

SWAP ALGORITHMS IN NULL MODEL ANALYSIS

Nicholas J. Gotelli^{1,3} and Gary L. Entsminger²

Null model analysis is an important research tool in community ecology (Gotelli 2001). Researchers compare community data with randomized data to ask how communities would appear if they were structured only by stochastic factors (Gotelli and Graves 1996). Such tests move beyond conventional statistical analyses, and provide a benchmark for patterns that might be expected in the absence of species interactions (Nitecki and Hoffman 1987). However, the use of null models in community ecology is controversial, and has been so since the first null model analyses appeared in the literature (e.g., Williams 1947).

Beginning with the exchanges between Diamond (1975) and Connor and Simberloff (1979), much of the controversy has centered on null model analysis of binary presence–absence matrices, in which each row is a species, each column is a site, and the entries indicate the presence (1) or absence (0) of a species at a site (Simberloff and Connor 1979). A standard approach has been to create null matrices in which the matrix elements are reshuffled, but the row and column totals of the original matrix are preserved (Connor and Simberloff 1979, Stone and Roberts 1990, Manly 1995, Sanderson et al. 1998).

However, creating a set of random matrices with fixed row and column totals is a challenging statistical problem. One cannot simply fill an empty matrix with randomly placed 1s, because eventually a point will be reached in which any further placements will violate either row or column totals. “Fill” and “swap” algorithms are the two solutions to this problem (Gotelli and Entsminger 2001). In a “fill” algorithm, an empty matrix is filled one cell at a time until row and column constraints are violated. At that point the algorithm backtracks by removing one of the previously filled

cells and continues forward until the matrix is successfully filled (Sanderson et al. 1998). “Swap” algorithm begin with the observed matrices as the starting point. They repeatedly swap subelements of the observed matrix to create new random matrices (Connor and Simberloff 1979). Both swap and fill algorithms generate reshuffled matrices that preserve row and column totals of the original matrix, although their statistical distributions are different from one another (Sanderson et al. 1998, Gotelli and Entsminger 2001), depending on the precise details of the algorithm used.

In a recent issue of *Ecology*, Manly and Sanderson (2002) criticized a swapping algorithm for null model analysis of presence–absence matrices presented by Gotelli (2000). In brief, Manly and Sanderson (2002) created a single, random matrix, tested it with the swapping algorithm in Gotelli (2000), and obtained a statistically significant result ($P < 0.05$). From this outcome, they concluded that the swapping algorithm was defective.

In this paper, we provide a more detailed analysis of the swapping algorithm and re-evaluate the conclusions of Manly and Sanderson (2002). Our main points are: (1) We could not reproduce their result. When we tested their random matrix with our implementation of Gotelli’s algorithm, we obtained the identical, nonsignificant result that Manly and Sanderson did. (2) We then created and tested 100 additional random matrices of the same dimension and fill as Manly and Sanderson’s single random matrix. When we tested this set of random matrices with Gotelli’s algorithm, we appropriately rejected the null hypothesis ($P < 0.05$ in each tail) for only 10 of the 100 random matrices. (3) Manly and Sanderson claimed that Gotelli’s method requires the calculation of the mean and normalized variance of the simulated data, but the method neither requires nor implements these calculations. These new analyses are consistent with previous benchmark tests of the swap algorithm (Gotelli 2000, Gotelli and Entsminger 2001). Our results provide no support for Manly and Sanderson’s (2002) conclusion that the swap algorithm is “irredeemably flawed.”

It is a challenging mathematical problem to construct a random matrix with a set of fixed row and column

Manuscript received and accepted 6 September 2002. Corresponding Editor: D. R. Strong.

¹ Department of Biology, University of Vermont, Burlington, Vermont 05405 USA

² Acquired Intelligence, Incorporated, 99 Schillhammer Road, Jericho, Vermont 05465 USA

³ Email: ngotelli@zoo.uvm.edu

TABLE 2. Consistent results of different swap algorithms.

Algorithm	Source	No. significant trials	Tail probability	Average <i>C</i> score	1 SD
Gotelli Swap	EcoSim 3.0	0	0.079 (0.015)	13.767 (0.009)	0.092 (0.004)
Gotelli Swap	EcoSim 7.0	0	0.078 (0.019)	13.756 (0.013)	0.124 (0.007)
Independent Swap	EcoSim 7.0	0	0.077 (0.008)	13.752 (0.004)	0.124 (0.007)
Manly Swap	Manly and Sanderson (2002)	3	0.071 (0.022)	13.750 (0.013)	0.127 (0.009)
Gotelli Swap	Manly and Sanderson (2002)	20	0.041 (0.009)	13.730 (0.012)	0.113 (0.006)

Notes: All tests use the random 15×15 matrix presented in Manly and Sanderson (2002: Table 1). Out of 20 simulation trials, results are presented for the number of significant results ($P < 0.05$), average P value, C score, and standard deviation. Standard deviations for each of these averages are given in parentheses. Only Manly and Sanderson's (2002) implementation of the Gotelli Swap generated nonrandom results for their random test matrix.

the 100 for which the null hypothesis was rejected ($P < 0.05$) in either tail of the distribution. For a set of 100 random matrices, a well-tempered null model should reject the null hypothesis for approximately 5 matrices in each tail of the distribution. We obtained this result for all 3 of the swap algorithms we tested. There was no evidence of an excessive Type I error rate for the Gotelli Swap, which confirms other benchmark tests (Gotelli 2000, Gotelli and Entsminger 2001).

Third, we were unable to reproduce the results reported in Manly and Sanderson (2002). When we applied the Gotelli Swap as implemented in EcoSim 3.0 to the random matrix they presented in their Table 1, we never rejected the null hypothesis in 20 trials, and generated an average P value of 0.079 (Table 1). The

00efT* [(378ot* [(37sults)-392(En426(tical0,-)19e)-383(obtais-192(Enbsimu).)(in)-32(En57(th1n)-g 0]nfT* [(00com274]TJ -lental T

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, London, UK.
