Magnetic Resonance in Medicine 51:418–422 (2004)

# Comparison of Fourier and Wavelet Resampling Methods

Angela R. Laird,[1] Baxter P. Rogers,[2] and M. Elizabeth Meyerand[2*]

**Resampling can be used to compute the null distribution of any test statistic for the purpose of measuring the significance of the measured value. This study investigated how well the spatial and temporal correlations of simulated and experimentally observed fMRI time series were preserved under Fourier and wavelet resampling methods. The null distributions of a test statistic estimated by each resampling method were compared. In addition, both resampling methods were applied to locate activated voxels in an fMRI dataset and ROC analysis showed that wavelet resampling performed more accurately than Fourier resampling. Magn Reson Med 51:418–422, 2004. © 2004 Wiley-Liss, Inc.**

**Key words: fMRI; wavelet resampling; Fourier resampling; statistical inference**

Frequently in fMRI data analysis, a test statistic is computed in order to quantify some parameter of interest and the null distribution of this test statistic is unknown. With the advance of high-speed computer processors, it is now possible to utilize nonparametric statistics that involve resampling the data in order to computationally obtain this null distribution. Resampling can be applied in time series analysis methods in order to generate multiple realizations of the process from which the original time series is assumed to originate. To obtain the null distribution of a test statistic using a resampling technique, the chosen test statistic is computed in the observed time series. Then the series is randomly resampled and the test statistic is computed in the resampled time series. This step is repeated many times and the distribution of values from the resampled series forms the null distribution of that test statistic (1). In order to obtain the most accurate estimate of the null distribution, it is important that the resampled time series possess the same statistical properties as the experimental data.

A crucial step in this method of statistical inference is determining the manner in which the data are to be resampled. The simplest technique is to shuffle or randomize the values of the time series. A version of this method was used by Biswal et al. (2) to obtain parameter confidence intervals. However, in order to justify the randomization, it is required that the data resampled is exchangeable (3,4). Thus, for observations that are not approximately independent, such as fMRI data, resampling in the time domain will fail to preserve the temporal autocorrelations.

An alternative to time domain resampling is to transform the observed discrete time series, $x[t_n]$, to the Fourier domain (5). Since the components of the Fourier transform are approximately independent at each frequency, fMRI data can be resampled in the Fourier domain. If a random phase shift, $\theta[f_n]$, is added to the Fourier phase $\phi[f_n]$ over all frequencies $f_n$, a resampled Fourier transform is constructed. After taking the inverse Fourier transform, the resampled time series has the same power spectrum as $x[t_n]$, and thus the same temporal autocorrelation function (5,6). However, when analyzing multivariate data it is important to also consider the spatial autocorrelations. Adding the same random phase shift to the Fourier phases of all voxels ensures the preservation of the spatial autocorrelations (6).

Schreiber and Schmitz (7) proposed an improvement to Fourier resampling based on iteratively refined amplitude-adjusted Fourier transform (AAFT) surrogates. In this method the resampled time series is iteratively corrected to have both the same power spectrum and amplitude distribution as the observed time series. These two properties cannot be exactly and simultaneously preserved. However, the iteration process should continue until the amplitude distribution is exactly preserved and the discrepancy between the power spectra is minimal.

In a recent study, Bullmore et al. (3) demonstrated that an alternative to Fourier resampling is to resample after transformation to the wavelet domain. To resample a time series using the wavelet method, the series is decomposed using the fourth-order Daubechies wavelet into its wavelet coefficients, which are approximately independent. This particular wavelet was determined to be optimal for fMRI time series, as it maximizes the decorrelation between the wavelet coefficients. The wavelet coefficients are permuted within each level of detail and then reconstructed to obtain the resampled time series. While wavelet resampling has been shown to effectively preserve the temporal autocorrelations present in fMRI data (3), no method has been suggested thus far for preserving the spatial autocorrelations. The purpose of this study was to investigate the preservation of these time series properties and to compare the relative strengths and weaknesses of Fourier and wavelet resampling techniques in fMRI studies.

## MATERIALS AND METHODS

In this study, data from three right-handed subjects were resampled using the Fourier and wavelet methods. Data were acquired on a 1.5 T GE Signa LX using the following scan parameters: flip angle 90°, TE/TR 40/2000 ms, FOV 24 × 24 cm, 22 slices, slice thickness 7 mm, gap 1 mm, and a 64 × 64 imaging matrix. The subjects performed an event-related finger-tapping exercise where visual cues for tapping were presented randomly in 2-sec blocks. A cluster of 25 voxels was randomly selected from the frontal

[1]Research Imaging Center, University of Texas Health Science Center, San Antonio, Texas.
[2]Department of Medical Physics, University of Wisconsin, Madison, Wisconsin.
*Correspondence to: M. Elizabeth Meyerand, 1530 Medical Sciences Center, 1300 University Avenue, Madison, WI 53706.
Received 15 November 2002; revised 18 August 2003; accepted 16 September 2003.
DOI 10.1002/mrm.10671
Published online in Wiley InterScience (www.interscience.wiley.com).

Table 1
Percentages of Comparisons for Which Temporal Autocorrelations Were Preserved Using the Five Methods of Fourier and Wavelet Resampling

| | Fourier resampling | AAFT resampling | Wavelet resampling—different permutation vector | Wavelet resampling—same permutation vector | 2D wavelet resampling |
|---|---|---|---|---|---|
| Dataset 1 | 100% | 100% | 92.8% | 93.2% | 79.2% |
| Dataset 2 | 100% | 100% | 91.2% | 90.8% | 78.8% |
| Dataset 3 | 100% | 100% | 94.8% | 94.6% | 81.2% |
| Average | 100% | 100% | 92.9% | 92.9% | 79.7% |

The numbers indicate the degree to which the observed temporal autocorrelations fit into the confidence intervals of the resampled temporal autocorrelations.

lobe to be resampled 1000 times using five methods of resampling: Fourier resampling, AAFT resampling, wavelet resampling using a different permutation vector at each voxel time series, wavelet resampling using the same permutation vector at each voxel time series, and 2D wavelet resampling.

Once the five resampled datasets were constructed, we compared the degree to which the temporal autocorrelations of the observed data were preserved in each set. The autocorrelation coefficients at 20 lags for each of the 25 time series were computed in the resampled datasets and the 95% confidence intervals over the 1000 resamplings were then defined. Our aim was to investigate if the autocorrelation coefficients of the observed data fit within the 95% confidence intervals of the resampled data over all lags for all voxels. Next, the spatial autocorrelations of the voxels in the cluster were compared in the same manner by computing the 95% confidence intervals of the $25 \times 25$ matrix of correlation coefficients for each voxel pair.

We then performed a separate analysis involving the resampling of a single time series using the 1D Fourier and wavelet methods. In this analysis, we constructed two simulated time series by adding a 3% signal change of a block task and an event-related task waveform to Gaussian noise. These simulated time series were resampled 5000 times. We computed the null distributions of the correlation coefficients for both methods of resampling. Our motivation for these computations was to determine if the obtained distributions were identical for both methods of resampling, as this is required if the assumptions of both methods are true. We also performed the same procedure on experimentally obtained time series data. A single voxel was selected out of the previously analyzed cluster and resampled 5000 times using both methods. Then a voxel was selected and resampled from a dataset obtained with exactly the same scan parameters as detailed above, with the exception of the utilization of a block task design. The null distributions were computed for the observed time series in the same manner as for the simulated time series.

Next, we performed two separate analyses on a whole-brain dataset of an event-related finger-tapping task (scan parameters detailed previously). Correlation coefficients were computed between the task waveform and the time series. The time series were resampled 1000 times via Fourier resampling and wavelet resampling (same permutation vector at each voxel) to compute two null distributions. To correct for the multiple comparisons problem, the maximum statistic in the brain was recorded for each resampling over all 1000 resamplings (a number deemed satisfactory by Nichols et al. (4)) to form the null distribution of that statistic. A threshold at the $\alpha = 0.05$ level was applied to the statistical images.

As a final investigation in the differences of Fourier and wavelet resampling, an ROC analysis was performed. Simulated data was created using whole-brain resting data. Activation was added to three clusters of voxels in the right and left motor cortex and supplementary motor area. Activation clusters consisted of 25 voxels, such that there was a total of 75 active voxels in the dataset. To accurately simulate the BOLD response to neural activity, activation was added to range from a 1–5% increase in signal intensity. Forty percent of the total active voxels were at 0–1%, 20% at 1–2%, 20% at 2–3%, 10% at 3–4%, and 10% at 4–5%. Once the simulated datasets were constructed, correlation coefficients between the voxel time series and the task waveform were computed. We resampled all time series using both the Fourier resampling method and the wavelet resampling method that uses the same vector to permute the coefficients. We used the computed distributions to calculate the ROC curves of true-positive rate vs. false-positive rate for both methods following the procedure of the established ROC technique (8,9).

## RESULTS

The results of the comparison of the preservation of the temporal and spatial autocorrelations of the five different methods of Fourier and wavelet resampling are shown in Tables 1 and 2. The percentages reported in Table 1 are percentages of the total comparisons for which the observed temporal autocorrelation coefficients fit into the resampled confidence intervals. As expected, the Fourier and AAFT resampling methods preserved these autocorrelations exactly. Wavelet resampling using the same and different permutation vectors performed similarly, preserving the temporal autocorrelations for an average of 92.9% of the comparisons, while the 2D method was less accurate, preserving only 79.7% of the comparisons.

Table 2 presents the percentages of the total comparisons for which the observed spatial autocorrelations fit into the resampled confidence intervals. Again, the Fourier and AAFT methods exactly preserved the autocorrelations. Wavelet resampling with a different permutation

Table 2
Percentages of Comparisons for Which the Observed Matrix of Spatial Autocorrelations Fit Into the Confidence Intervals of the Resampled Spatial Autocorrelations as Determined by the Five Methods of Fourier and Wavelet Resampling

| | Fourier resampling | AAFT resampling | Wavelet resampling—different permutation vector | Wavelet resampling—same permutation vector | 2D wavelet resampling |
|---|---|---|---|---|---|
| Dataset 1 | 100% | 100% | 44.0% | 92.3% | 63.0% |
| Dataset 2 | 100% | 100% | 73.0% | 90.7% | 71.0% |
| Dataset 3 | 100% | 100% | 65.7% | 93.0% | 75.3% |
| Average | 100% | 100% | 60.9% | 92.0% | 69.8% |

vector at each voxel preserved the correlation coefficients with an average comparison percentage of only 60.9%. Using the same permutation vector at each voxel greatly improved the ability to preserve the spatial relationships between voxel time series, with an average comparison percentage of 92.0%. 2D wavelet resampling did not perform as well, preserving an average of 69.8% of the total comparisons.

The next stage of this study investigated if identical null distributions were computed for a test statistic using Fourier and wavelet resampling. Figure 1 presents the null distributions obtained via both methods of 1D resampling for the simulated and experimentally observed block stimulus time series. It is clear from Fig. 1a,c that the two distributions obtained for the simulated time series did not agree. This was true also for Fig. 1b,d for the experimental data. In Fig. 2 the same sort of effect was seen for the event-related stimulus. The Kolmogorov-Smirnov test for determining if two samples come from populations with a common distribution was then applied to each of the four comparisons of Fourier and wavelet resampling. The null hypotheses that the distributions were the same for each type of resampling were rejected in all four cases ($\alpha = 0.05$). Quantile–quantile plots are shown in Fig. 3 to further illustrate this result. In these plots, the straight line (red) indicates the reference line. If two samples come from the same distribution, their qq-plot should fall along this line. The middle values of these plots characterize the bulk of the data, while the high and low values at the tails are of the most interest when classifying distributions. In three of the four cases, there was a clear deviation at the tails (Fig. 3a,c,d). The result of the fourth case in experimental data with a block task waveform was less convincing (Fig. 3b). However, even in this case the K-S test was able to reject the null hypothesis that the distributions were identical.

Statistical maps are shown in Fig. 4 ($\alpha = 0.05$) for the results of the correlation analysis using Fourier and wavelet resampling as a method of statistical inference. Correlation values ranged from −0.5334 to 0.8009 in the observed data. Fourier resampling determined the minimum critical value to be −0.5735 and the maximum critical value to be 0.5872. Thus, 34 voxels were determined to be positively correlated at the selected level of significance, and no voxels were negatively correlated. The voxels that were significantly correlated were located in a cluster in the left motor cortex. Wavelet resampling determined the critical values to be −0.4239 and 0.4254. In this case, 83

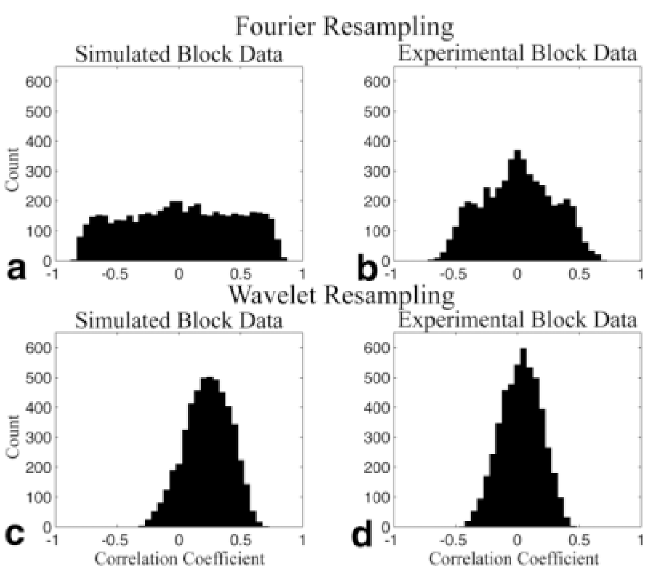



FIG. 1. Histograms of the distributions of the correlation coefficients with the block task waveform for simulated and experimental data obtained via 1D Fourier (**a,b**) and wavelet resampling (**c,d**). The distributions determined by the two different methods are not identical.

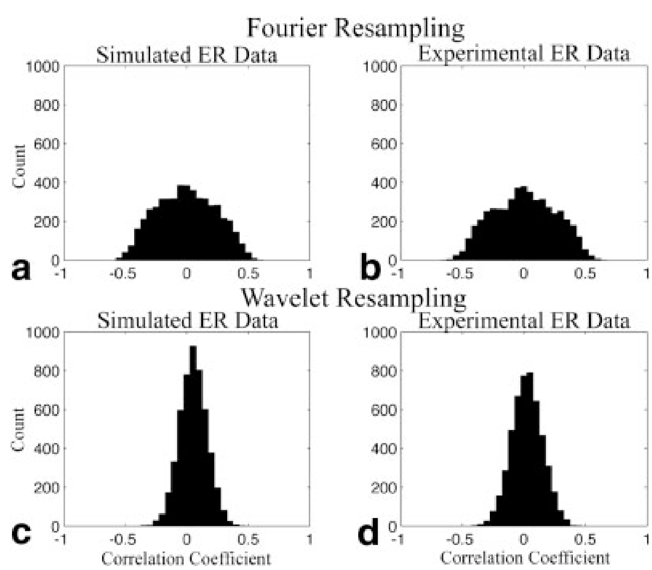



FIG. 2. Histograms of the distributions of the correlation coefficients with the event-related task waveform for simulated and experimental data obtained via 1D Fourier (**a,b**) and wavelet resampling (**c,d**). Again, the distributions determined by the two methods differ.

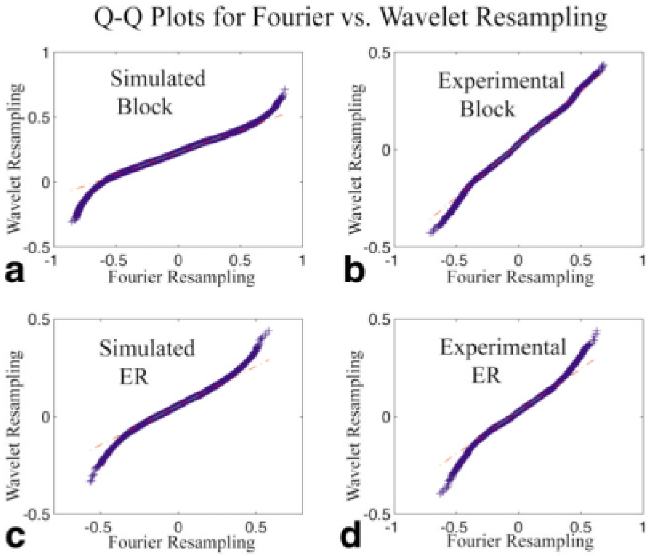



FIG. 3. Quantile–quantile plots for the distributions computed via 1D Fourier and wavelet resampling. The reference line (red) indicates the case of identical distributions and the values at the tails are of the most interest when classifying distributions. In three of the four plots seen there is a clear deviation at the tails (**a,c,d**). The result of the fourth case in experimental data with a block task waveform is less convincing (**b**). However, the two-sample Kolmogorov-Smirnov tests rejected the null hypothesis that the distributions were identical for all four cases ($\alpha = 0.05$).

voxels were determined to be positively correlated and 21 were negatively correlated, for a total of 104 voxels that were significantly correlated with the reference waveform. A cluster of these voxels was located in the left motor cortex and this cluster was somewhat larger than the corresponding one found via Fourier resampling. In addition, two positively correlated voxels were found in the supplementary motor area: a positively correlated cluster was found in the ipsilateral cerebellum and a negatively correlated cluster was found in the right motor cortex.

Our last investigation of the differences between Fourier and wavelet resampling was the ROC analysis. The results of that analysis are shown in Fig. 5 and the plot of true-positive rate vs. false-positive rate clearly shows a greater performance of the wavelet resampling method in the analysis of an fMRI whole-brain activation dataset.

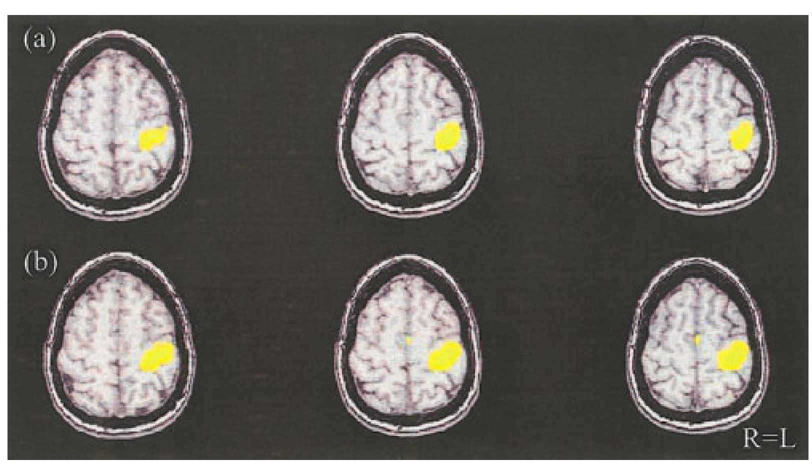
## DISCUSSION

The purpose of this study was to compare the different techniques of Fourier and wavelet resampling such that researchers may more accurately determine which approach is applicable to their individual needs. We first tested how the Fourier and wavelet resampling methods were able to preserve the temporal and spatial autocorrelations of fMRI time series. It is important to reiterate that preservation of the spatiotemporal properties of the time series under investigation is crucial when using the nonparametric methods to computationally determine the proper null distribution of the test statistic. In this particular analysis, the cluster investigated was randomly selected from a region in the frontal lobe. As the method of resampling follows a hypothesis-testing framework, we felt it necessary to investigate the preservation of the spatiotemporal properties of this inactive region of the brain, a region under the null hypothesis. It is possible that the results presented are influenced by the region of the brain studied. More work is needed to determine if there is a spatial heterogeneity to the preservation of temporal and spatial autocorrelations of resampled fMRI time series.

In addition, we computed the null distributions for the correlation coefficient between the task waveform and the observed time series using both methods in order to verify that the distributions obtained were not identical for either the simulated or experimental time series analyzed, using both block and event-related task designs. Although not the preferred method of analyzing fMRI data, correlation analysis was selected in this study due to its simplicity, ease of use, and familiarity to the fMRI community. For the analyses of a real dataset, it appeared that the method of Fourier resampling was more conservative than wavelet resampling at the 5% significance level. That is, fewer voxels were significantly correlated with the reference waveform according to the results of the analysis of this method.

We conclude, based on the results of the ROC analysis, that wavelet resampling is a superior method to Fourier resampling in the analysis of fMRI data. Although we

FIG. 4. Significant positive correlations between the reference waveform and the voxel time series in contiguous axial slices as determined by Fourier resampling (**a**) and the method of wavelet resampling that uses the same permutation vector at each voxel (**b**).



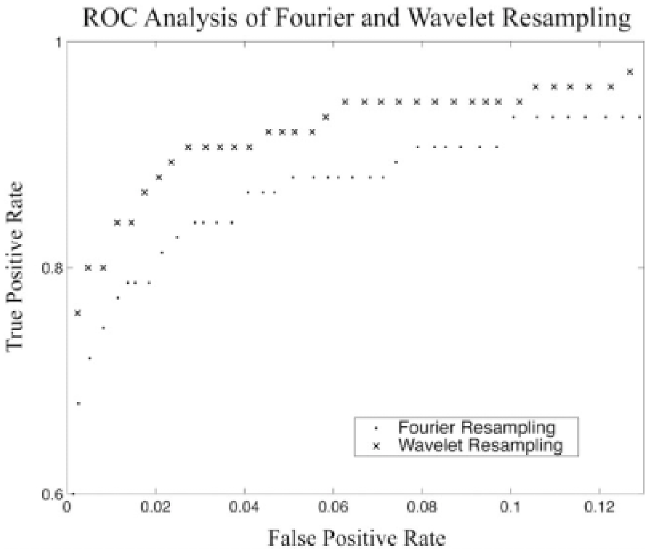



FIG. 5. ROC curve of true-positive rate vs. false-positive rate for Fourier and wavelet resampling. It is apparent from this curve that wavelet resampling performs better as a method of statistical inference in fMRI than Fourier resampling.

followed the traditional ROC method in fMRI, it is unclear at this point if the chosen analysis parameters adversely affected the outcome. For instance, it is unknown whether the cluster size, location, or activation amplitude influenced the results in any way. Further details should be investigated in determining the relationship between these factors and ROC results in fMRI analysis methods.

## REFERENCES

1. Good P. Permutation tests: a practical guide to resampling methods for testing hypotheses. New York: Springer; 1994.
2. Biswal BB, Taylor PA, Ulmer JL. Use of jackknife resampling techniques to estimate the confidence intervals of fMRI parameters. J Comput Assist Tomogr 2001;25:113–120.
3. Bullmore E, Long C, Suckling J, Fadili J, Calvert G, Zelaya F, Carpenter TA, Brammer M. Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. Hum Brain Mapp 2001;12:61–78.
4. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 2002;15:1–25.
5. Davison AC, Hinkley DV. Bootstrap methods and their applications. Cambridge, UK: Cambridge University Press; 1997.
6. Prichard D, Theiler J. Generating surrogate data for time-series with several simultaneously measured variables. Phys Rev Lett 1994;73:951–954.
7. Schreiber T, Schmitz A. Improved surrogate data for nonlinearity tests. Phys Rev Lett 1996;77:635–638.
8. Constable RT, Skudlarski P, Mencl E, Pugh KR, Fulbright RK, Lacadie C, Shaywitz SE, Shaywitz BA. Quantifying and comparing region-of-interest activation patterns in functional brain MR imaging: methodology considerations. Magn Reson Imag 1998;16:289–300.
9. Skudlarski P, Constable RT, Gore JC. ROC analysis of statistical methods used in functional MRI: individual subjects. Neuroimage 1999;9:311–329.