Running head: SERIAL CORRELATIONS

# Methods for detecting 1/$f$ noise

Simon Farrell

University of Bristol

Eric-Jan Wagenmakers

University of Amsterdam

Roger Ratcliff

The Ohio State University

Dr. Simon Farrell

Department of Psychology

University of Bristol

8 Woodland Road

Clifton, Bristol BS8 1TN

UK

e-mail: Simon.Farrell@bristol.ac.uk

fax: +44 (0)117 928 8588

**Abstract**

This unpublished document describes simulations comparing the spectral classifier of T. Thornton and D. L. Gilden (2005) for detecting $1/f$ noise to the ARFIMA model selection approach endorsed by E.-J. Wagenmakers, S. Farrell, & R. Ratcliff (2004).

## Methods for detecting 1/*f* noise

## Methods for detecting 1/*f* noise

Thornton and Gilden (2005) are primarily concerned with the presentation of a method for distinguishing 1/*f* noise from stochastic processes that may mimic the statistical properties of 1/*f* noise, but are not true 1/*f* noise; that is, distinguishing persistent serial correlations from transient correlations. The method presented by Thornton and Gilden is a spectral classifier, in which the likelihood of a time series (or set of time series) is estimated by comparing the power spectrum (the frequency-domain representation) of a time series to a library of spectra derived from two candidate models of serial correlations. One model considered by Thornton and Gilden is 1/*f* noise, treated by Thornton and Gilden as fractional Brownian motion (fBm) with added Gaussian white noise (white noise possessing no systematic serial correlations). This added white noise is sometimes interpreted as independent variability in motor processes (e.g., Gilden, 1997). The other model they consider is an autoregressive moving average model, the ARMA(1, 1) model, in which the value of a series at time $t$ depends only on the state of the system at time $t-1$; that is,

$$X_t = \phi X_{t-1} + \epsilon_t + \theta \epsilon_{t-1} \tag{1}$$

For a full description of ARMA models, see Brockwell and Davis (1996) or Wagenmakers, Farrell, and Ratcliff (2004). The purpose in comparing these two models is that, although possessing only transient correlations, the temporal statistics of an ARMA(1, 1) time series can resemble those of 1/*f* noise (Wagenmakers et al., 2004). Thornton and Gilden present some simulation results showing that their method is well suited to classifying series generated from these two specific models.

It is clear that the method presented by Thornton and Gilden (2005) is more rigorous than the method that has been previously employed. Previous investigations (Gilden, 1997, 2001; Gilden & Wilson, 1995; Van Orden, Holden, & Turvey, 2003) of serial correlations in psychology have used an approximate procedure in which a power spectrum is constructed for a time series, and a line fit to this spectrum in log-log co-ordinates to determine the extent of $1/f$ noise in the series (see Gilden, 2001; Wagenmakers et al., 2004 for a description). Wagenmakers et al. (2004) showed that this method was inappropriate, as short-range stochastic processes not possessing the characteristics of $1/f$ noise could be misidentified as $1/f$ noise. Accordingly, Wagenmakers et al. (2004), Wagenmakers, Farrell, and Ratcliff (2005) argued that $1/f$ cannot be considered in isolation, but must be accompanied by examination of alternative models, such as the ARMA model, that can give rise to similar temporal patterns as $1/f$ noise, but that do not possess long-range dependence. It is promising to see recognition of this important point in the method presented by Thornton and Gilden.

Despite the implication in Thornton and Gilden's presentation, however, it is not clear that their spectral classification method is also superior to a method previously suggested for distinguishing persistent and transient correlations. Wagenmakers et al. (2004) presented a method in which short-range processes, represented by the ARMA model, are compared to an extension of the ARMA model that incorporates long-range dependencies (see also Beran, Bhansali, & Ocker, 1998). This ARFIMA model (fractionally integrated ARMA model; see, e.g., Beran, 1994) incorporates an additional parameter that scales the extent of long-range dependence (see Appendix of Wagenmakers et al., 2004 for a formal description). Wagenmakers et al. advocated an approach in which ARMA and ARFIMA models were competitively tested against each other in a model selection framework. Specifically, Wagenmakers et al. advocated determination of the maximum likelihood of a time series under the ARMA model and the ARFIMA model, and then selection of one of these models using an information metric such as Akaike's information criterion (AIC; Akaike, 1974).

Thornton and Gilden (2005) claim that their spectral classifier method is superior to the ARMA/ARFIMA comparison framework. However, Thornton and Gilden did not rigorously compare their approach to model selection using the ARMA/ARFIMA framework. In the following simulations, we compare the performance of Thornton and Gilden's spectral classifier to that of the ARMA/ARFIMA modelling framework. In previous applications of the ARFIMA method, we have compared the ARMA(1, 1) to the ARFIMA(1, $d$, 1) model as an example of the modelling approach (Wagenmakers et al., 2004), and have also compared a wide range of ARMA and ARFIMA models (Wagenmakers et al., 2005). To highlight the similarity between the spectral classifier approach and the ARFIMA methods previously applied, we restrict the method to competitively test the ARFIMA(0, $d$, 0) model against an ARMA(1, 1) model (note that these models are not nested).[1] The ARFIMA(0, $d$, 0) model is essentially identical to the fBm plus white noise (fBmW) model considered by Thornton and Gilden: Both ARFIMA(0, $d$, 0) and fBmW are characterised by a spectrum that is linear in log-log co-ordinates, with some whitening at high frequencies.[2]

*Details of the simulations*

Simulations[3] were run in a replication of the maximum likelihood (ML) simulations reported by Thornton and Gilden (2005), but comparing the spectral classifier to the ARFIMA modelling approach. We investigated the same parameter sets as Thornton and Gilden: in 20 steps, $\alpha$ was varied linearly over the range [0, 1], $\beta$ was varied linearly over the range [0, 2], $\phi$ was varied linearly over the range [.1, .95], and $\theta$ varying non-linearly according to $\theta = -.0854k + .002k^2$, where $k = 0...19$. Under each point in the 20×20 fBmW and ARMA(1, 1) parameter spaces, 1000 series were generated and given to the spectral classifier and the ARFIMA method for analysis. One-over-$f$ (fBmW) series were generated using spectral synthesis as in Thornton and Gilden (2005); spectral synthesis involves construction of a Fourier spectrum from given parameter values, and generation of a pseudo-random time series from this spectrum using the Discrete Fourier Transform (DFT)[4]. ARMA(1, 1) series were generated using time-domain filtering, using a lead-in of 500 trials to avoid non-stationarity.

*Spectral classifier.* The spectral classifier was implemented as in Thornton and Gilden (2005). First, for each generating model, 2000 series of 1024 observations were generated for each point in the 20×20 parameter space. Each set of series then allowed the calculation of a mean spectrum and spectrum covariance matrix (containing the variance of mean spectrum values, and the covariance in spectral power between different frequencies).

For the benchmarking simulations, the log-likelihood (ln$L$) of each test series was calculated under the fBmW and ARMA(1, 1) models.[5] The maximum ln$L$ and associated parameter estimates for each test series were obtained by finding the point in the 20×20 parameter space with the highest associated ln$L$. The test series was then classified as fBmW or ARMA(1, 1) respectively by determining whether the difference in ln$L$ between the two models ($\ln L_{fBmW} - \ln L_{ARMA(1,\ 1)}$) was $\geq 0$.

*ARFIMA method.* For each test series, maximum ln$L$ estimates for both the ARMA(1, 1) and ARFIMA(0, $d$, 0) models were found using the `fracdiff` package in R (Maechler, 2005). This package uses an approximate ML procedure (Haslett & Raftery, 1989) to estimate the parameters of the ARFIMA($p$, $d$, $q$) model.[6] The ARMA(1, 1) model was estimated by fitting an ARFIMA(1, $d$, 1), restricting $d$ to the range $[0, 10^{-8}]$; this was done to ensure that differences in fit between the ARMA and ARFIMA were due only to the models, and not the method used to fit them. Series were classified as originating from the fBmW or ARMA(1, 1) model by calculation of the AIC metric from the resulting maximum ln$L$ values by

$$AIC = -2\ln L + 2V, \tag{2}$$

where $V$ is the number of free parameters in the model, and was 2 for the ARFIMA(0, $d$, 0) model and 3 for the ARMA(1, 1) model (as each model incorporates an addition parameter for the variance of innovations). A series was classified as fBmW by determining whether $AIC_{ARFIMA} \leq AIC_{ARMA(1,\ 1)}$.

*Simulation results*

Figure 1 shows the classification proportions for the spectral classifier, while Figure 2 shows classification proportions for the ARFIMA method. In each figure, the panel on the left shows the proportion of times fBmW series were correctly classified as long-range dependent ($1/f$), while the panel on the right shows the proportion of ARMA(1, 1) series that were mistakenly classified as $1/f$ noise. Comparison of the two models reveals the ARFIMA method to be more accurate at classifying fBmW series (giving high rates of "yes" classifications for these series) and less accurate at classifying ARMA(1, 1) models, giving high false alarm rates in these cases.

Although these figures are informative in comparing results to those reported in Thornton and Gilden (2005), they do not indicate which of the methods, if any, is more accurate overall. A more accurate picture of the data is given in Figure 3. This figure shows distributions of ln$L$ differences for the two generating models. Each data point entering the distributions is the difference in ln$L$ between the fitted fBmW and ARMA models (spectral classifier: top panel) or the ARFIMA and ARMA models (ARFIMA method: bottom panel). The figure shows that the two methods give strikingly similar results. For both procedures, the distribution of ln$L$ differences for series from the fBmW model is approximately symmetric and centred just above 0. For the ARMA(1, 1) generating model, the ln$L$ differences are negatively skewed, with a peak at around 0.

To reinforce this point, Figure 4 shows the receiver-operating characteristic (ROC) function for the two methods. Each ROC curve was constructed by varying the criterion, along the dimension of ln$L$ difference, by which series were classified as fBmW or ARMA, and determining hit rates (proportion of fBmW series above the criterion) and false alarm rates (proportion of ARMA series above the criterion). Figure 4 shows that the two methods give overwhelmingly similar results. Also shown in Figure 4 is the area under the ROC curve ($A$) for each method. The $A$ values differ only trivially, highlighting the equivalent accuracy of classification of the spectral classifier and ARFIMA method.

*Spectral classifier and ARFIMA: Summary*

In summary, the claims of Thornton and Gilden (2005) are unfounded: Both the spectral classifier and ARFIMA method involve comparison of long-range dependent models to alternative short-range models, and are equally effective in discriminating between transient and persistent correlations. There are differences in implementation in each approach: fBmW involves ML estimation in the frequency domain, while the ARFIMA method involves ML estimation in the time domain (using the autocovariance function; Haslett & Raftery, 1989). The ARFIMA method is also more desirable on pragmatic grounds: Unlike the spectral classifier, it does not require re-computation of a library of covariances if series of a different length are examined, or if one wishes to examine additional models [e.g., ARFIMA(1, $d$, 1)].

## References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.

Beran, J. (1994). *Statistics for long-memory processes.* New York, NY: Chapman & Hall.

Beran, J., Bhansali, R. J., & Ocker, D. (1998). On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. *Biometrika*, *85*, 921-934.

Brockwell, P. J., & Davis, R. A. (1996). *Introduction to time series and forecasting.* New York, NY: Springer.

Gilden, D. L. (1997). Fluctuations in the time required for elementary decisions. *Psychological Science*, *8*, 296-301.

Gilden, D. L. (2001). Cognitive emissions of $1/f$ noise. *Psychological Review*, *108*, 33-56.

Gilden, D. L., & Wilson, S. G. (1995). Streaks in skilled performance. *Psychonomic Bulletin & Review*, *2*, 260-265.

Haslett, J., & Raftery, A. E. (1989). Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. *Journal of the Royal Statistical Society (C)*, *38*, 1-50.

Maechler, M. (2005). *fracdiff: Fractionally differenced ARIMA(*p, d, q*) models.* R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org.

Thornton, T. L., & Gilden, D. L. (2005). Provenance of correlations in psychological data. *Psychonomic Bulletin & Review*, *12*, 409-441.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, *132*, 331-350.

Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of $1/f$ noise in human cognition. *Psychonomic Bulletin & Review*, *11*, 579-615.

Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2005). Human cognition and a pile of sand: A discussion on serial correlations and self-organized criticality. *Journal of Experimental Psychology: General*, *134*, 108-116.

## Author Note

The second author was supported by a VENI grant from the Dutch Organization for Scientific Research (NWO). The third author was supported by NIMH grant R37-MH44640. We are very grateful for Tom Thornton's assistance in clarifying the specifics of the spectral classification method.

## Footnotes

[1]Note that Thornton and Gilden devote several pages to criticism based on the supposed necessity of nesting in the ARMA/ARFIMA framework, but then acknowledge in their Footnote 4 that there is no necessity for nesting in the ARFIMA method

[2]The form of the spectral density of the ARFIMA(0, $d$, 0) model is given by:

$$S(f) \sim |1 - e^{-if}|^{-2d} \tag{3}$$

(e.g., Beran, 1994). Note that $2d = \alpha$ from the fBmW model (Beran, 1994).

[3]To allow other researchers to verify these results, and to make both methods freely available to those interested in analysing their own data, we have made MATLAB and R code available on the web, at http://seis.bris.ac.uk/∼pssaf.

[4]Examination of series of length $2^k$ allows use of the Fast Fourier Transform (FFT), which is much faster than the DFT.

[5]The log-likelihood was calculated, under the assumption of Gaussian variates, as

$$\ln L = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{\Sigma}| - \frac{1}{2}\mathbf{v}^T\mathbf{\Sigma}^{-1}\mathbf{v}, \tag{4}$$

where $N$ is the number of points in the spectrum ($N$=8 for series of length 1024), $|\mathbf{\Sigma}|$ is the determinant of the covariance matrix, $\mathbf{v}$ is a vector of differences between the predicted and observed spectrum values, and $\mathbf{\Sigma}^{-1}$ is the inverse of the covariance matrix.

[6]We have previously advocated the use of exact ML estimation of ARFIMA models using the Ox program. We use the R method for ARFIMA fitting here as the approximate ML estimation procedure is much faster; this makes little difference in application to individual series, but is appreciably faster for the simulations conducted here.

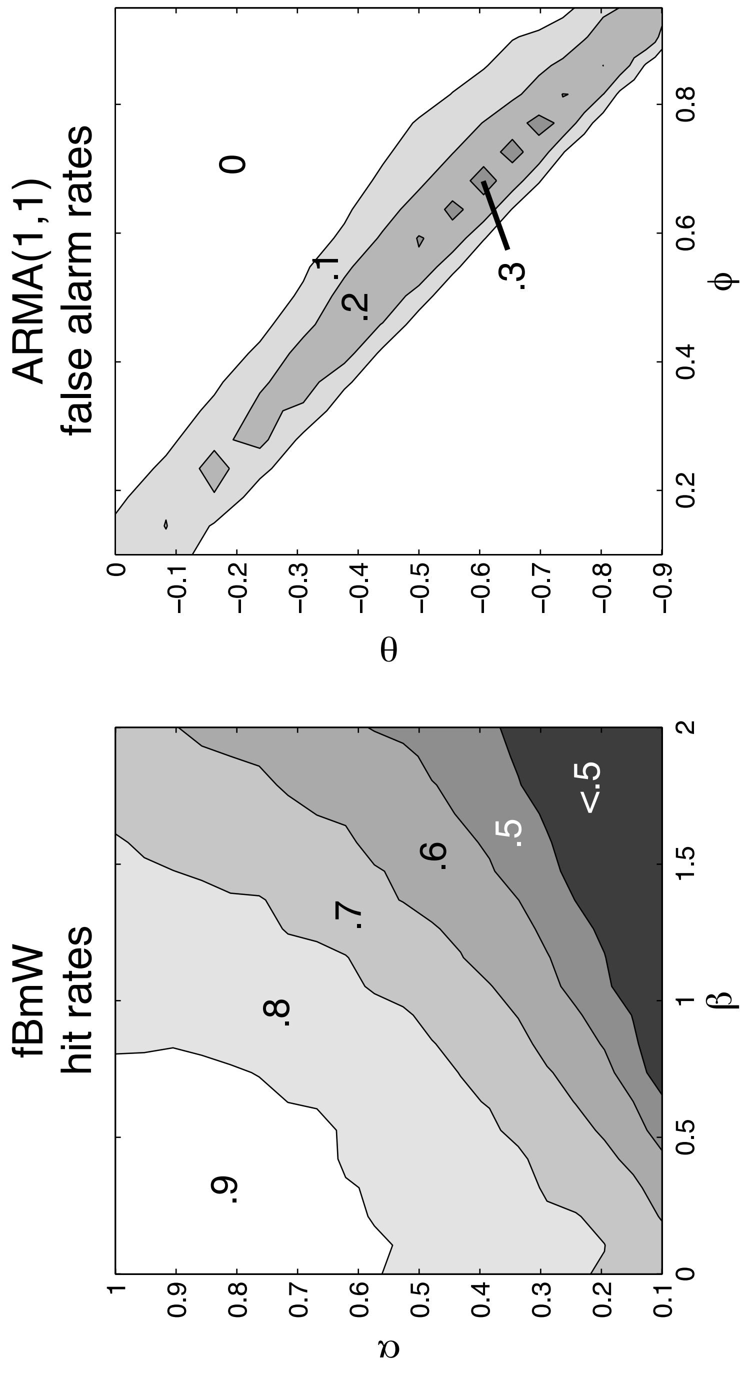
## Figure Captions

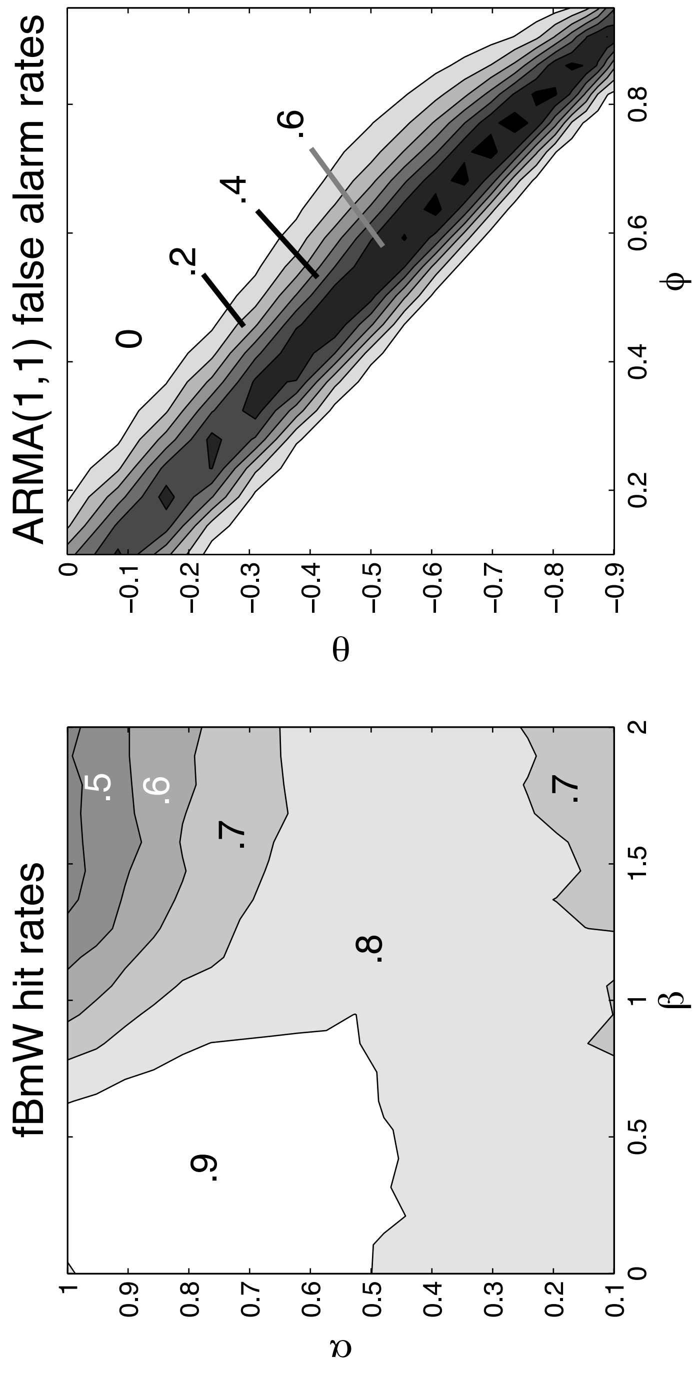
*Figure 1.* Contour plots of the probability of accepting a series as fBmW under the spectral classifier. Series were generated from the fBmW model (left panel) and the ARMA(1, 1) model (right panel). For Figures 1 and 2, bands increase in increments of .1, and the number in each band indicates the bottom end of the band (e.g., .7 indicates detection rates varied from .7 to .8).

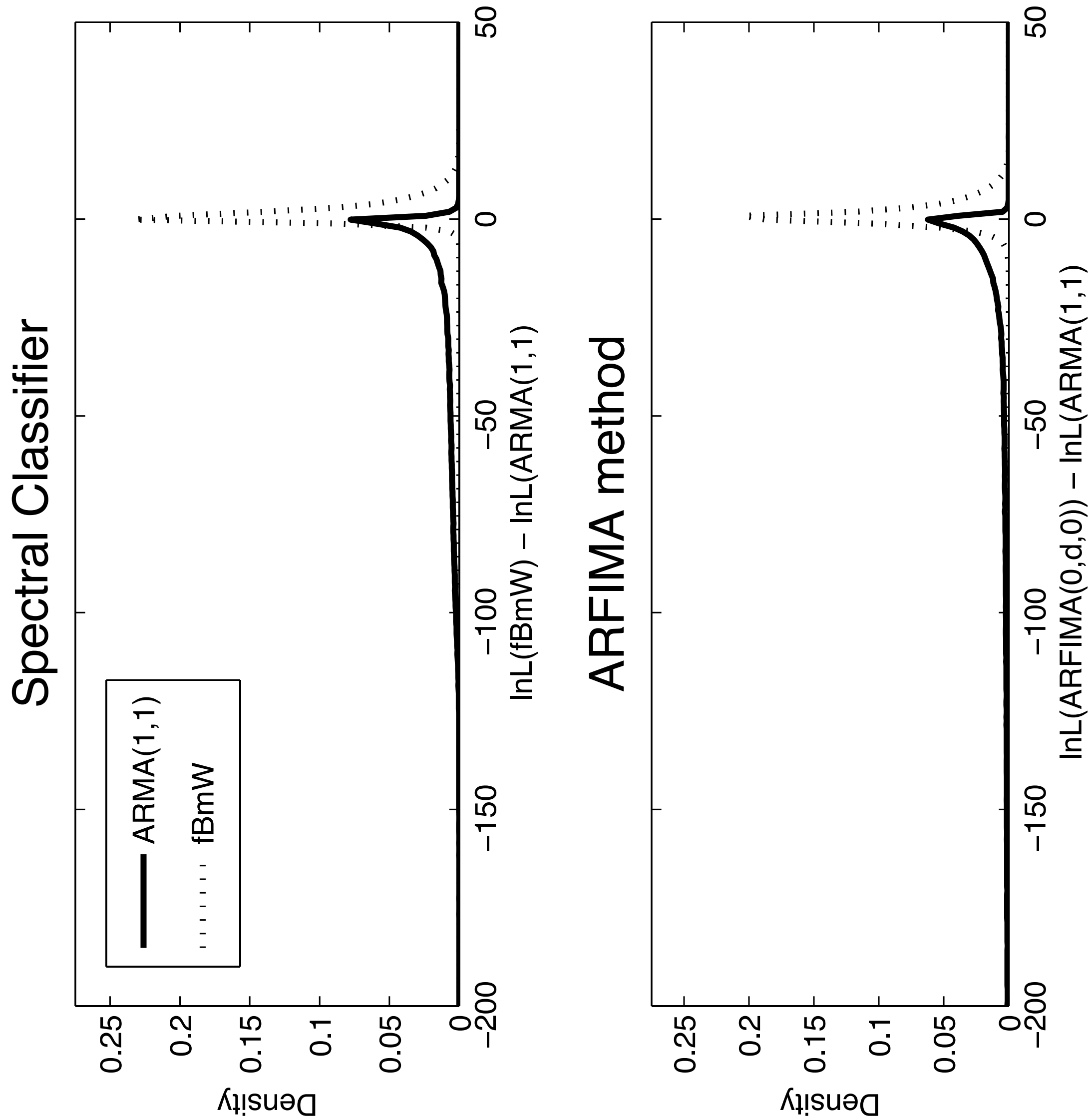
*Figure 2.* Contour plots of the probability of accepting a series as fBmW under the ARFIMA method. Series were generated from the fBmW model (left panel) and the ARMA(1, 1) model (right panel).

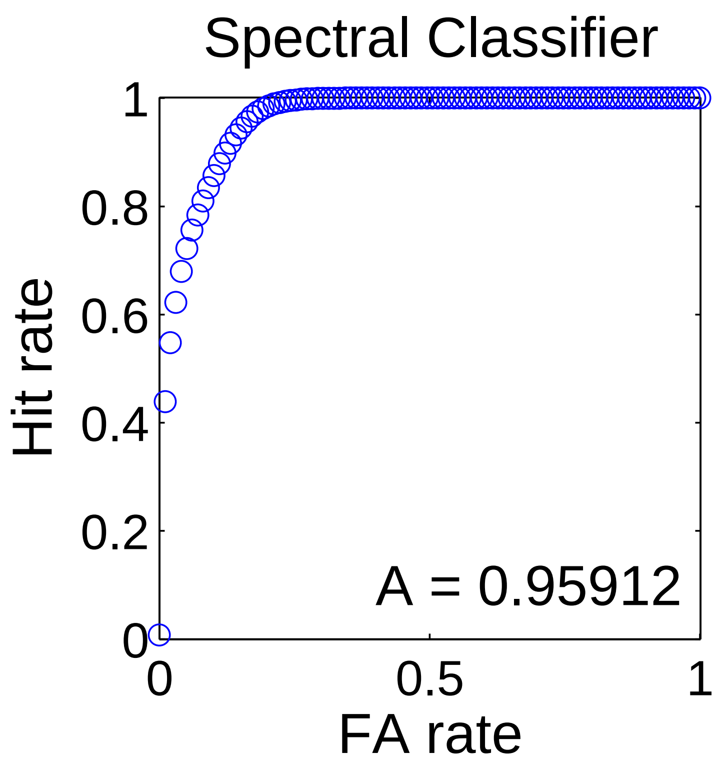
*Figure 3.* Distributions of differences in log-likelihood for the fBmW and ARMA(1, 1) generating models. Each observation entering the distributions is the difference in $\ln L$ between the two fitted models. The top panel shows the difference distributions for the spectral classifier ($\ln L_{fBmW} - \ln L_{ARMA(1,\ 1)}$); the bottom panel shows the difference distributions for the ARFIMA method ($\ln L_{ARFIMA(0,\ d,\ 0)} - \ln L_{ARMA(1,\ 1)}$).

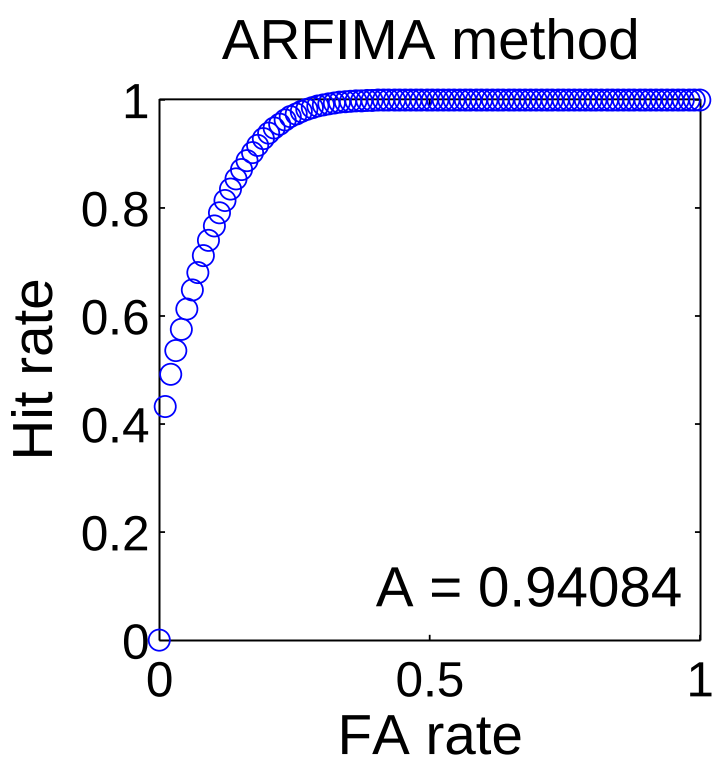
*Figure 4.* ROC functions generated from the difference distributions in Figure 3. Each point was calculated by choosing a criterion (along the dimension of $\ln L$ differences) and taking the area under the fBmW (top panel) or ARFIMA (bottom panel) distribution above the criterion as the hit rate, and the area under the ARMA(1, 1) distribution above the criterion as the false alarm (FA) rate. Criteria were chosen to give a constant spacing of FA rates.


fBmW
hit rates
α
β
.9
.8
.7
.6
.5
<.5
ARMA(1,1)
false alarm rates
θ
φ
0
.1
.2
.3


ARMA(1,1) false alarm rates
θ
φ
fBmW hit rates
α
β


Spectral Classifier
ARMA(1,1)
fBmW
Density
lnL(fBmW) − lnL(ARMA(1,1))
ARFIMA method
Density
lnL(ARFIMA(0,d,0)) − lnL(ARMA(1,1))


Spectral Classifier
1
0.8
0.6
0.4
0.2
0
Hit rate
A = 0.95912
0
0.5
1
FA rate


ARFIMA method
1
0.8
0.6
0.4
0.2
0
Hit rate
A = 0.94084
0
0.5
1
FA rate