *Annals of the Institute of Statistical Mathematics*, v. 40, p. 165–178.
- Mielke, P.W. and Yao, Y.C., 1990, On g-sample empirical coverage tests: Exact and simulated null distributions of test statistics with small and moderate sample sizes: *Journal of Statistical Computation and Simulation*, v. 35, p. 31–39.
- Mielke, P.W., Berry, K.J., and Brier, G.W. 1981. Application of multi-response permutation procedures for examining seasonal changes in monthly mean sea-level pressure patterns: *Monthly Weather Review*, v. 109, p. 120–126.
- Mielke, P.W., Berry, K.J., and Johnson, E.S., 1976, Multi-response permutation procedures for a priori classifications: *Communications in Statistics—Theory and Methods*, v. A5, p. 1409–1424.
- Mielke, P.W., Berry, K.J., and Medina, J.G., 1982, Climax I and II—Distortion resistant residual analyses: *Journal Applied Meteorology*, v. 21, p. 788–792.
- Mielke, P.W., Berry, K.J., Brockwell, P.J., and Williams, J.S., 1981, A class of nonparametric tests based on multiresponse permutation procedures: *Biometrika*, v. 68, p. 720–724.
- Mielke, H.W., Anderson, J.C., Berry, K.J., Mielke, P.W., Chaney, R.L., and Leech, M., 1983, Lead concentrations in inner-city soils as a factor in the child lead problem: *American Journal of Public Health*, v. 73, p. 1366–1369.
- Narula, S.C., and Wellington, J.F., 1982, The minimum sum of absolute errors regression—State of the art survey: *International Statistical Review*, v. 50, p. 317–326.
- Pielou, E.C., 1984, *The interpretation of ecological data*: John Wiley and Sons, New York, 263p.
- R Development Core Team, 2012, *R—A language and environment for statistical computing*: Vienna, Austria, R Foundation for Statistical Computing, accessed May 05, 2013, at <http://www.R-Project.org>.
- Rao, J.S., 1976, Some tests based on arc-lengths for the circle: *Sankhya, Series B*, v. 38, p. 329–338.
- Reich, R.M., Mielke, P.W., Jr., and Hawksworth, F.G., 1991, Spatial analysis of Ponderosa pine trees infected with dwarf mistletoe: *Canadian Journal Forest Research*, v. 21, p. 1808–1815.
- Schroeder, R.L., and Vangilder, L.D., 1997, Tests of wildlife habitat models to evaluate oak mast production: *Wildlife Society Bulletin*, v. 25, p. 639–646.
- Seneta, E., 1983, The weighted median and multiple regression: *Australian Journal of Statistics*, v. 25, p. 370–377.
- Sherman, B., 1950, A random variable related to the spacing of sample values: *Annals of Mathematical Statistics*, v. 21, p. 339–361.

- Slauson, W.L., 1988, Graphical and statistical procedures for comparing habitat suitability data: U.S. Fish and Wildlife Service, Biological Report, v. 89, n. 6.
- Solow, A.R., 1989, A randomization test for independence of animal locations: *Ecology*, v. 70, p. 1546–1549.
- Terrell, J.W., Cade, B.S., Carpenter, J., and Thompson, J.M., 1996, Modeling stream fish habitat limitations from wedged-shaped patterns of variation in standing stock: *Transactions of the American Fisheries Society*, v. 125, p. 104–117.
- Tucker, D.F., Mielke, P.W., Jr., and Reiter, E.R., 1989, The verification of numerical models with multivariate randomized block permutation procedures: *Meteorology and Atmospheric Physics*, v. 40, p. 181–188.
- Van Valen, L., 1978, The statistics of variation: *Evolutionary Theory*, v. 4, p. 33–43.
- Whaley, F.S., 1983, The equivalence of three independently derived permutation procedures for testing the homogeneity of multidimensional samples: *Biometrics*, v. 39, p. 741–745.
- Wong, R.K.W., Chidambaram, N., and Mielke, P.W., 1983, Application of multi-response permutation procedures and median regression for covariate analyses of possible weather modification effects on hail responses: *Atmosphere-Ocean*, v. 21, p. 1–13.
- Zimmerman, G.M., Goetz, H., and Mielke, P.W., 1985, Use of an improved statistical method for group comparisons to study effects of prairie fire: *Ecology*, v. 66, p. 606–611.