USING ATS TO TURN TIME SERIES ESTIMATION AND MODEL DIAGNOSTICS INTO FAST REGRESSION ESTIMATION AND MODEL DIAGNOSTICS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Jeremy Troisi

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. William S. Cleveland, Chair

Department of Statistics

Dr. Bowei Xi

Department of Statistics

Dr. Anindya Bhadra

Department of Statistics

Dr. Wen-Wen Tung

Department of Earth and Atmospheric Sciences

Approved by:

Dr. Jun Xie

Head of the Departmental Graduate Program

I dedicate this to my dad. Although there were obviously numerous others that were in my corner in the struggle to bring this work to a conclusion, he put time and effort to help me at the times I needed the support the most.

ACKNOWLEDGMENTS

I am grateful to have had the opportunity to work with Professor Cleveland. His approach to data analysis is principled, practical, and effective.

I would like to thank my committee members Professors Bowei Xi, Anindya Bhadra, and Wen-Wen Tung for reviewing my dissertation and my defense.

I am grateful for the guidance and assistance of many others in the department, particularly Professor Hao Zhang for his insistence in completing my dissertation while remotely embedded within my career and Doug Crabill for being the technical guru that he is and utilizing that knowledge to help aspiring PhD candidates use the computational resources they need to complete their dissertations. I am grateful for the work of Patti Foster, Mary Sigman, and Holly Graef for running things so smoothly.

I am grateful to Marty Schnepp, Alex Mann, and numerous other teachers at Holt High School for getting me started on my path towards pursuing advanced scientific studies.

Finally, I am grateful to my dad, my girlfriend Denise, and the rest of my family and friends for their support.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

ABSTRACT

Troisi, Jeremy Ph.D., Purdue University, May 2019. Using ATS to Turn Time Series Estimation and Model Diagnostics into Fast Regression Estimation and Model Diagnostics. Major Professor: William S. Cleveland.

The Average Transform Smooth (ATS) statistical methods [McRae, Mallows, and Cleveland], [1], are applied to measurements of a non-gaussian random variable to make them close to gaussian. This gaussianization makes use of the well known concept of variance stabilizing transformation, but takes it further by first averaging blocks of r measurements, transforming next, and then smoothing. The smoothing can be nonparametric, or can be the fitting of a parametric model. The gaussianization makes analysis simpler and more effective.

In this work ATS is applied to the periodogram of a stationary parametric time series, and makes use of the periodogram large sample properties given the true power spectrum [Brillinger], [2], to develop a new approach to parametric time series model estimation and model diagnostics. The ATS results and the theory are reformulated as a regression model, PPS-REG, involving true power spectrum and the periodogram. PPS-REG has attractive properties: iid gaussian error terms with mean 0 and a known variance; accurate estimation; much faster estimation than the classical maximum likelihood when the time series is large; enables the use of the very powerful classical regression model diagnostics; bases the diagnostics on the power spectrum, adding substantially to the standard use of the autocovariance function for diagnosing the fits of models specified in the time domain.

1. PERIODOGRAM THEORY AND ATS RESULTS

This work develops new methods of estimation and of model diagnostics to fit parametric models to numeric time series $x_0, \ldots, x_{n-1}$. To make notation simpler we will assume, without loss of generality, that n is even, in part, because we can easily insure this in practice. The first step is to compute the periodogram $I(f_j)$ at the Fourier frequencies, $f_j, j = 0 \ldots n/2$. For time series with large n that is not very small, computing I using the FFT makes sense. This is provided routinely by time series software. Our approach is based on ATS and theoretical results for the periodogram of Brillinger as described in the abstract.

We start by invoking the theoretical results of the theory of I. Let $S(f)$ be the true power spectrum of the time series model for $0 \leq f \leq 1/2$. Then the $I(f_j)/S(f_j)$ for all j are independent chi-squared distributions. For $j = 0, n/2$ the degrees of freedom are 1 and for j otherwise they are 2. For this reason we will discard the periodogram values for frequencies 0 and $n/2$. In fact, because it makes sense to subtract the mean of the data in a stationary time series analysis, the periodogram at the DC frequency 0 is 0.

Next we take the mean of the periodogram in disjoint blocks of Fourier frequencies, each with r consecutive values of f_j starting with $j = 1$ and ending with the last complete block, and not including $f_{n/2}$. We divide each block average $S(f)$ where f is the midpoint of the r frequencies of the block.

The transformation we take is ln, the natural log. This is quite natural because the log of the power spectrum is typically the one studied in applications. The question is what should r be? If r is large, we risk introducing bias in the power spectrum. If r is too small we risk poor gaussianization. Here we appeal to the work of Cleveland, Mallows, and McRae, [1], who found that for an r as small as 4 the n ln block averages are well approximated by independent normals with constant variances. This is great news because

it means, especially with large n, that there will be little bias. So we use $r = 4$ in the regression model.

2. FORMULATION OF THE REGRESSION MODEL

The ATS power spectrum results have been used to carry out estimation of the power spectrum itself using non-parametric regression. This is quite useful for applications where frequency analysis is critical to subject matter study, for example atmospheric science.

Here we extend it further by using it to estimate the parameters of time domain parametric models. As discussed earlier, we do this by expressing the ATS results as a regression model with gaussian errors, and estimate the parameters using regression methods. This leads also to model diagnostics in the frequency domain, departing from the standard, which is time domain diagnostics using autocorrelations. The former is far better than the latter because the statistical variability of the former is much simpler than for the latter.

Let the averaged periodogram values be

$\overline{I(f_k^*)}$ for $k = 1 \ldots b$, where b is the number of blocks.

Let β be the unknown parameters of the regression model

Let $y_k = \ln(\overline{I(f_j)})$ be the response in the regression model.

Then the model is

$$y_k = S(f_k^*) - \psi(4) + ln(4) + \varepsilon$$

where ε are error terms with mean 0 and variance $\psi^{(1)}(4)$

ψ and $\psi^{(1)}$ are the digamma and trigamma functions respectively. These "corrections" arise from the taking the natural log.

All this is an approximation, but one that can expect will work quite well. In the next section to do a model validation study through a comprehensive simulation study. We study the regression model for auto-regressive, moving-average models (ARMA) and fractional auto-regressive, moving-average models (FARMA).

3. VERIFYING THE REGRESSION MODEL

We utilized the class of ARMA models followed by FARMA models as a mechanism for verification, because this represents a large class of utilized models. For all simulations we generated 100 independent runs of a simulation instance with Gaussian noise, $\varepsilon_t \sim Normal(0, \sigma^2 = 1)$. For ARMA cases, the simulation length was $n = 2^{20} = 1,048,576$ and for FARMA cases, the simulation length was $n = 2^{18} = 262,144$ and $d = 0.45 > 0$.

With these simulations, we performed model diagnostics of our log averaged periodogram versus frequency compared to the log power spectrum curve. Additionally, we checked residuals plots of the difference between the log averaged periodogram and the log power spectrum versus frequency. Finally, Quantile-Quantile (QQ) Normal Plots of the residuals were created to verify their distributional properties.

3.1 ARMA Models

3.1.1 AR(p = 1)

The $AR(p = 1)$ time series model is:

$$x_t = \phi x_{t-1} + \varepsilon_t = 0.95 x_{t-1} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi, \sigma^2)) = \frac{\sigma^2}{|1 - \phi e^{-i2\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi, \sigma^2))) + \varepsilon = \ln(\sigma^2) - \ln(|1 - \phi e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results: [3]

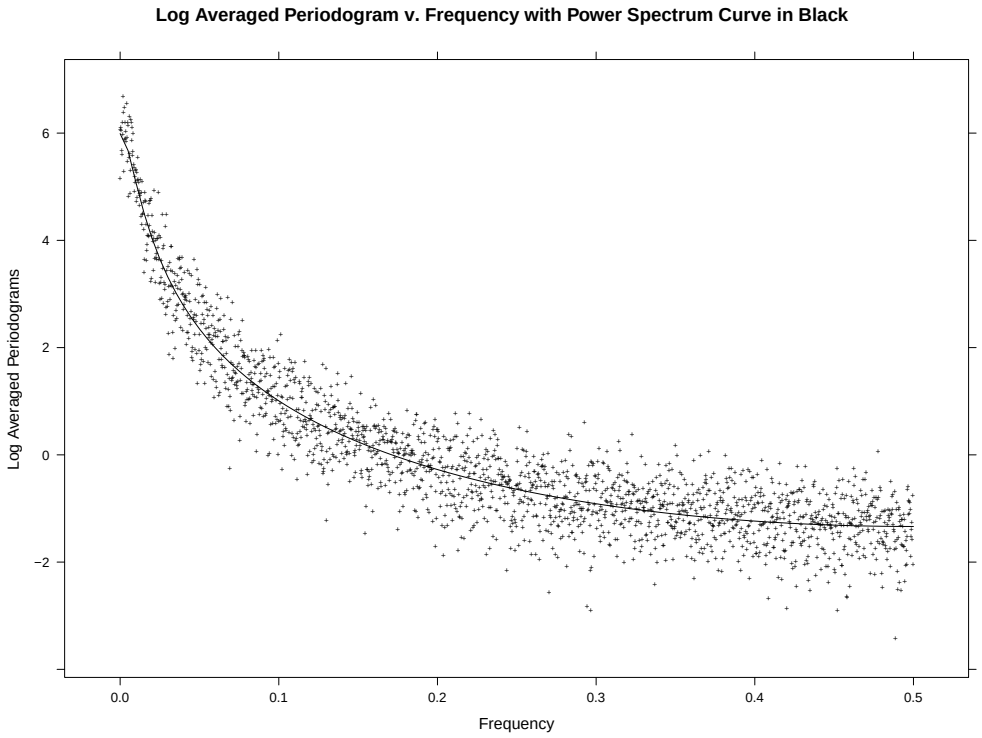

Fig. 3.1. AR(p = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

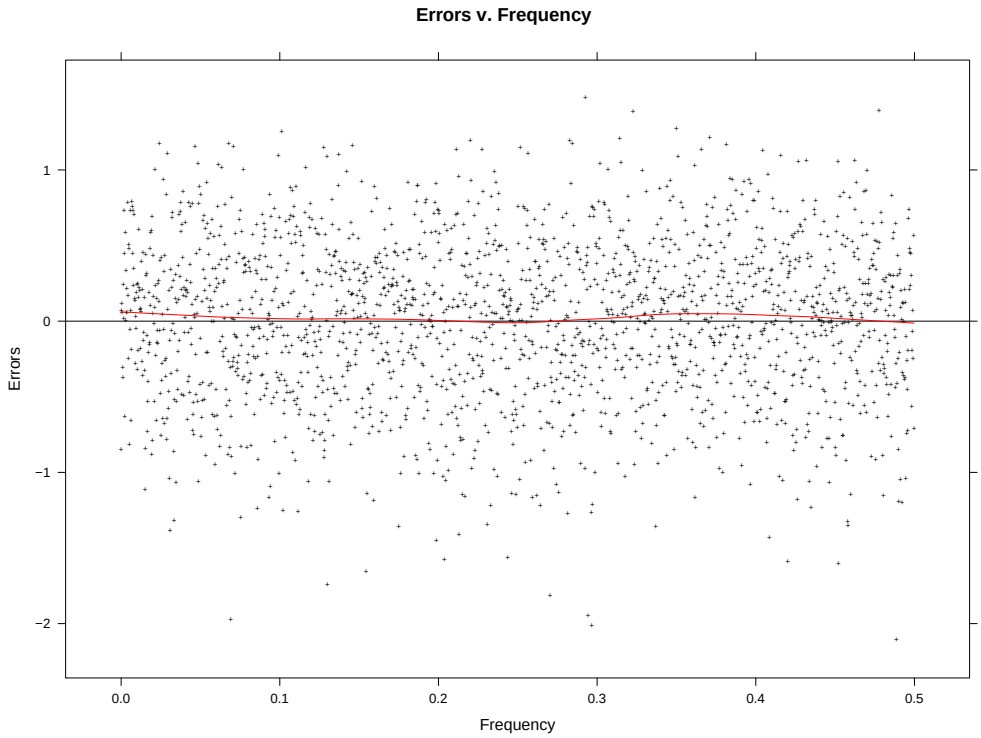

Fig. 3.2. AR(p = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third. [4]

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

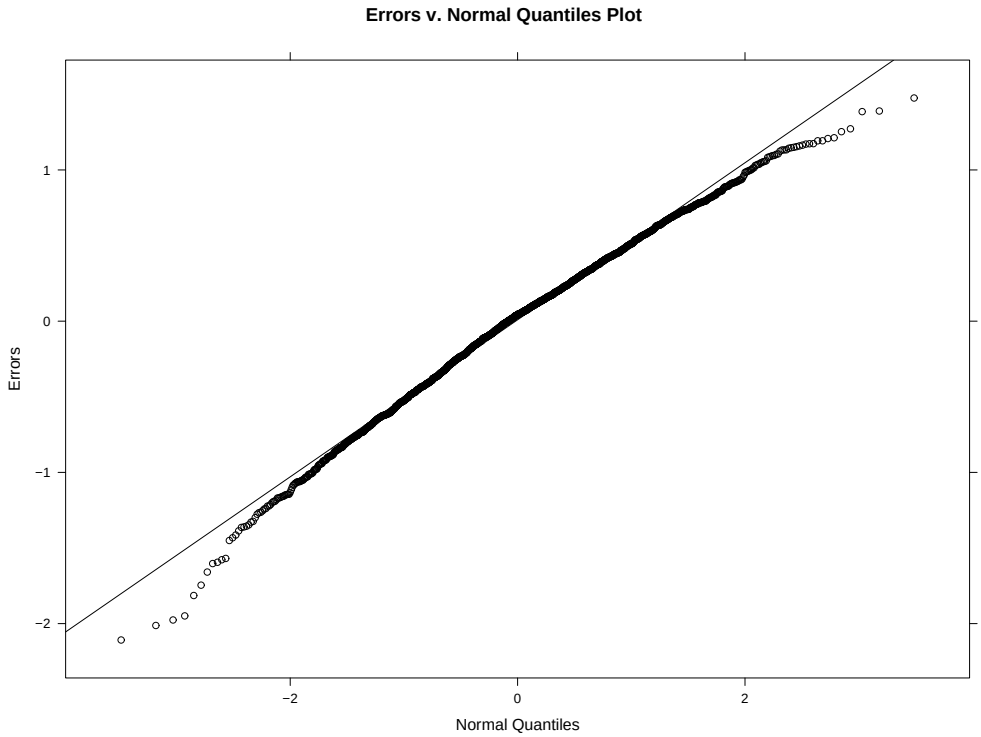

Fig. 3.3. AR(p = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.2 AR(p = 2)

The $AR(p=2)$ time series model is:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t = 1.9x_{t-1} - 0.95x_{t-2} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi_1,\phi_2,\sigma^2)) = \frac{\sigma^2}{|1-\phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi_1,\phi_2,\sigma^2))) + \varepsilon = \ln(\sigma^2) - \ln(|1-\phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

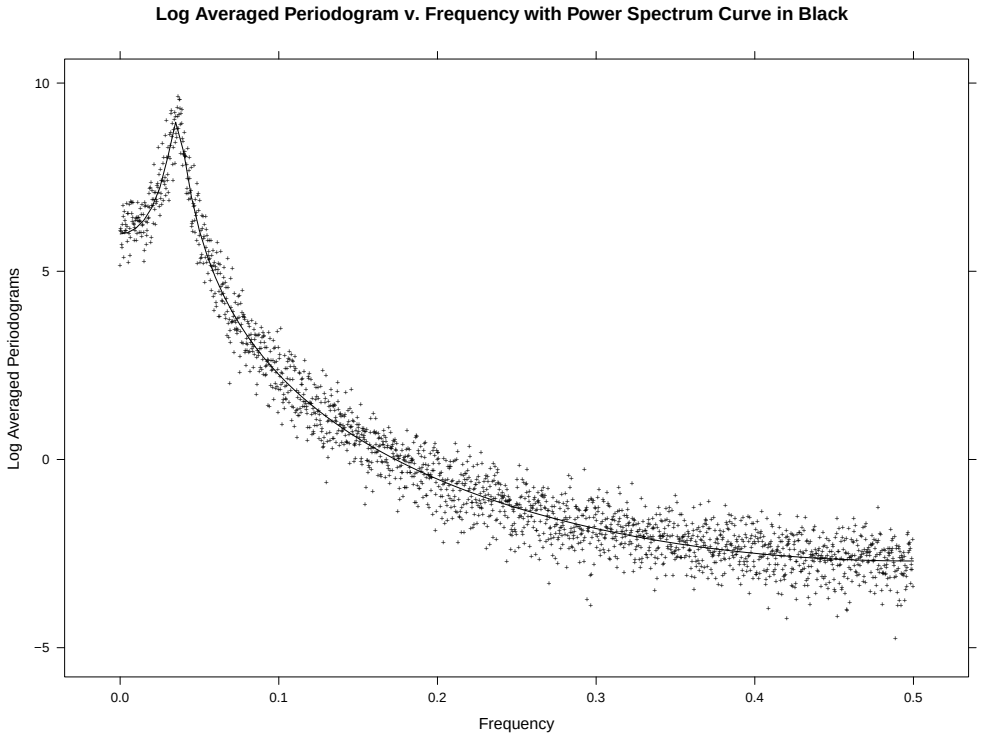

Fig. 3.4. AR(p = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

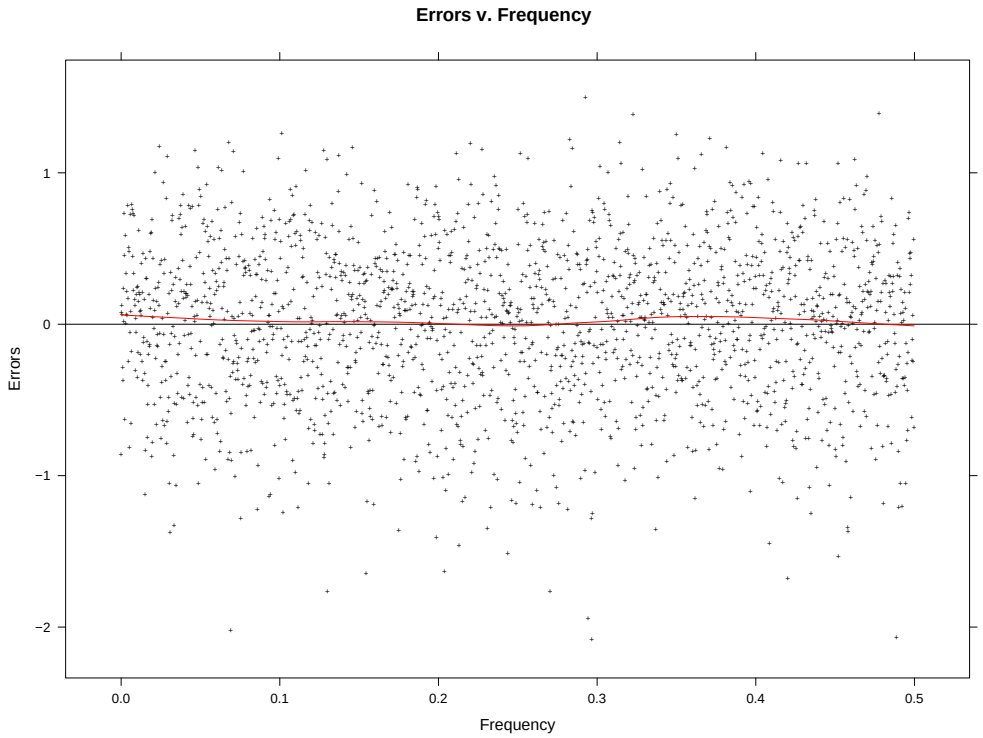

Fig. 3.5. AR(p = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

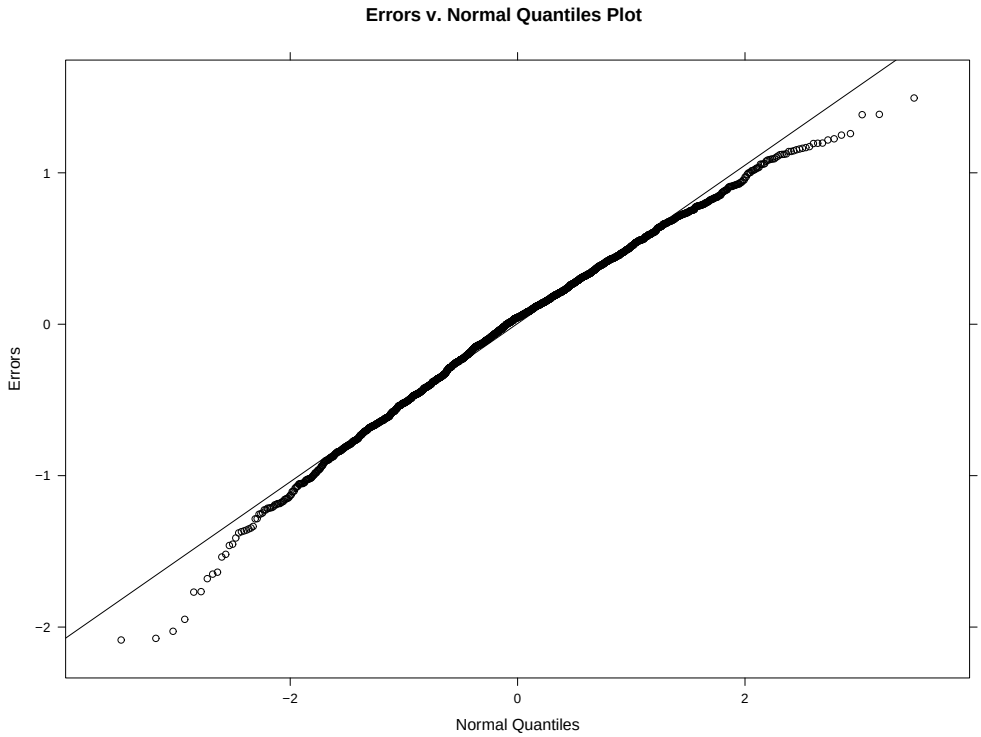

Fig. 3.6. AR(p = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.3 MA(q = 1)

The $MA(q = 1)$ time series model is:

$$x_t = \theta\varepsilon_{t-1} + \varepsilon_t = 0.95\varepsilon_{t-1} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\theta,\sigma^2)) = \sigma^2|1 + \theta e^{-i2\pi f_j}|^2$$

The regression model is:

$$y_j = \ln(S(f_j|(\theta,\sigma^2))) + \varepsilon = \ln(\sigma^2) + \ln(|1+\theta e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

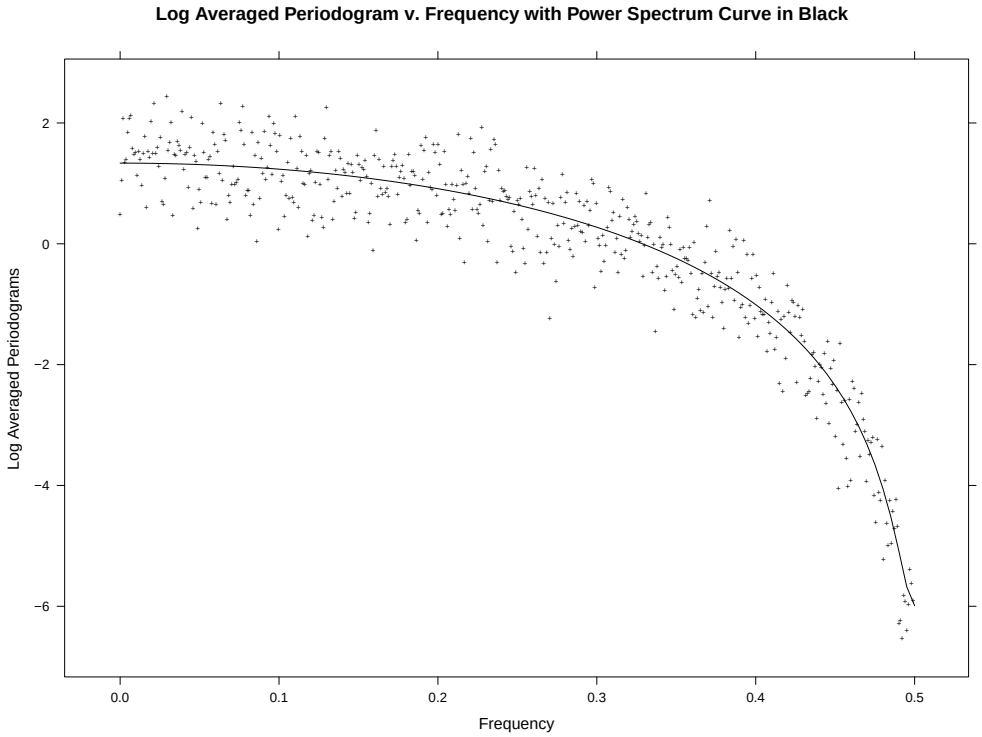

Fig. 3.7. MA(q = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

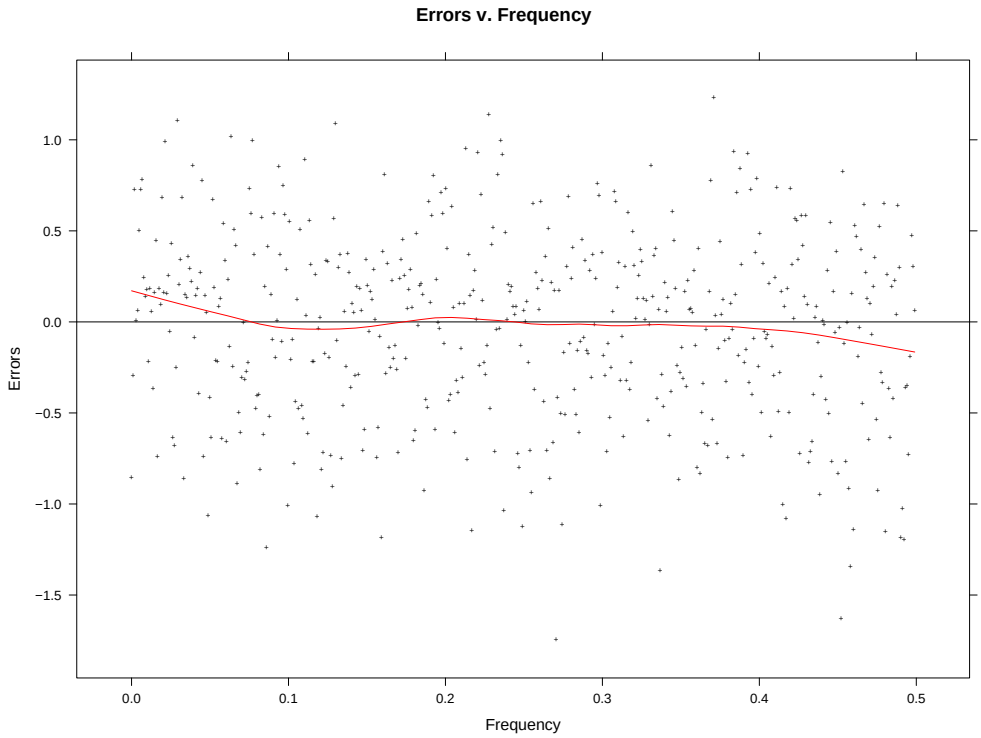

Fig. 3.8. MA(q = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not posses a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

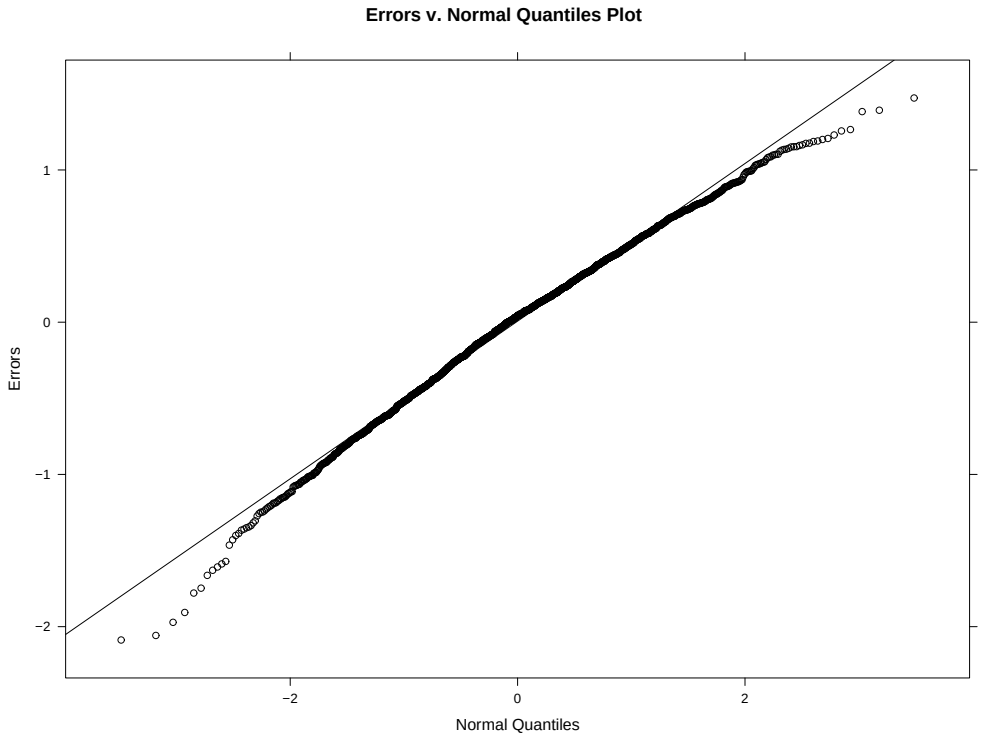

Fig. 3.9. MA(q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.4 MA(q = 2)

The $MA(q=2)$ time series model is:

$$x_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t = 1.9\varepsilon_{t-1} + 0.95\varepsilon_{t-2} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\theta_1, \theta_2, \sigma^2)) = \sigma^2 |1 + \theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2$$

The regression model is:

$$y_j = \ln(S(f_j|(\theta_1, \theta_2, \sigma^2))) + \varepsilon = \ln(\sigma^2) + \ln(|1 + \theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

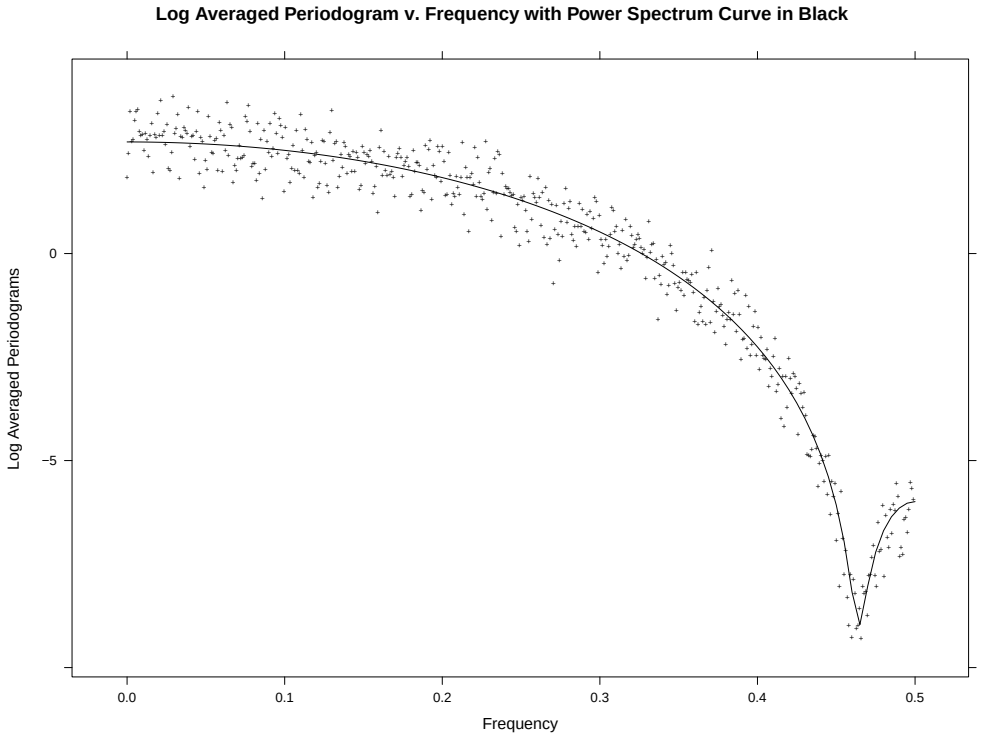

Fig. 3.10. MA(q = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

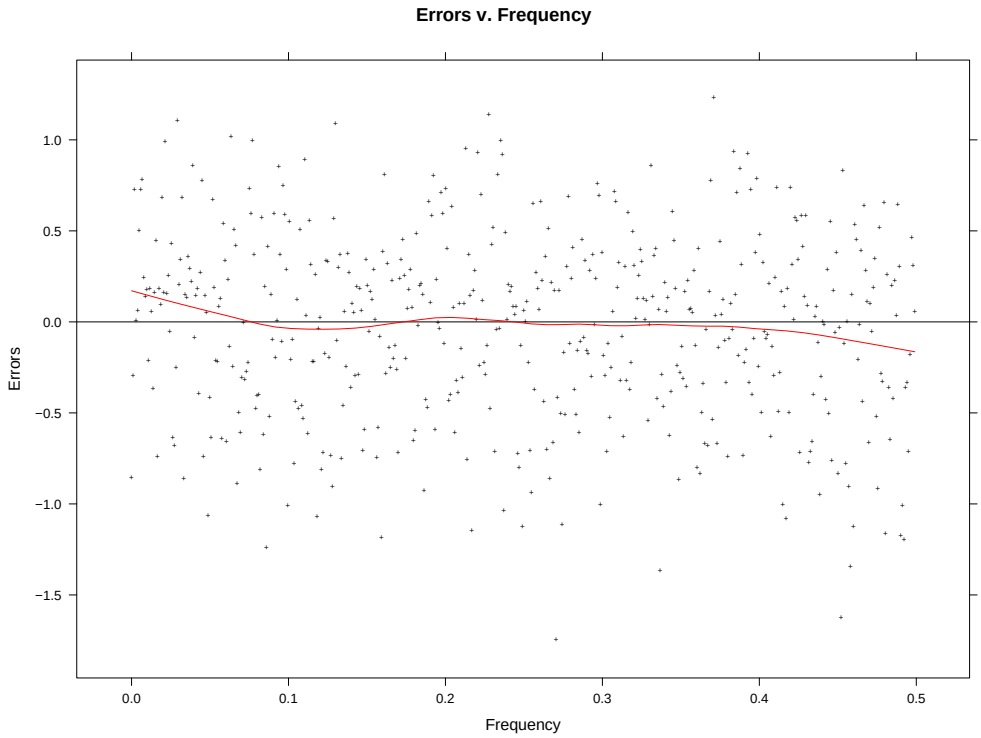

Fig. 3.11. MA(q = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

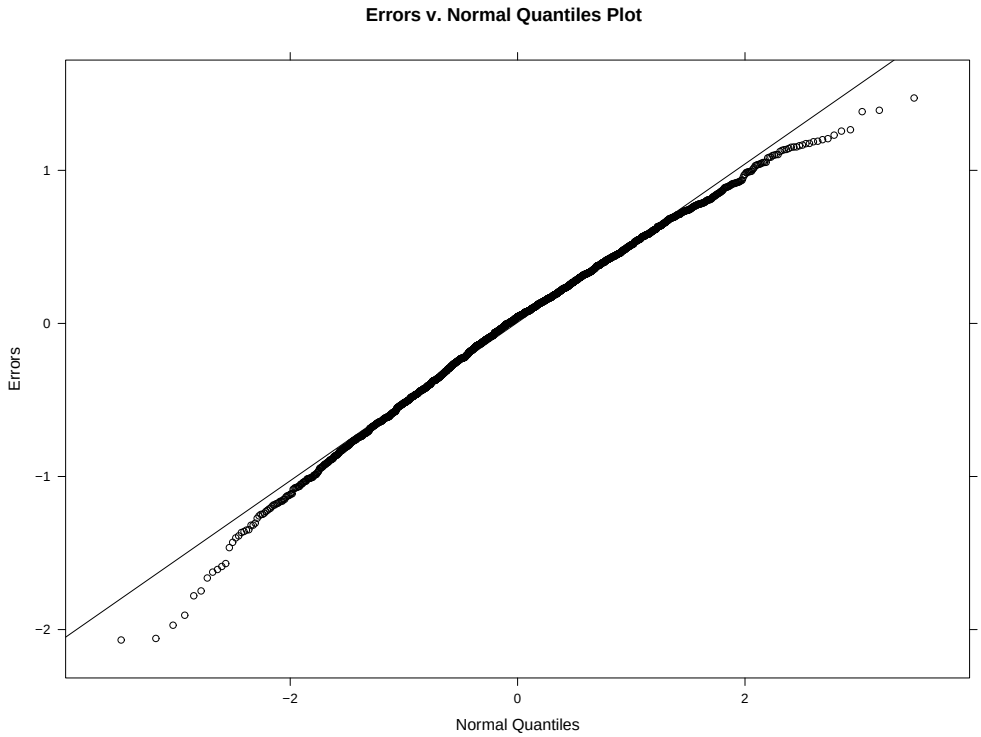

Fig. 3.12. MA(q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.5 ARMA(p = 1, q = 1)

The $ARMA(p = 1, q = 1)$ time series model is:

$$x_t = \phi x_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi, \theta, \sigma^2)) = \sigma^2 \frac{|1 + \theta e^{-i2\pi f_j}|^2}{|1 - \phi e^{-i2\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi,\theta,\sigma^2))) + \varepsilon = \ln(\sigma^2) + \ln(|1+\theta e^{-i2\pi f_j}|^2) - \ln(|1-\phi e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

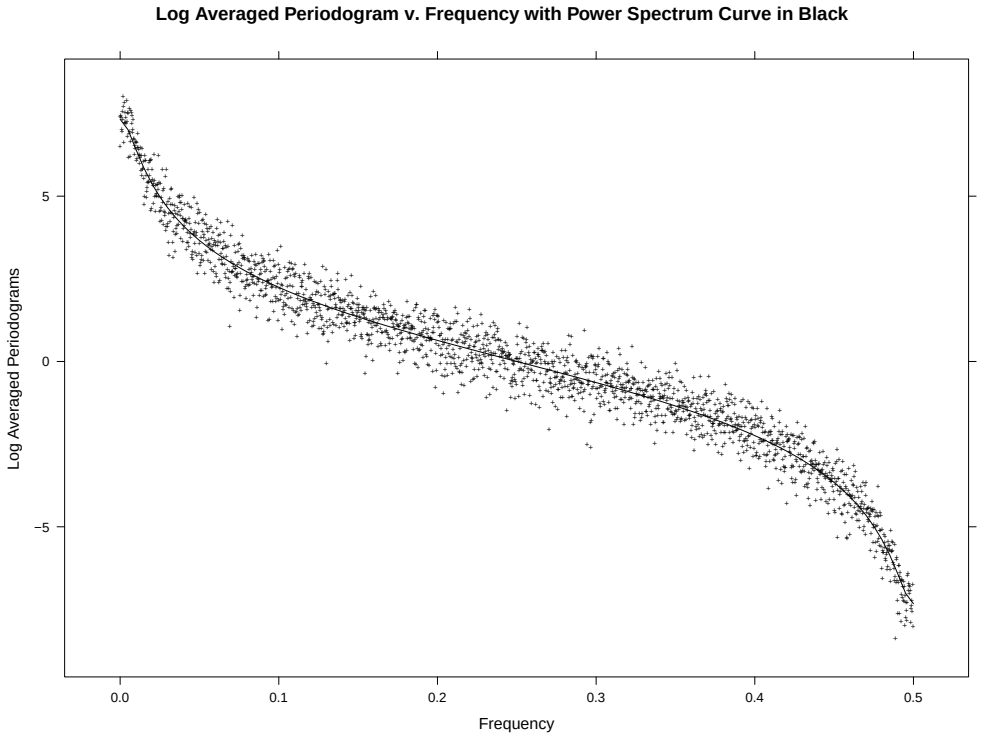

Fig. 3.13. ARMA(p = 1, q = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

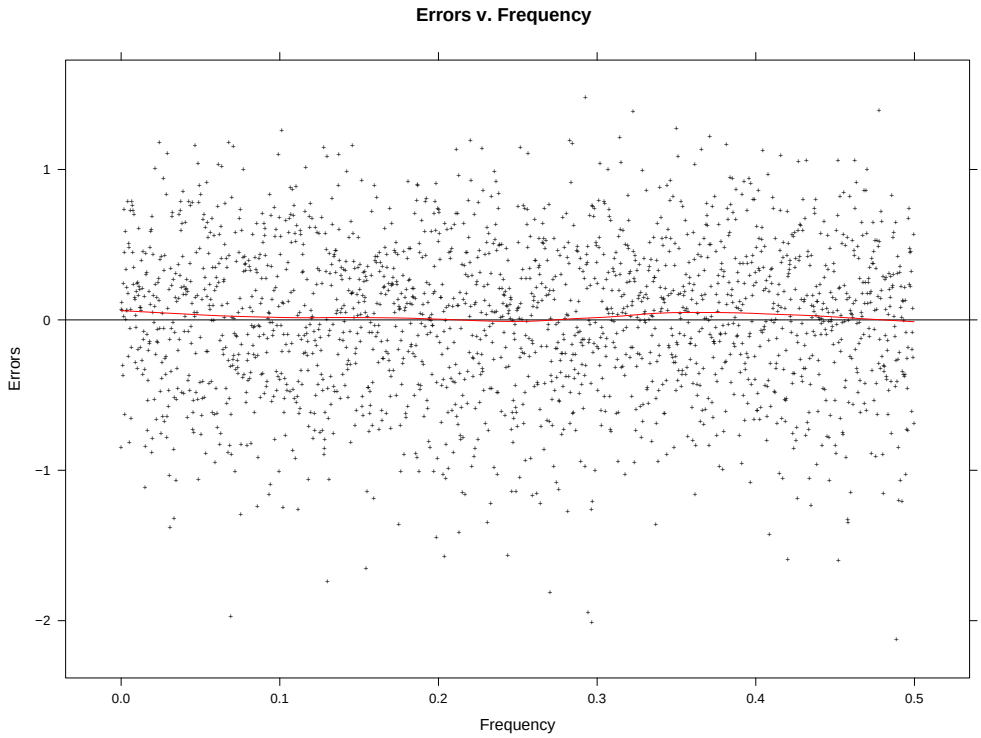

Fig. 3.14. ARMA(p = 1, q = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

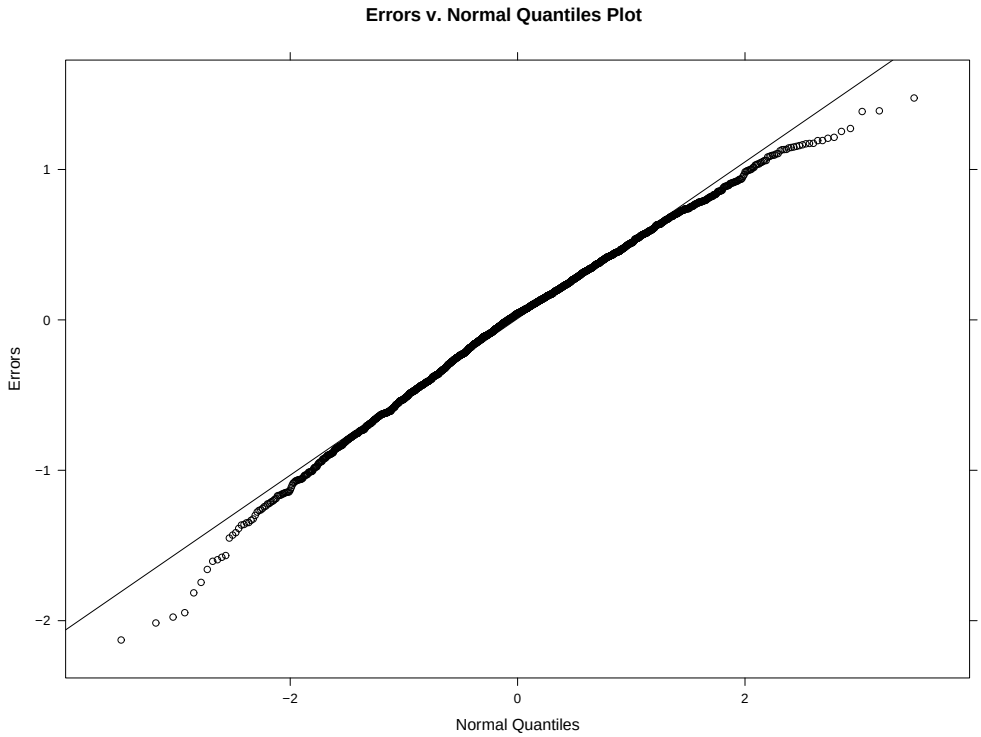

Fig. 3.15. ARMA(p = 1, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.6 ARMA(p = 2, q = 1)

The $ARMA(p=2,q=1)$ time series model is:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \theta\varepsilon_{t-1} + \varepsilon_t = 1.8x_{t-1} - 0.9x_{t-2} + 0.9\varepsilon_{t-1} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi_1,\phi_2,\theta,\sigma^2)) = \sigma^2 \frac{|1+\theta e^{-i2\pi f_j}|^2}{|1-\phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi_1, \phi_2, \theta, \sigma^2))) + \varepsilon$$

$$= \ln(\sigma^2) + \ln(|1 + \theta e^{-i2\pi f_j}|^2) - \ln(|1 - \phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

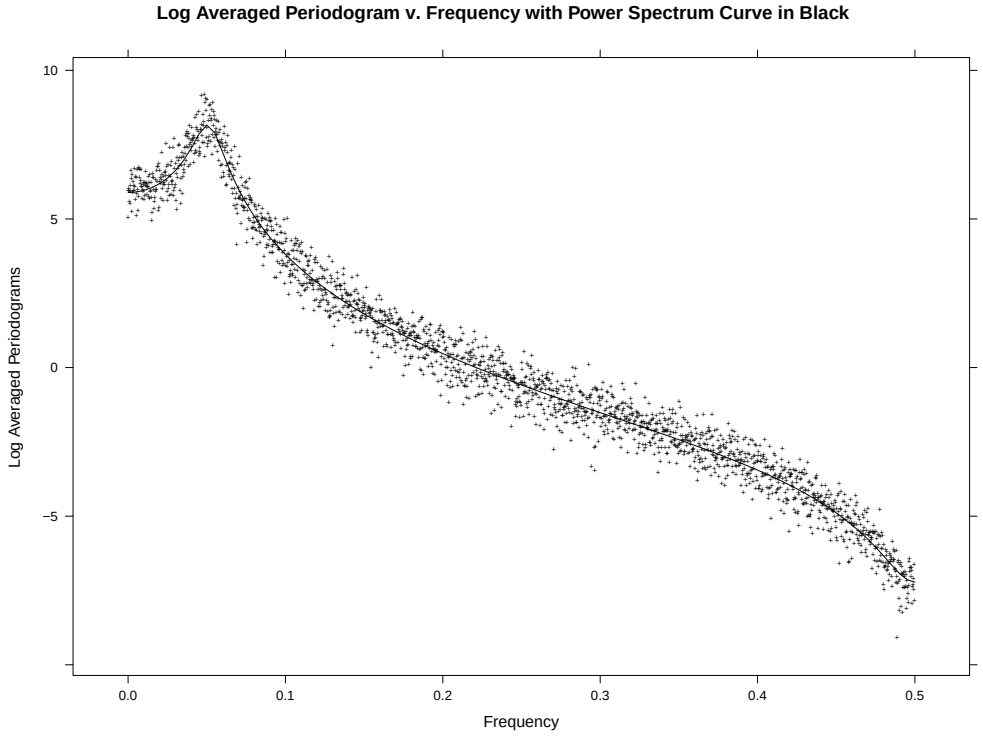

Fig. 3.16. ARMA(p = 2, q = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any

potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

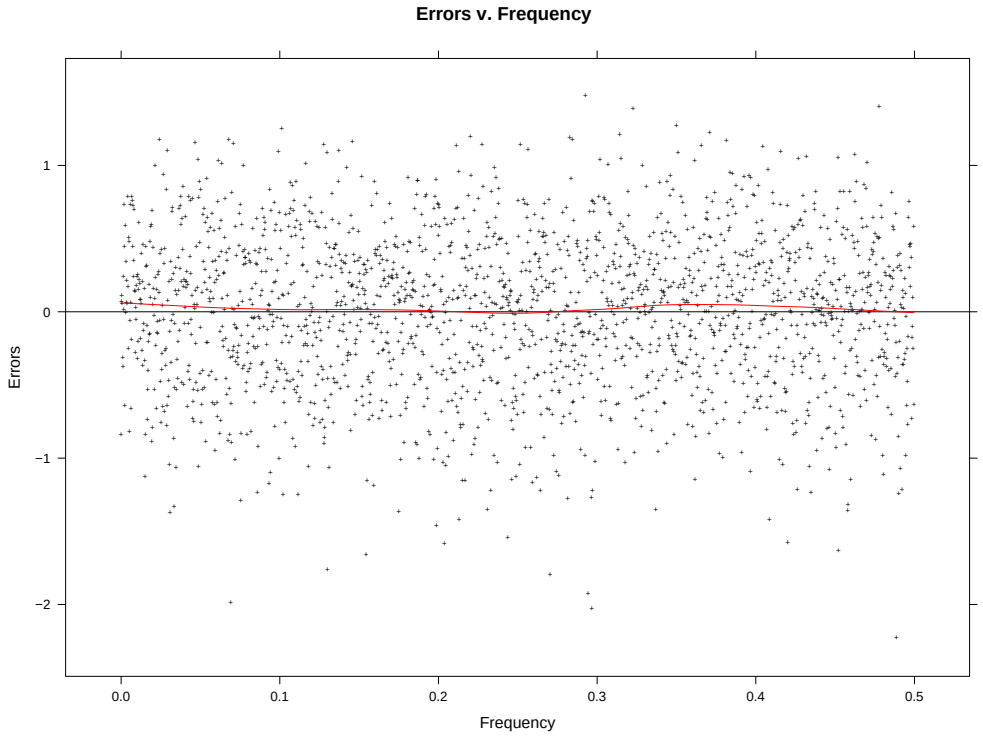

Fig. 3.17. ARMA(p = 2, q = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

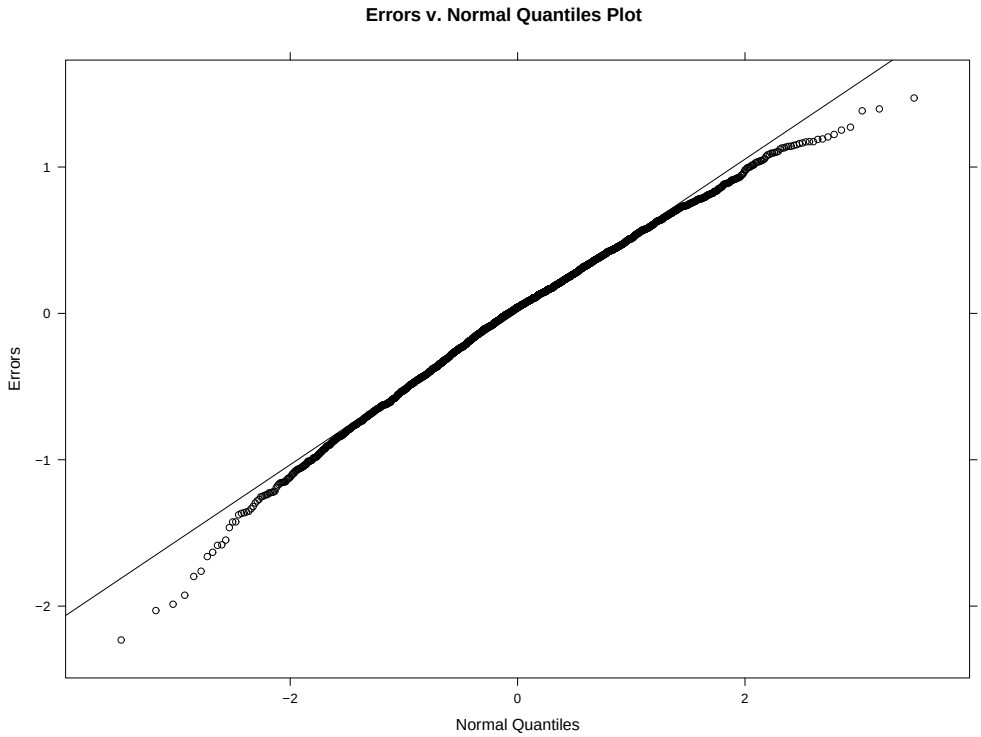

Fig. 3.18. ARMA(p = 2, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.7 ARMA(p = 1, q = 2)

The $ARMA(p=1,q=2)$ time series model is:

$$x_t = \phi x_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t = 0.9x_{t-1} + 1.8\varepsilon_{t-1} + 0.9\varepsilon_{t-2} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi,\theta_1,\theta_2,\sigma^2)) = \sigma^2 \frac{|1+\theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2}{|1-\phi e^{-i2\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi, \theta_1, \theta_2, \sigma^2))) + \varepsilon$$

$$= \ln(\sigma^2) + \ln(|1 + \theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2) - \ln(|1 - \phi e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram results:

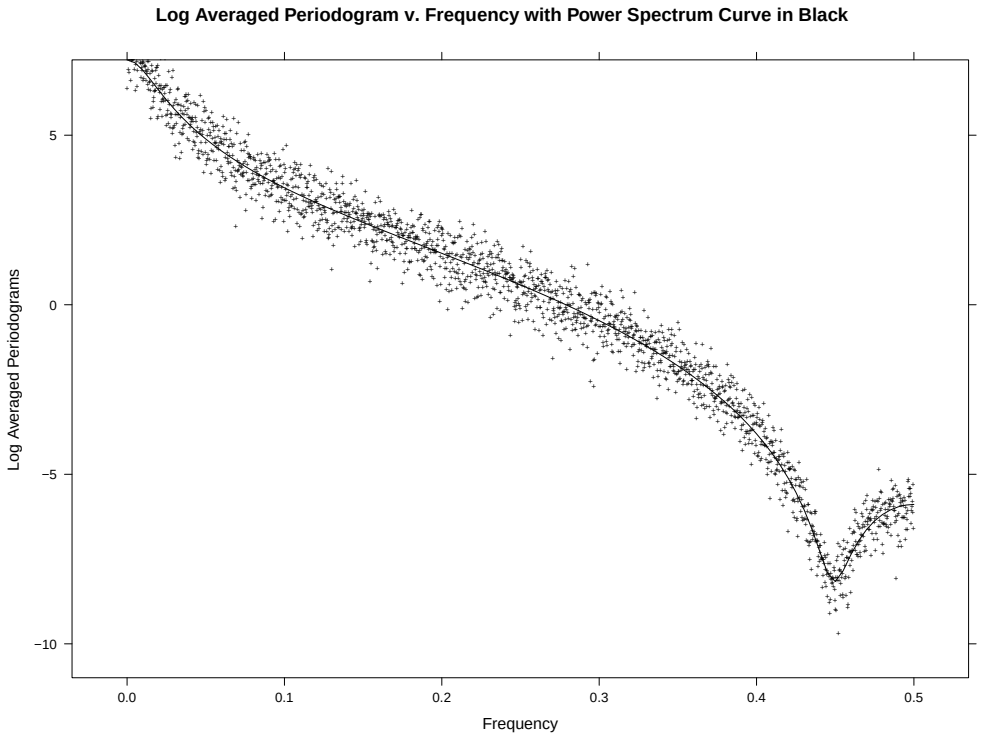

Fig. 3.19. ARMA(p = 1, q = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

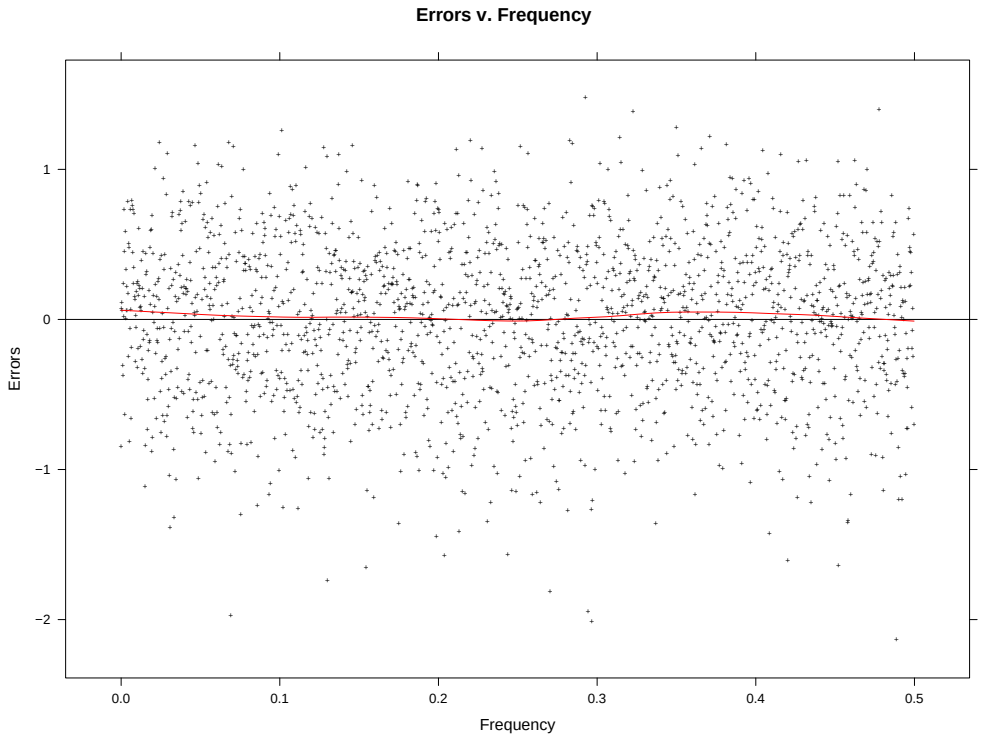

Fig. 3.20. ARMA(p = 1, q = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

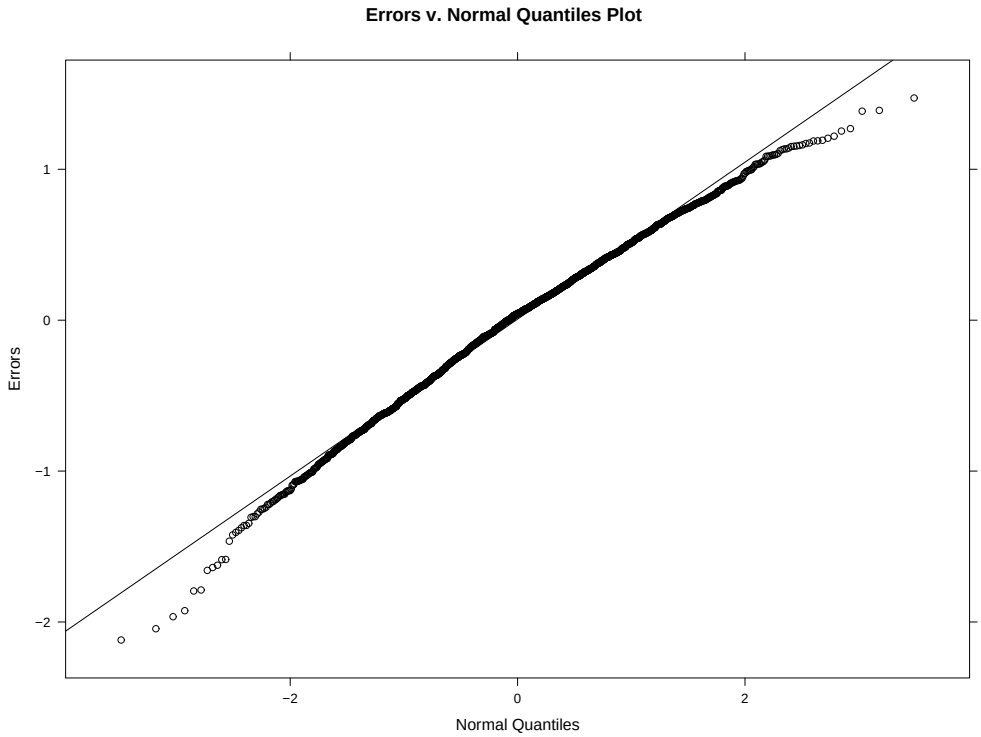

Fig. 3.21. ARMA(p = 1, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.1.8 ARMA(p = 2, q = 2)

The $ARMA(p=2,q=2)$ time series model is:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t = 1.8x_{t-1} - 0.9x_{t-2} + 1.8\varepsilon_{t-1} + 0.9\varepsilon_{t-2} + \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi_1,\phi_2,\theta_1,\theta_2,\sigma^2)) = \sigma^2 \frac{|1+\theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2}{|1-\phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi_1, \phi_2, \theta_1, \theta_2, \sigma^2))) + \varepsilon$$

$$= \ln(\sigma^2) + \ln(|1 + \theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2) - \ln(|1 - \phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2) + \varepsilon$$

The simulation yielded the following log averaged periodogram versus frequency results:

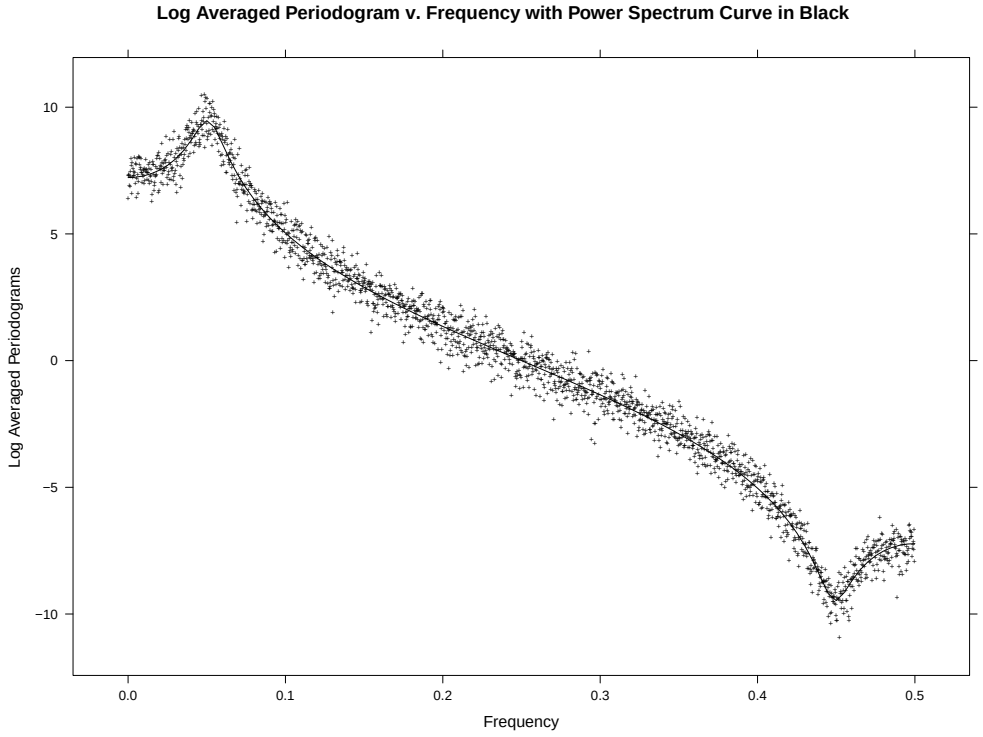

Fig. 3.22. ARMA(p = 2, q = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any

potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

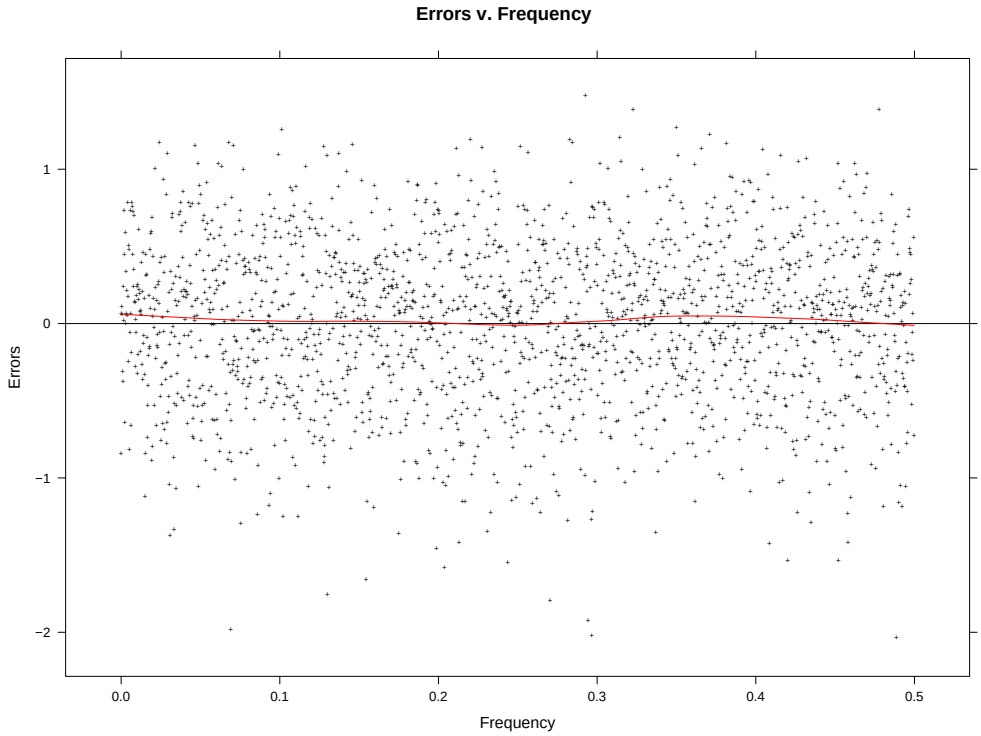

Fig. 3.23. ARMA(p = 2, q = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

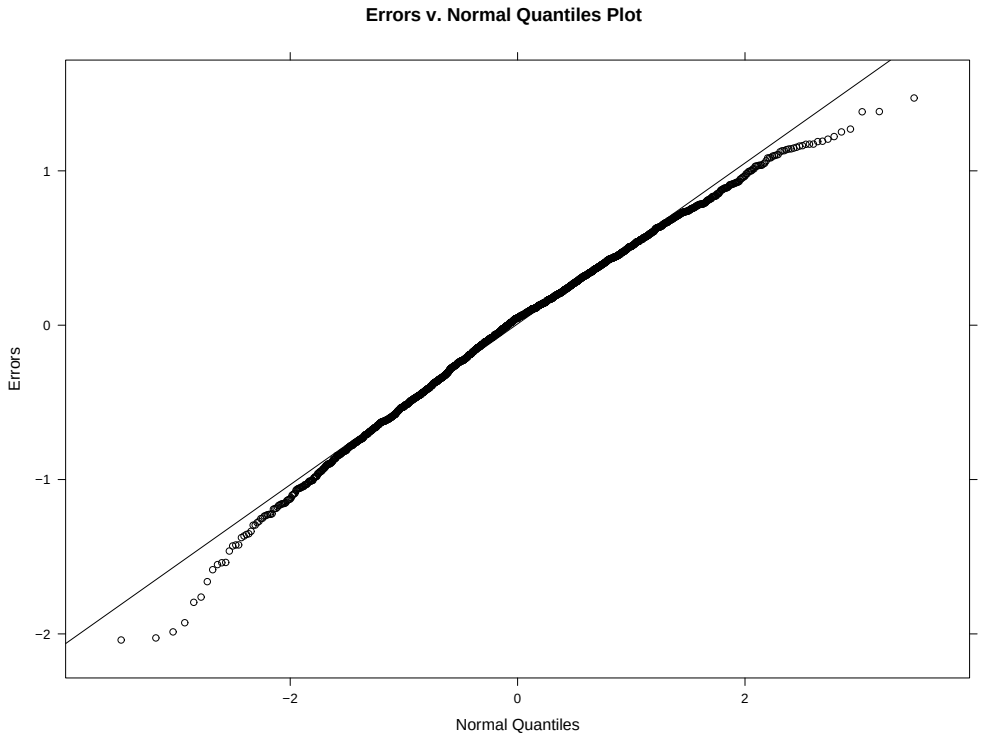

Fig. 3.24. ARMA(p = 2, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2 FARMA Models

3.2.1 FARMA(p = 0, d, q = 0)

The $FARMA(p = 0, d, q = 0)$ time series model is:

$$(1-B)^d x_t = (1-B)^{0.45} = \varepsilon_t$$

The power spectrum is: [5]

$$S(f_j|(d,\sigma^2)) = \sigma^2|1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$y_j = \ln(S(f_j|(d,\sigma^2))) + \varepsilon = \ln(\sigma^2) - d\cdot\ln(|1-e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

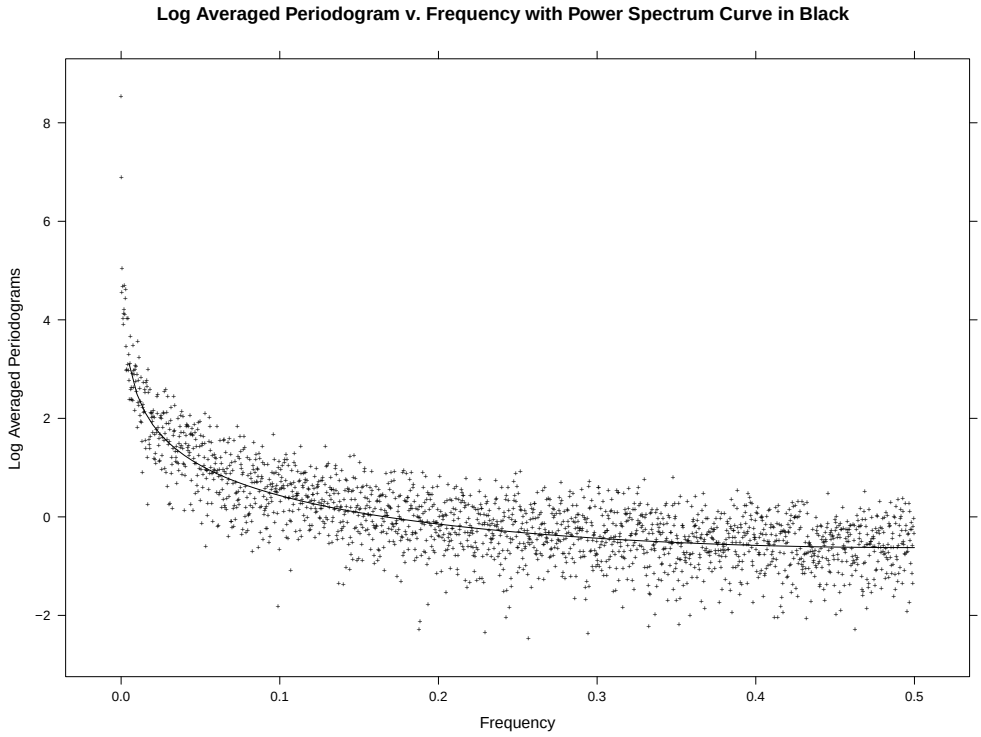

Fig. 3.25. ARFIMA(p = 0, d, q = 0) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any

potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

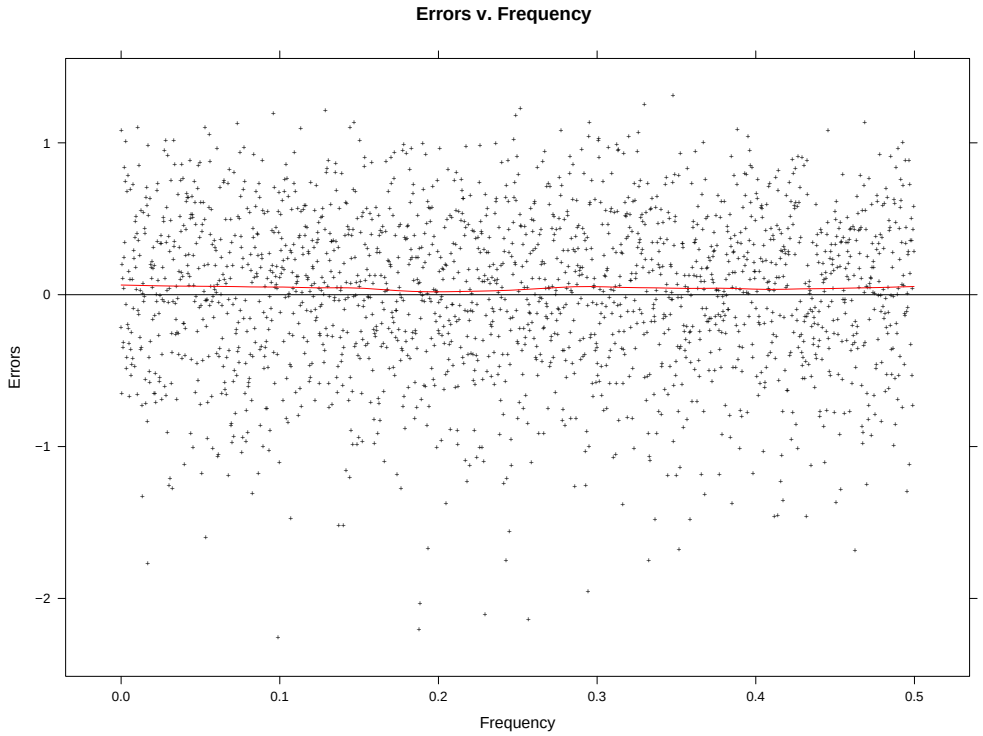

Fig. 3.26. ARFIMA(p = 0, d, q = 0) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

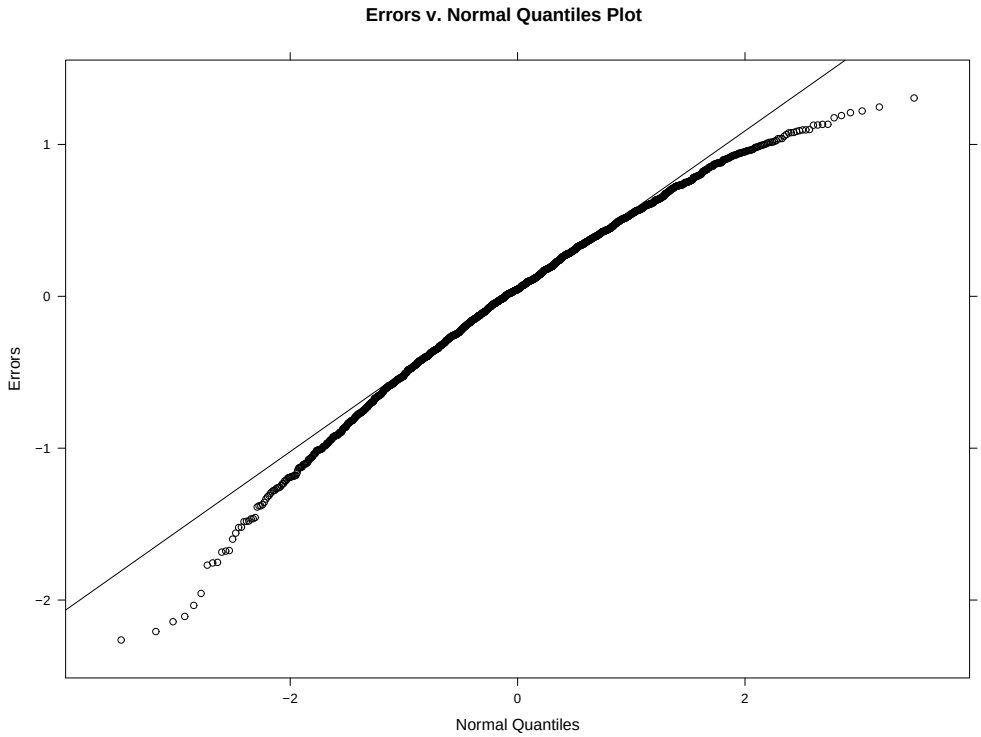

Fig. 3.27. ARFIMA(p = 0, d, q = 0) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.2 FARMA(p = 1, d, q = 0)

The $FARMA(p = 1, d, q = 0)$ time series model is:

$$(1 - \phi B)(1 - B)^d x_t = (1 - 0.95B)(1 - B)^{0.45} x_t = \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi, d, \sigma^2)) = \sigma^2 \frac{|1 - e^{-i2\pi f_j}|^{-2d}}{|1 - \phi e^{-i2\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi, d, \sigma^2))) + \varepsilon = \ln(\sigma^2) - d \cdot \ln(|1 - e^{-i2\pi f_j}|^2) - \ln(|1 - \phi e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

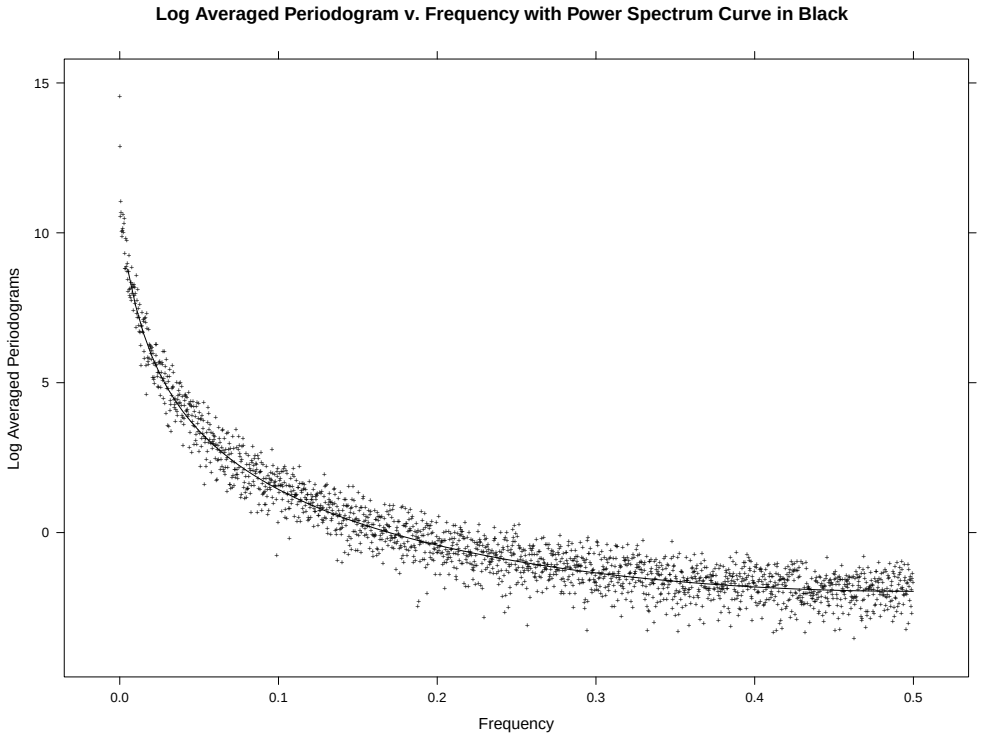

Fig. 3.28. ARFIMA(p = 1, d, q = 0) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum as these plots demonstrate. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

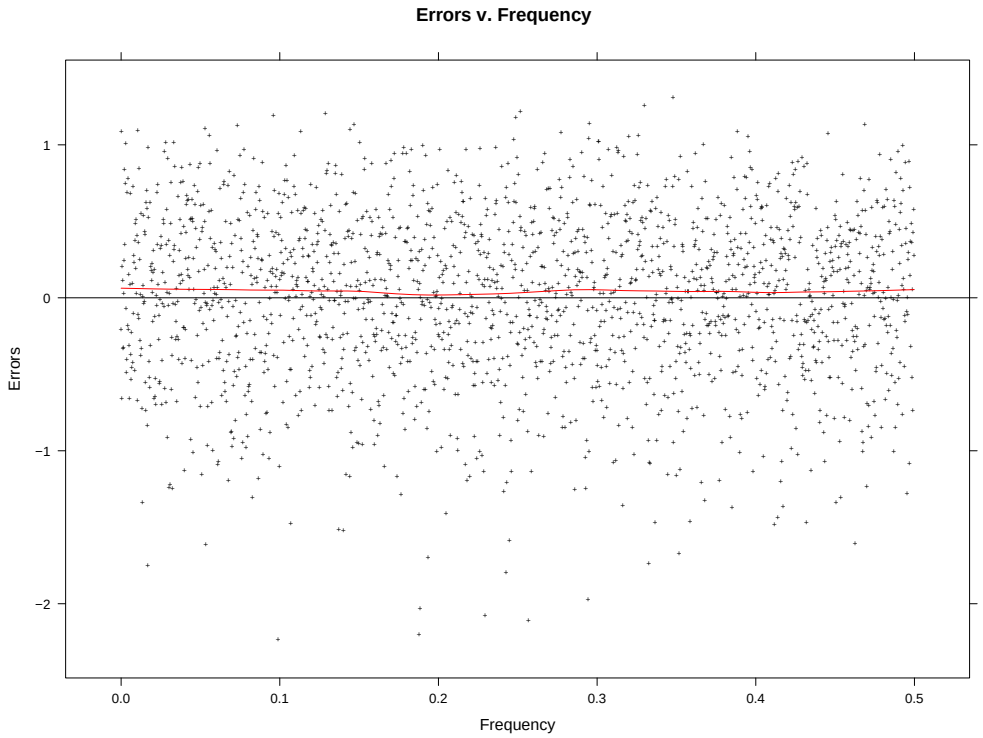

Fig. 3.29. ARFIMA(p = 1, d, q = 0) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

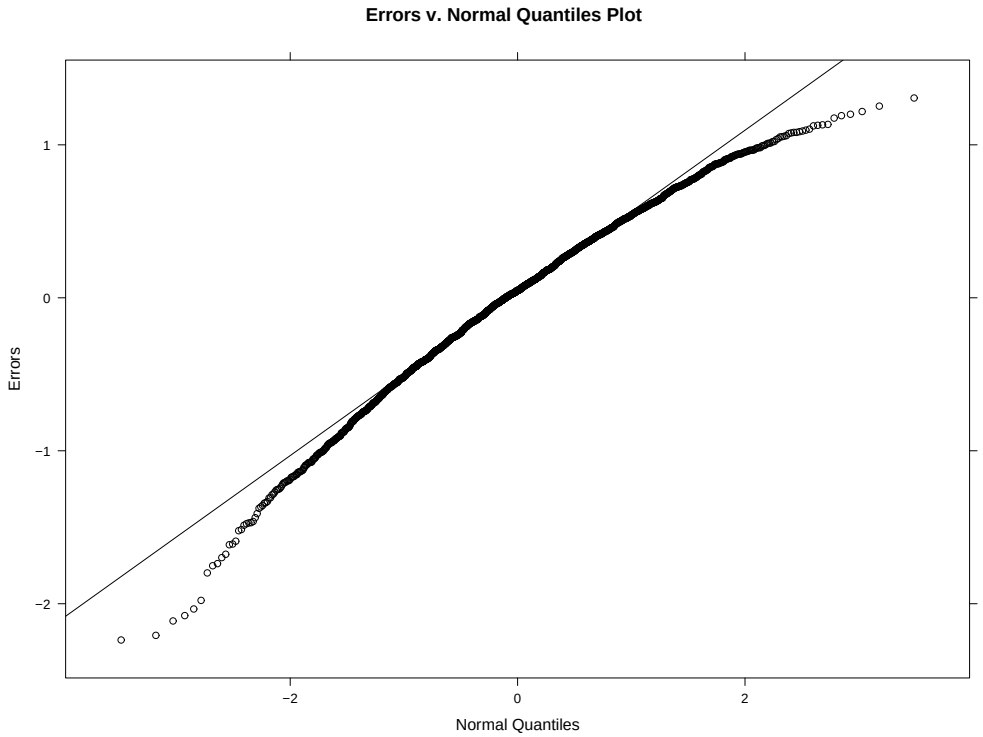

Fig. 3.30. ARFIMA(p = 1, d, q = 0) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.3 FARMA(p = 2, d, q = 0)

The $FARMA(p=2,d,q=0)$ time series model is:

$$(1-\phi_1 B-\phi_2 B^2)(1-B)^d x_t = (1-1.9B+0.95B^2)(1-B)^{0.45} x_t = \varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi_1,\phi_2,d,\sigma^2)) = \sigma^2 \frac{|1-e^{-i2\pi f_j}|^{-2d}}{|1-\phi_1 e^{-i2\pi f_j}-\phi_2 e^{-i2\pi f_j}|^2}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi_1, \phi_2, d, \sigma^2))) + \varepsilon$$

$$= \ln(\sigma^2) - d \cdot \ln(|1 - e^{-i2\pi f_j}|^2) - \ln(|1 - \phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

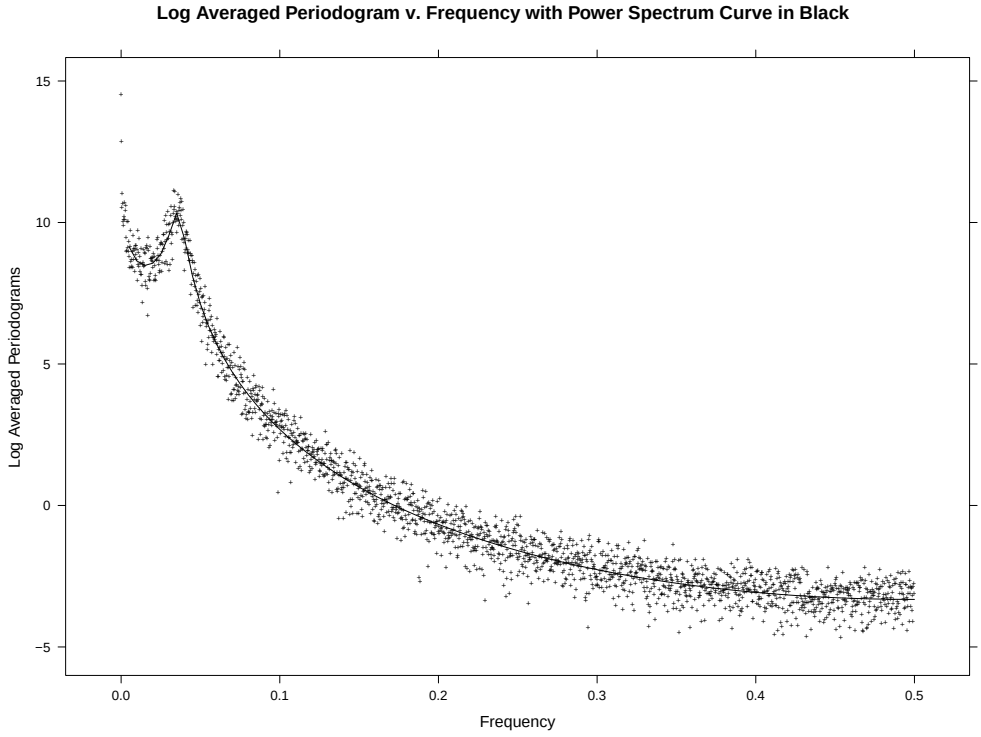

Fig. 3.31. ARFIMA(p = 2, d, q = 0) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any

potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

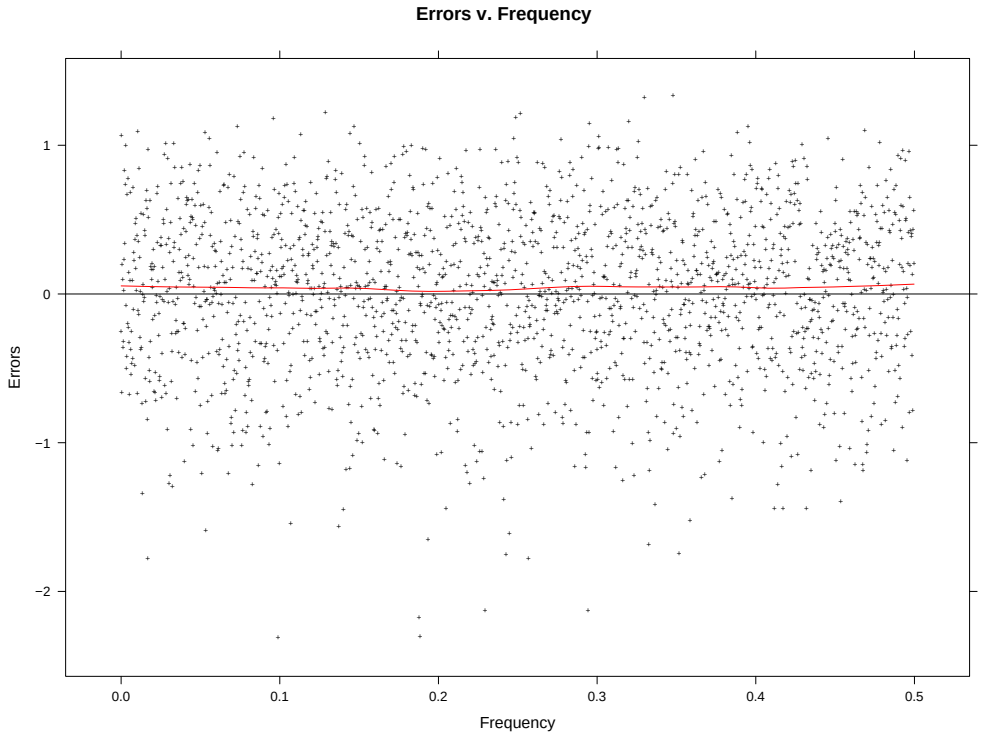

Fig. 3.32. ARFIMA(p = 2, d, q = 0) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

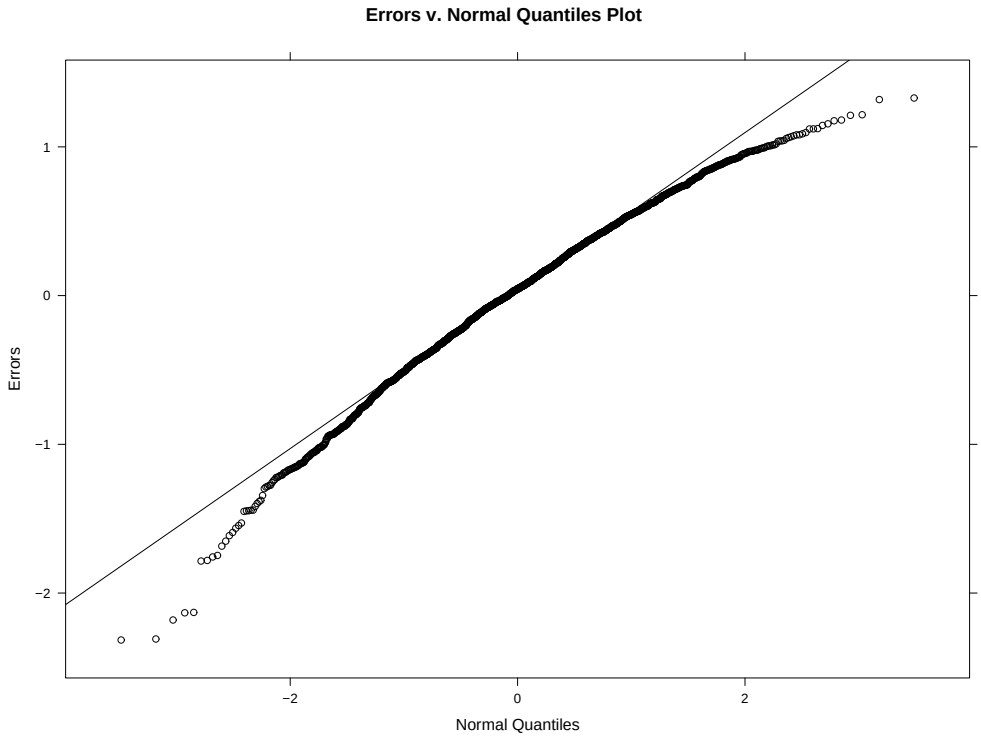

Fig. 3.33. ARFIMA(p = 2, d, q = 0) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.4 FARMA(p = 0, d, q = 1)

The $FARMA(p=0,d,q=1)$ time series model is:

$$(1-B)^d x_t = (1-B)^{0.45} = 0.95\varepsilon_{t-1} + \varepsilon_t = (1+0.95B)\varepsilon_t = (1+\theta B)\varepsilon_t$$

The power spectrum is:

$$S(f_j|(d,\theta,\sigma^2)) = \sigma^2|1+\theta e^{-i2\pi f_j}|^2|1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$y_j = \ln(S(f_j|(d,\theta,\sigma^2))) + \varepsilon = \ln(\sigma^2) + \ln(|1+\theta e^{-i2\pi f_j}|^2) - d \cdot ln(|1-e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

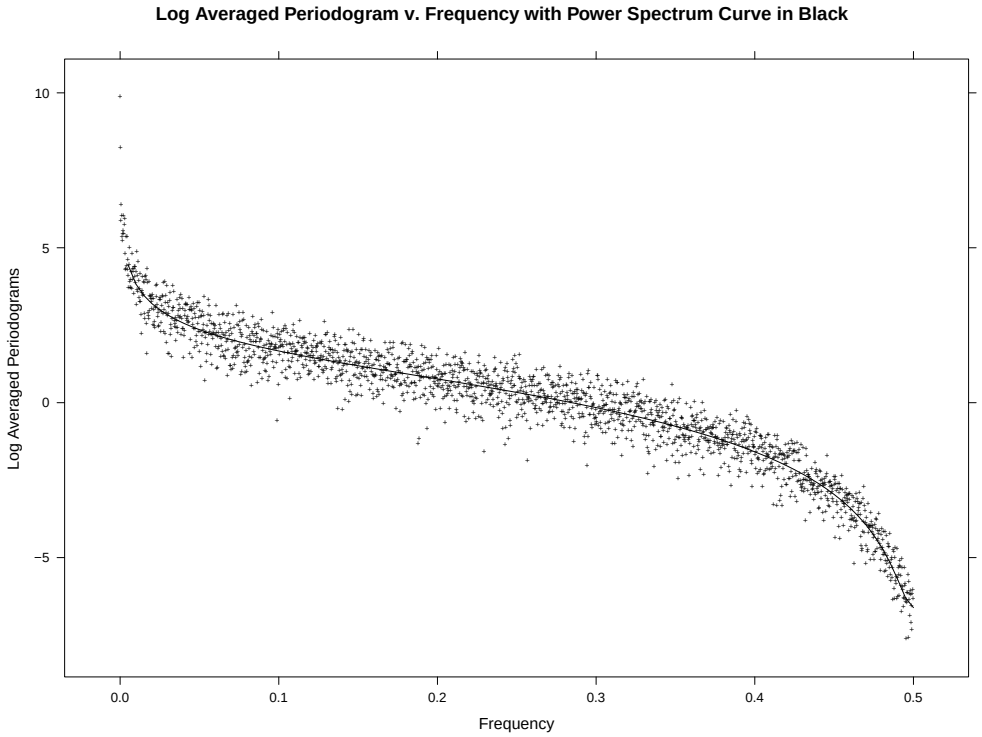

Fig. 3.34. ARFIMA(p = 0, d, q = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

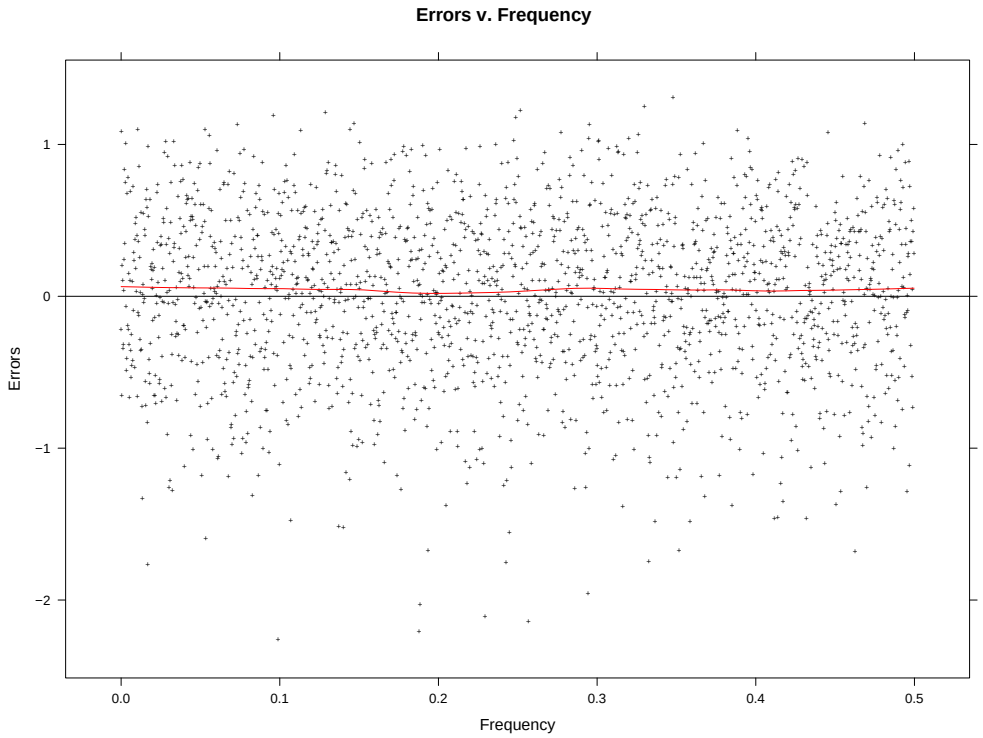

Fig. 3.35. ARFIMA(p = 0, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

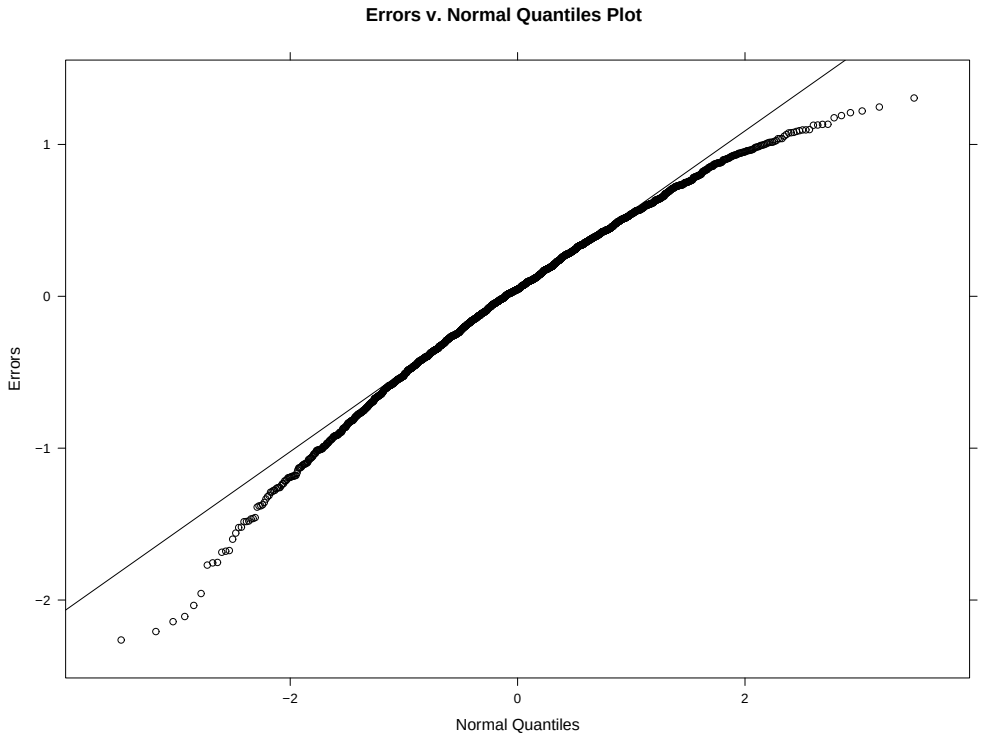

Fig. 3.36. ARFIMA(p = 0, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.5 FARMA(p = 0, d, q = 2)

The $FARMA(p=0,d,q=2)$ time series model is:

$$(1-B)^d x_t = (1-B)^{0.45} x_t = (1+1.9B+0.95B^2)\varepsilon_t = (1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The power spectrum is:

$$S(f_j|(d,\theta_1,\theta_2\sigma^2)) = \sigma^2|1+\theta_1 e^{-i2\pi f_j}+\theta_2 e^{-i4\pi f_j}|^2|1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$y_j = \ln(S(f_j|(d,\theta_1,\theta_2,\sigma^2))) + \varepsilon$$

$$= \ln(\sigma^2) + \ln(|1+\theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2) - d \cdot \ln(|1-e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

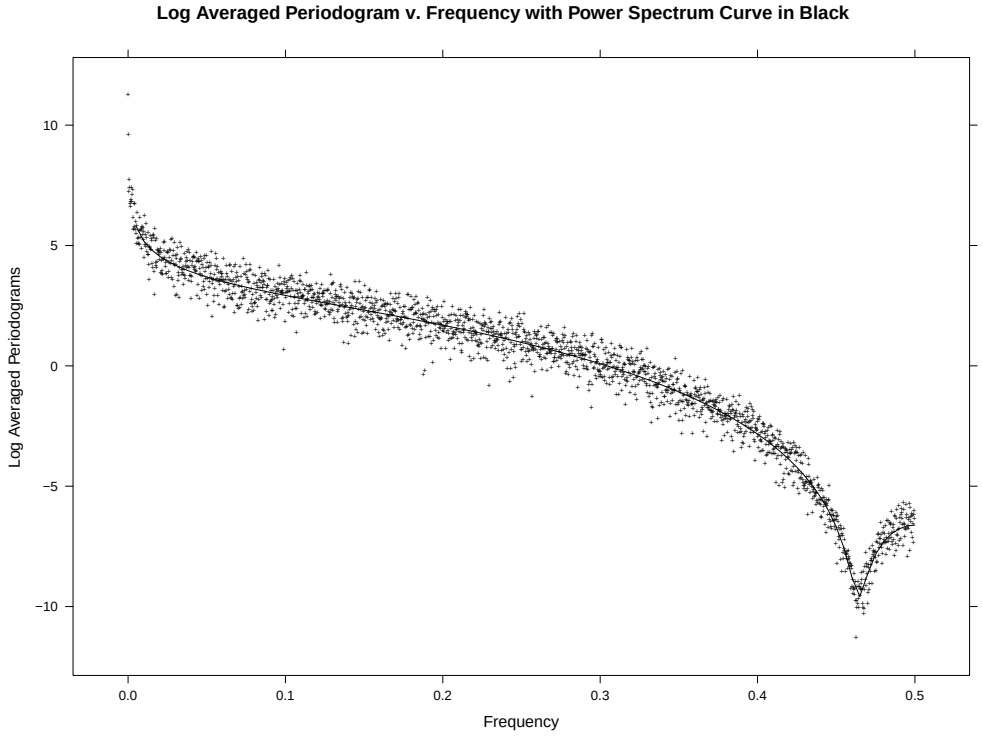

Fig. 3.37. ARFIMA(p = 0, d, q = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any

potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

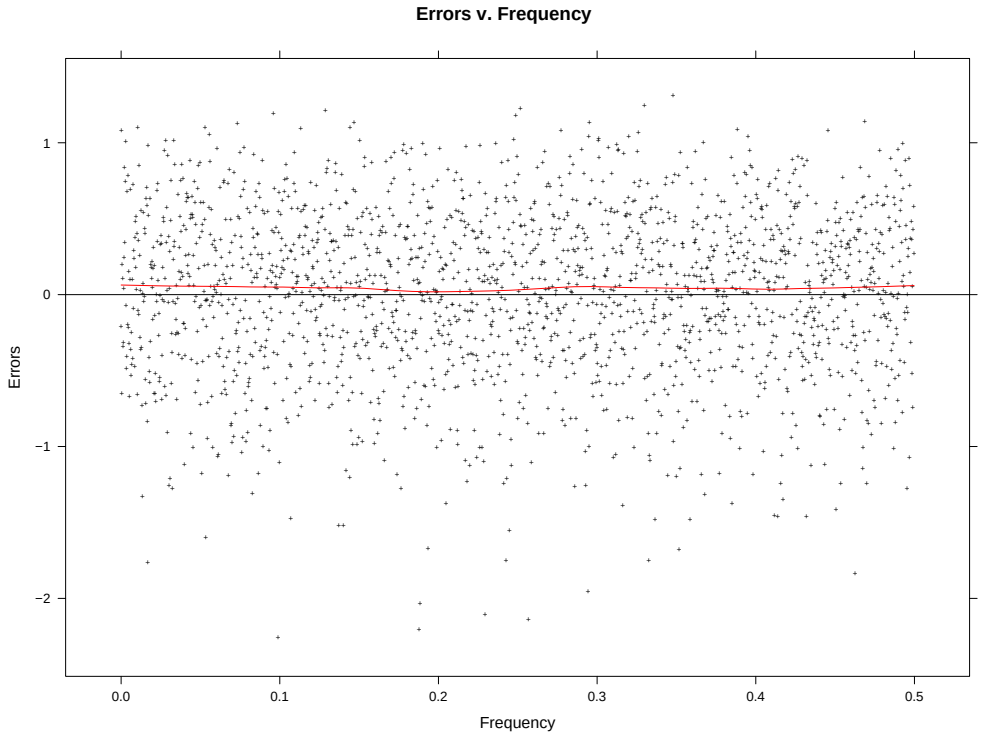

Fig. 3.38. ARFIMA(p = 0, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

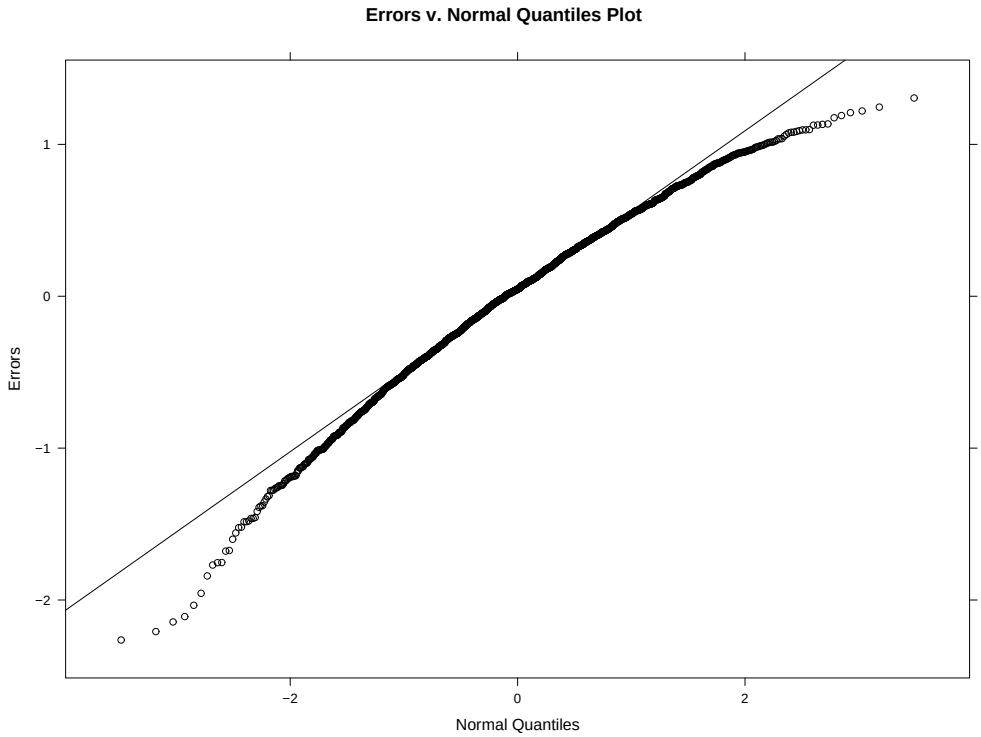

Fig. 3.39. ARFIMA(p = 0, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.6 FARMA(p = 1, d, q = 1)

The $FARMA(p=1,d,q=1)$ time series model is:

$$(1-\phi B)(1-B)^d x_t = (1-0.95B)(1-B)^{0.45} x_t = (1+0.95B)\varepsilon_t = (1+\theta B)\varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi,d,\theta,\sigma^2)) = \sigma^2 \frac{|1+\theta e^{-i2\pi f_j}|^2}{|1-\phi e^{-i2\pi f_j}|^2} |1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi, d, \theta, \sigma^2))) + \varepsilon$$

$$= \ln(\sigma^2) + \ln(|1 + \theta e^{-i2\pi f_j}|^2) - d \cdot \ln(|1 - e^{-i2\pi f_j}|^2) - \ln(|1 - \phi e^{-i2\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

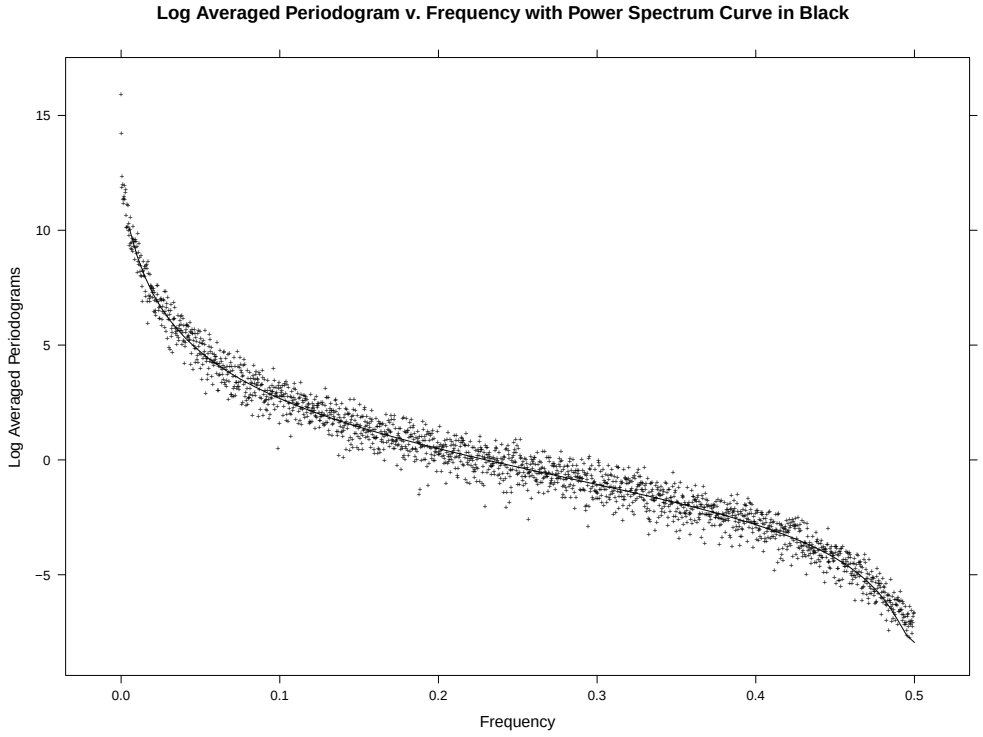

Fig. 3.40. ARFIMA(p = 1, d, q = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any

potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

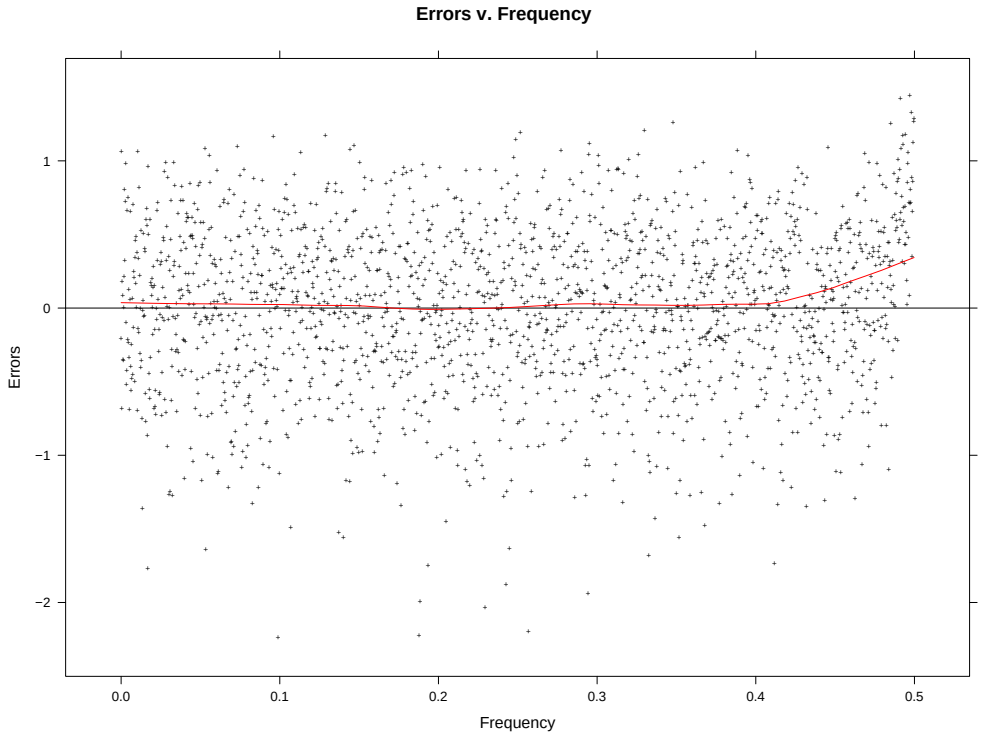

Fig. 3.41. ARFIMA(p = 1, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

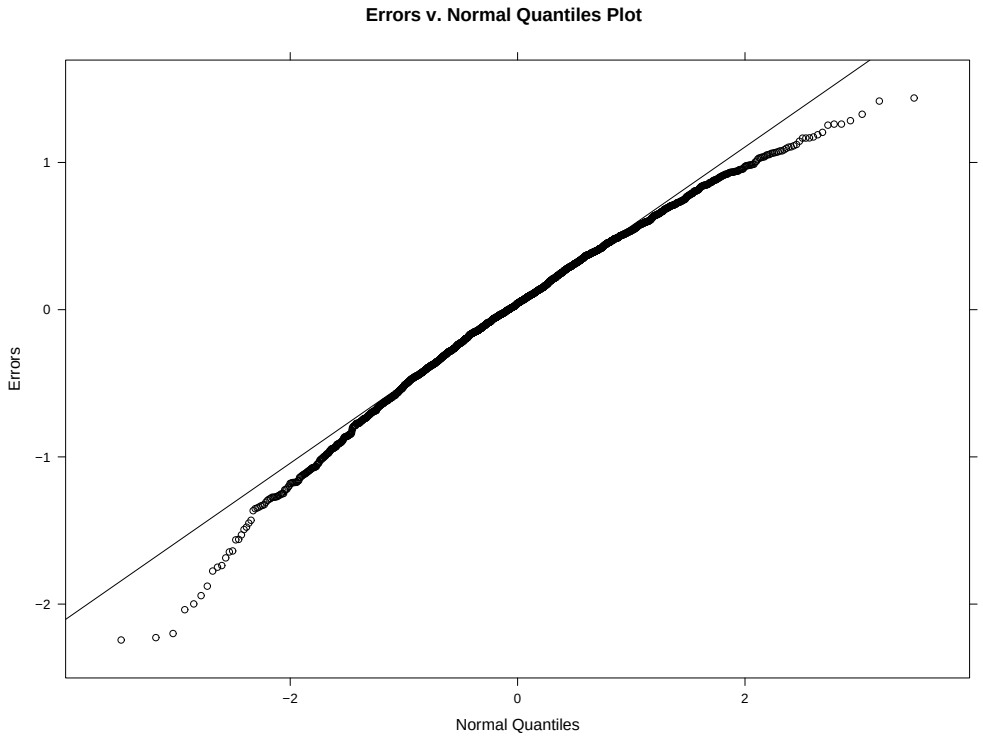

Fig. 3.42. ARFIMA(p = 1, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.7 FARMA(p = 2, d, q = 1)

The $FARMA(p=2,d,q=1)$ time series model is:

$$(1-\phi_1 B-\phi_2 B^2)(1-B)^d x_t = (1-1.5B+0.75B^2)(1-B)^{0.45}x_t = (1+0.75B)\varepsilon_t = (1+\theta B)\varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi_1,\phi_2,d,\theta,\sigma^2)) = \sigma^2 \frac{|1+\theta e^{-i2\pi f_j}|^2}{|1-\phi_1 e^{-i2\pi f_j}-\phi_2 e^{-i4\pi f_j}|^2}|1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi_1,\phi_2,d,\theta,\sigma^2))) + \varepsilon = \ln(\sigma^2) + \ln(|1+\theta e^{-i2\pi f_j}|^2) - d \cdot ln(|1-e^{-i2\pi f_j}|^2) - \ln(|1-\phi_1 e^{-i2\pi f_j} - \phi_2 e^{-i4\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

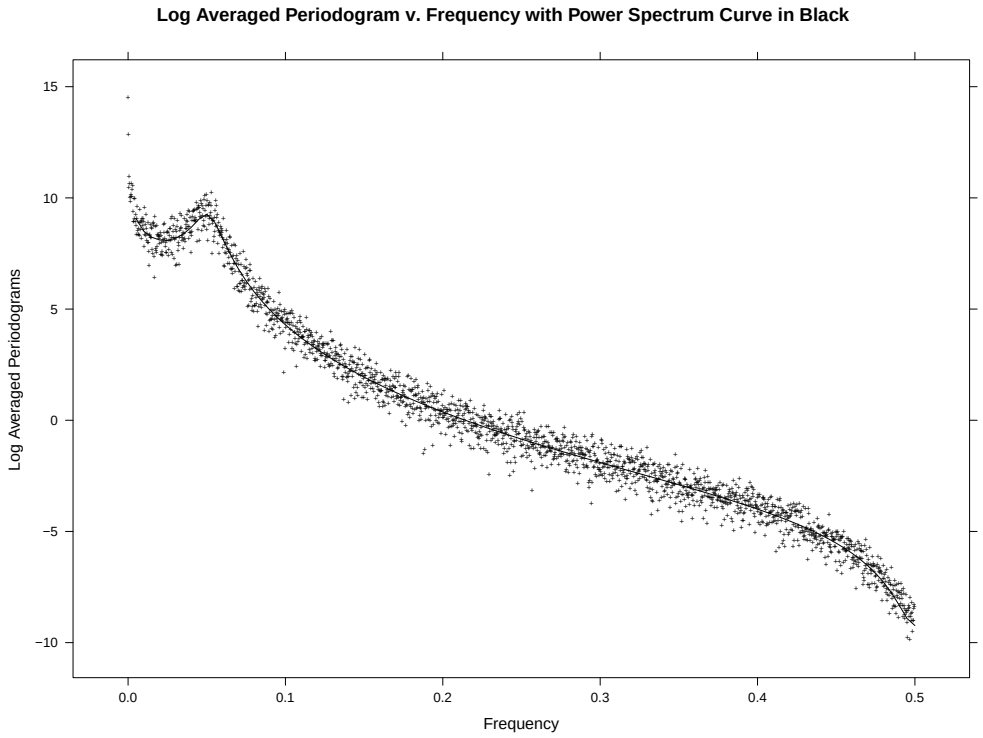

Fig. 3.43. ARFIMA(p = 2, d, q = 1) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

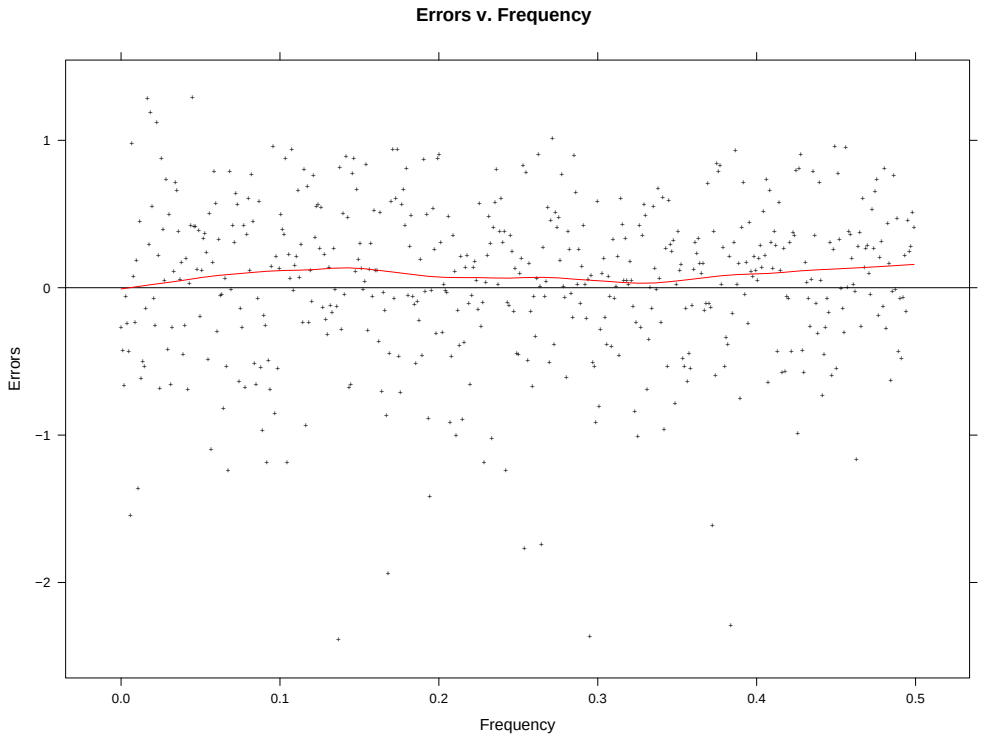

Fig. 3.44. ARFIMA(p = 2, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

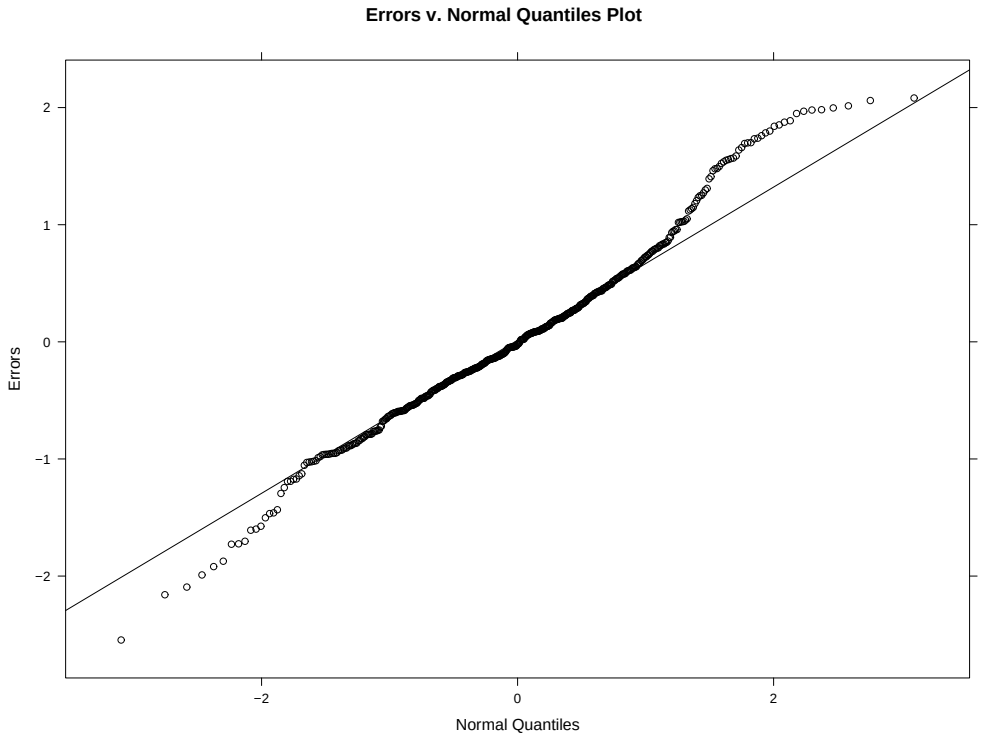

Fig. 3.45. ARFIMA(p = 2, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.8 FARMA(p = 1, d, q = 2)

The $FARMA(p = 1, d, q = 2)$ time series model is:

$$(1-\phi B)(1-B)^d x_t = (1-0.8B)(1-B)^{0.45} x_t = (1+1.5B+0.75B^2)\varepsilon_t = (1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi,d,\theta_1,\theta_2,\sigma^2)) = \sigma^2 \frac{|1+\theta_1 e^{-i2\pi f_j}+\theta_2 e^{-i4\pi f_j}|^2}{|1-\phi e^{-i2\pi f_j}|^2} |1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$y_j = \ln(S(f_j|(\phi, d, \theta_1, \theta_2, \sigma^2))) + \varepsilon = \ln(\sigma^2) - \ln(|1 - \phi e^{-i2\pi f_j}|^2) - d \cdot \ln(|1 - e^{-i2\pi f_j}|^2) + \ln(|1 + \theta_1 e^{-i2\pi f_j} + \theta_2 e^{-i4\pi f_j}|^2) + \varepsilon$$

The simulations yielded the following log averaged periodogram versus frequency results:

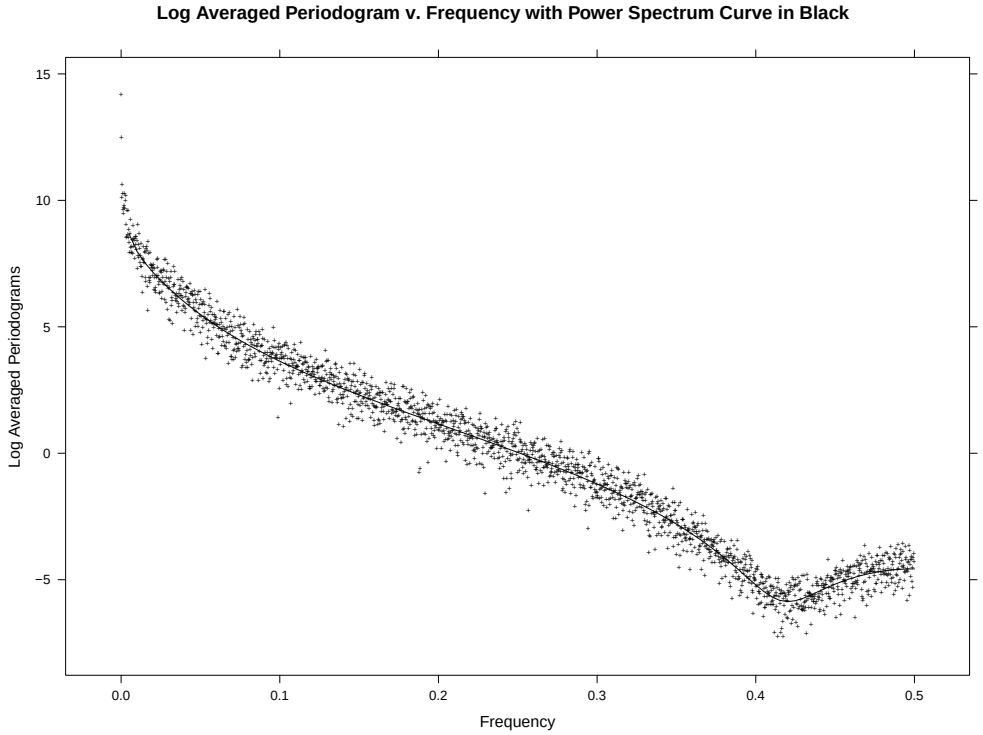

Fig. 3.46. ARFIMA(p = 1, d, q = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum as these plots demonstrate. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

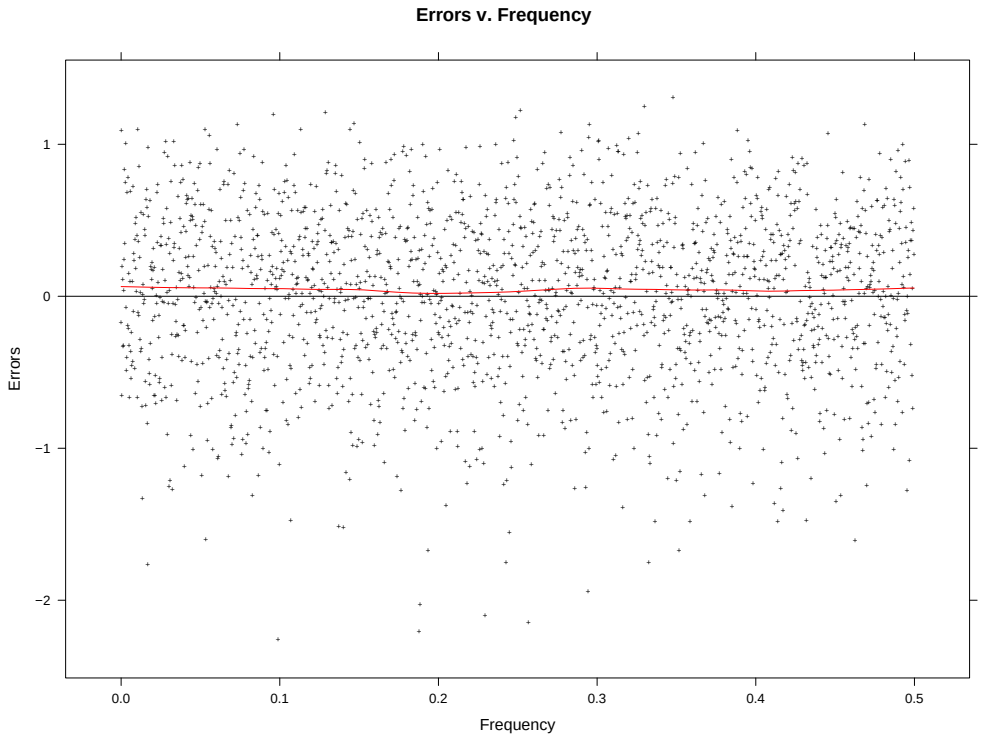

Fig. 3.47. ARFIMA(p = 1, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

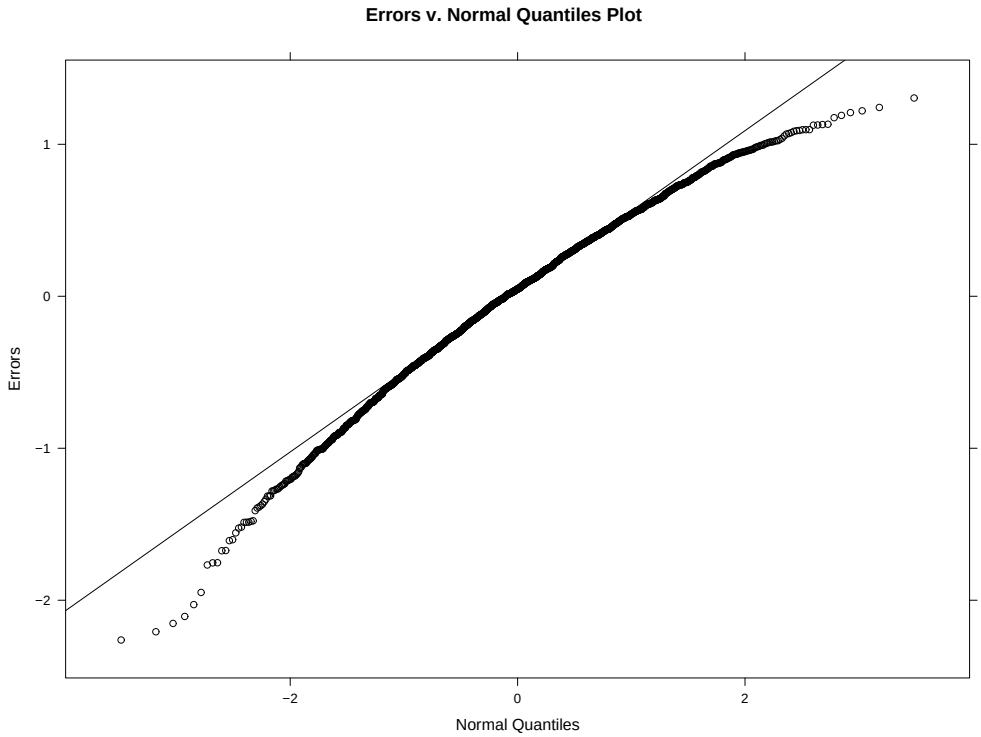

Fig. 3.48. ARFIMA(p = 1, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

3.2.9 FARMA(p = 2, d, q = 2)

The $FARMA(p=2,d,q=2)$ time series model is:

$$(1-\phi_1 B-\phi_2 B^2)(1-B)^d x_t = (1-1.5B+0.75B^2)(1-B)^{0.45}x_t$$

$$= (1+1.25B+0.5B^2)\varepsilon_t = (1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The power spectrum is:

$$S(f_j|(\phi_1,\phi_2,d,\theta_1,\theta_2,\sigma^2)) = \sigma^2 \frac{|1+\theta_1 e^{-i2\pi f_j}+\theta_2 e^{-i4\pi f_j}|^2}{|1-\phi_1 e^{-i2\pi f_j}-\phi_2 e^{-i4\pi f_j}|^2} |1-e^{-i2\pi f_j}|^{-2d}$$

The regression model is:

$$\begin{aligned} y_j = \ln(S(f_j|(\phi_1,\phi_2,d,\theta_1,\theta_2,\sigma^2))) + \varepsilon = \ln(\sigma^2) + \ln(|1+\theta_1 e^{-i2\pi f_j}+\theta_2 e^{-i4\pi f_j}|^2) \\ - \ln(|1-\phi_1 e^{-i2\pi f_j}-\phi_2 e^{-i4\pi f_j}|^2) \\ -d\cdot\ln(|1-e^{-i2\pi f_j}|^2)+\varepsilon \end{aligned}$$

The simulations yielded the following log averaged periodogram versus frequency results:

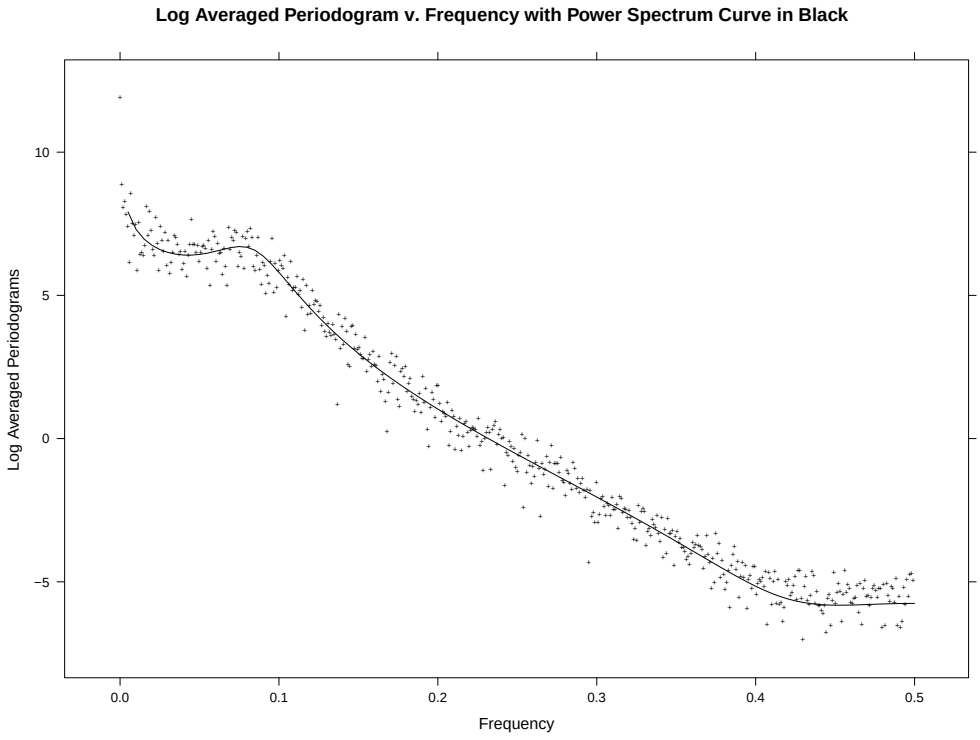

Fig. 3.49. ARFIMA(p = 2, d, q = 2) Log of the Averaged Periodogram v. Frequency

The log averaged periodogram ordinates are consistent with the true log power spectrum. However, a residuals plot will be a better diagnostic for visually identifying any potential deviations of our log averaged periodogram ordinates and the true log power spectrum:

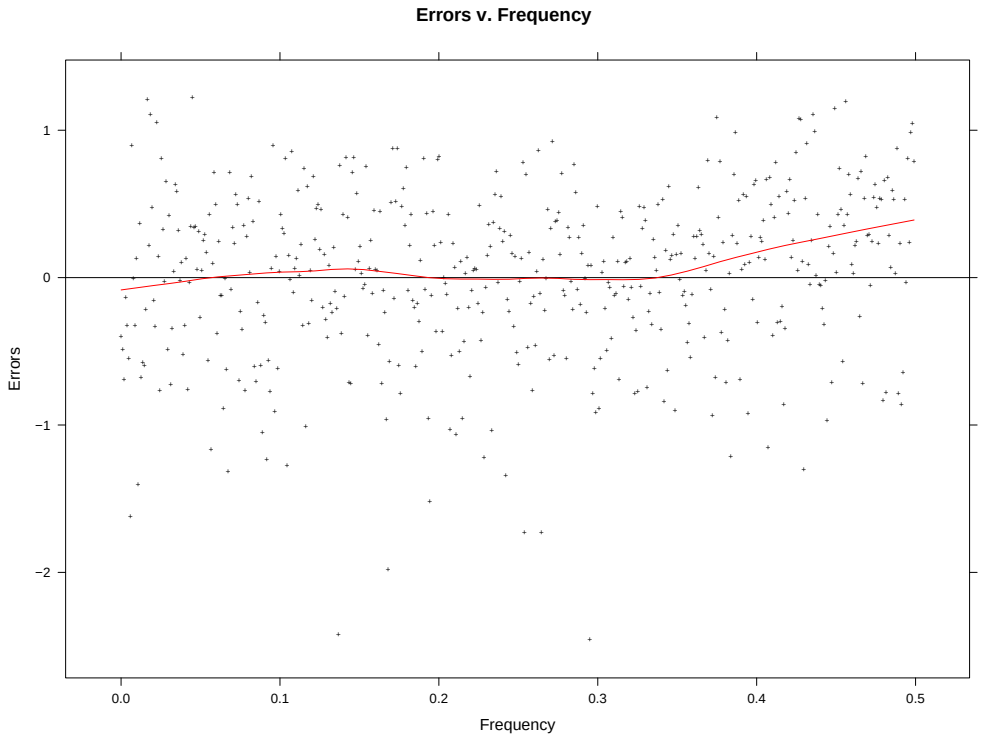

Fig. 3.50. ARFIMA(p = 2, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency

The residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals, Quantile-Quantile plots were constructed:

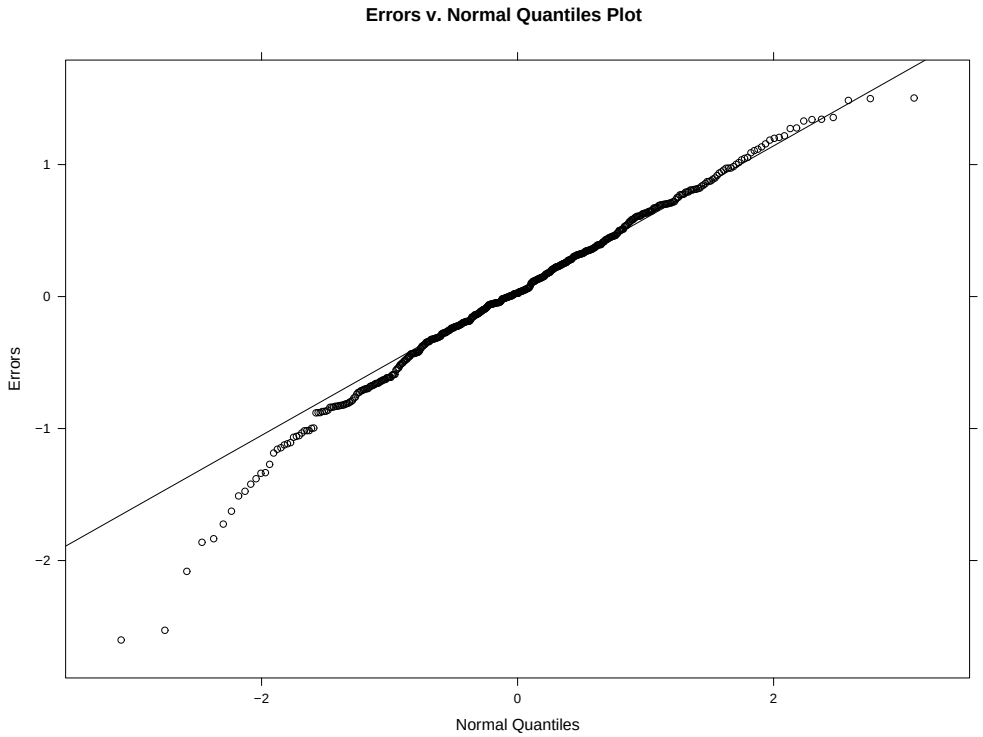

Fig. 3.51. ARFIMA(p = 2, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

4. ESTIMATING THE REGRESSION MODEL

All ARMA and FARMA PPS-REG models are nonlinear necessitating numerical or iterative methods to obtain parameter estimates. Parameter estimates were obtained with Nonlinear Least Squares (NLS) between the log averaged periodograms and the log power spectrum function. The *nls()* routine in R [6] was used to carry out these parameter estimations.

Identification of a model is done by a feed forward, feed back loop exploratory analysis on a case by case basis. For the purposes of effectively computing estimation results for 100 simulation results at a time, lower order parameter models providing initial estimates of only $\sigma^2_{(0)} = e^{\overline{y_k}}$ sufficed. Some higher order parameter ARMA models simulation runs required initial estimates from Conditional Sum of Squares, so it was utilized for all simulation runs. Higher order parameter FARMA models had ARMA parameter estimates initialized with Conditional Sums of Squares and $d^{(0)} = 0.25$.

Comparison of PPS-REG estimates to Maximum Likelihood Estimation (MLE) was done with scatterplots and QQ plots. Scatterplots identified the relationship of parameter estimates between PPS-REG and MLE estimation methods. QQ plots assessed the distributional properties of the parameter estimates per estimation method. To obtain ARMA model MLEs, the R routine *arima()* [7] was used and to obtain FARMA model MLEs, the R routine *fracdiff()* [8] was used.

When Auto-Regressive (AR) Filtration was an available technique, to validate the technique, we utilized multiple model diagnostics. First, to assert correct model selection, plots of log averaged periodogram versus frequency with PPS-REG estimated parameters log power spectrum curve were constructed. Residuals plots of the difference between the log averaged periodogram and PPS-REG estimated log power spectrum versus frequency assessed the fit of our model selection. Distributional properties of these residuals were evaluated with Quantile-Quantile (QQ) plots. These diagnostic plots were followed by the

same scatterplots and QQ plots as were utilized in determining the effectiveness of our regression model, but now on our AR filtered regression model.

All simulation parameters are identical to their respective Section 3 counterparts.

4.1 ARMA Models

4.1.1 AR(p = 1)

PPS-REG parameter estimates were initialized to be $(\phi^{(0)}, \sigma^2_{(0)}) = (0, e^{\overline{y_k}})$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

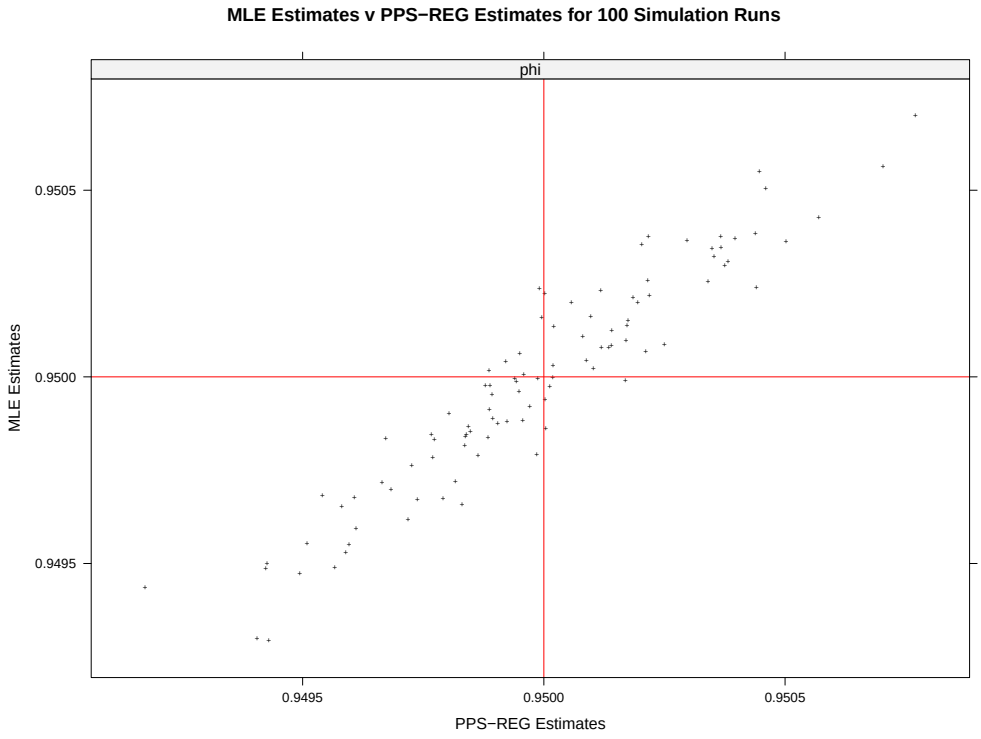

Fig. 4.1. AR(p = 1) Scatterplot of MLE estimates of ϕ v. PPS-REG estimates of ϕ

The results are comparable and more importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of ϕ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

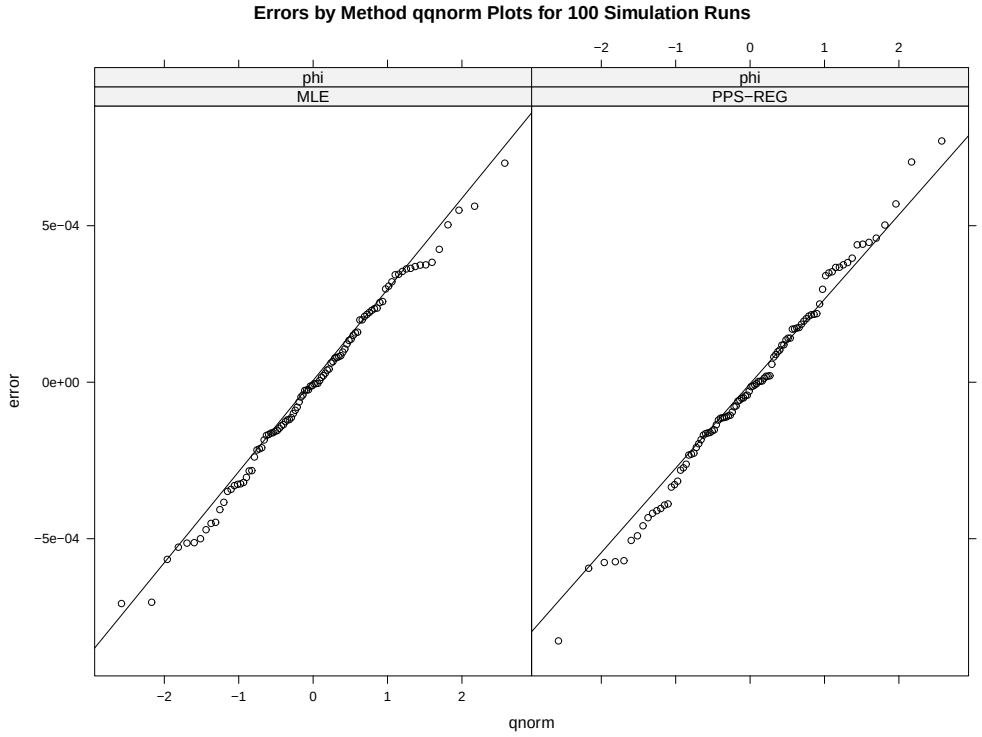

Fig. 4.2. AR(p = 1) Normal QQ-Plots of Estimation Error of ϕ by Estimation Method, PPS-REG v. MLE

Both estimation methods across differing simulations yield $\hat{\phi}$ estimates that are normally distributed about the simulated ϕ value.

4.1.2 AR(p = 2)

PPS-REG parameter estimates were initialized to be $(\phi_1^{(0)}, \phi_2^{(0)}, \sigma^2_{(0)}) = (0.95, 0, e^{\overline{y_k}})$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

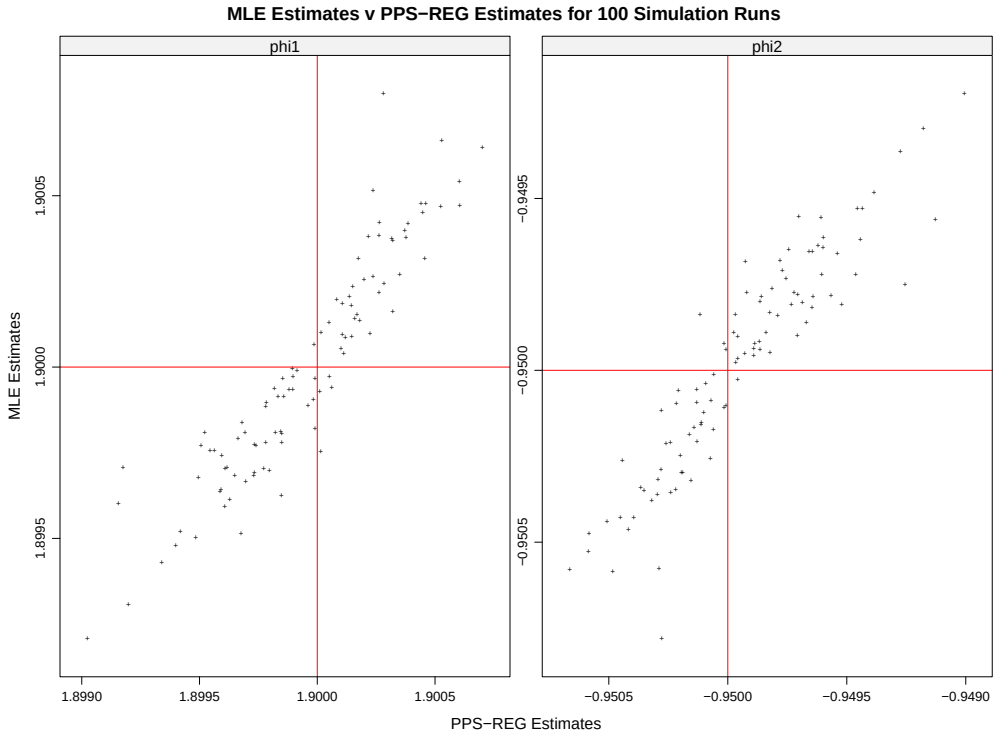

Fig. 4.3. AR(p = 2) Scatterplot of MLE estimates v. PPS-REG estimates by ϕ_1 and ϕ_2

The results are comparable and more importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of ϕ_1 and ϕ_2 indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

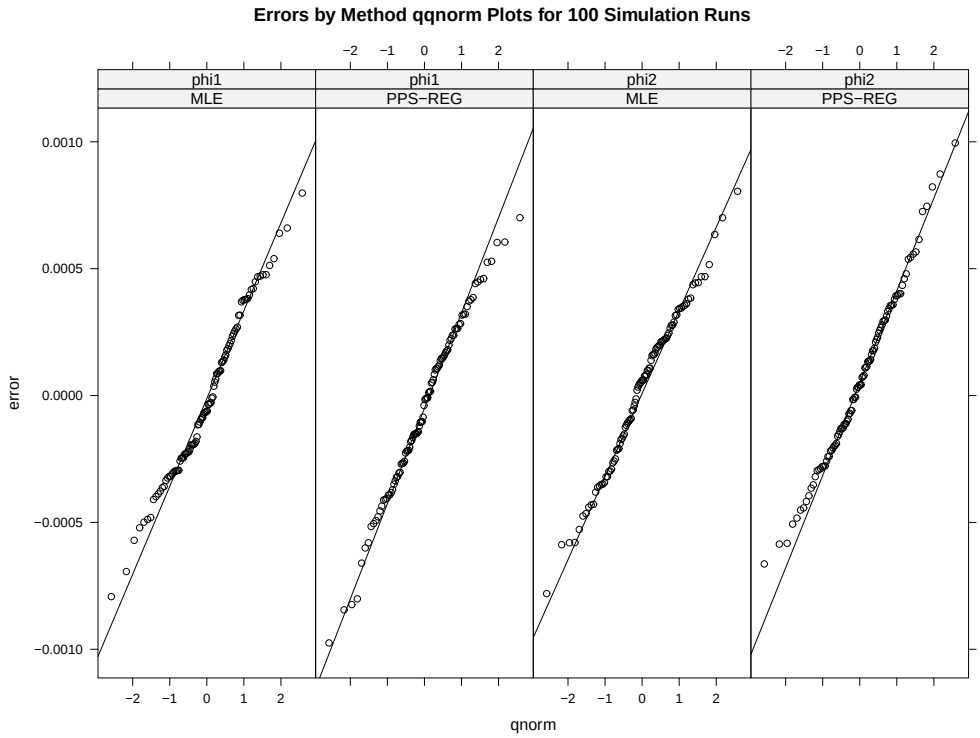

Fig. 4.4. AR(p = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ_1 or ϕ_2

Both estimation methods across differing simulations yield $\hat{\phi_1}$ and $\hat{\phi_2}$ estimates that are normally distributed about the simulated ϕ_1 and ϕ_2 values respectively.

4.1.3 MA(q = 1)

PPS-REG parameter estimates were initialized to be $(\theta^{(0)}, \sigma^2_{(0)}) = (0, e^{\overline{y_k}})$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

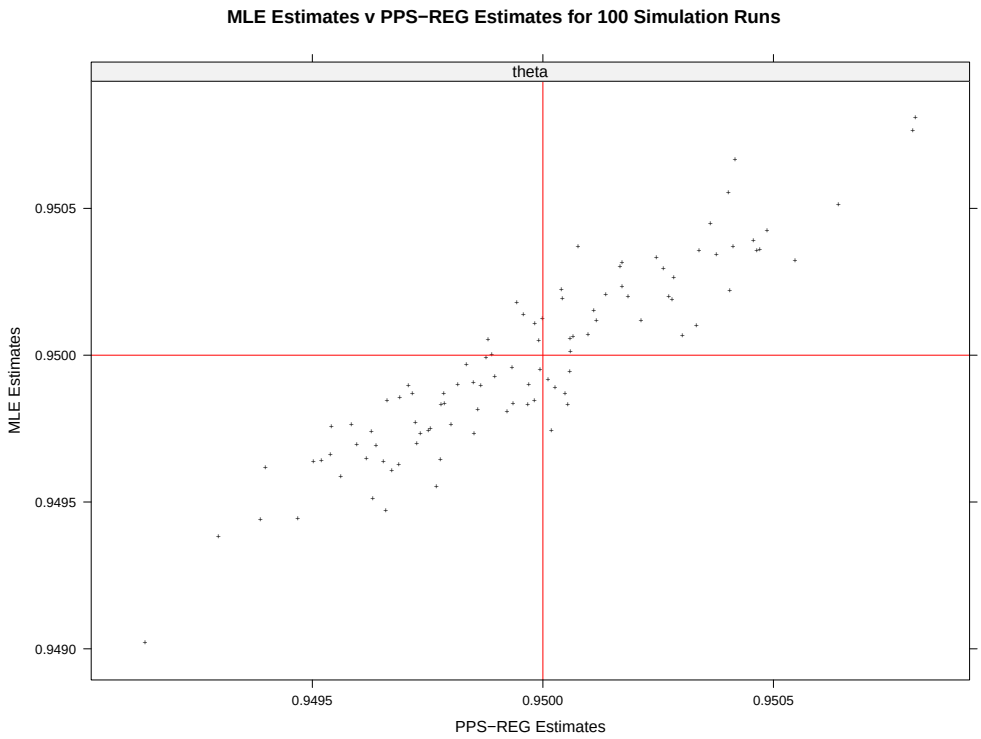

Fig. 4.5. MA(q = 1) Scatterplot of MLE estimates of ϕ v. PPS-REG estimates of ϕ

The results are comparable and more importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of θ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

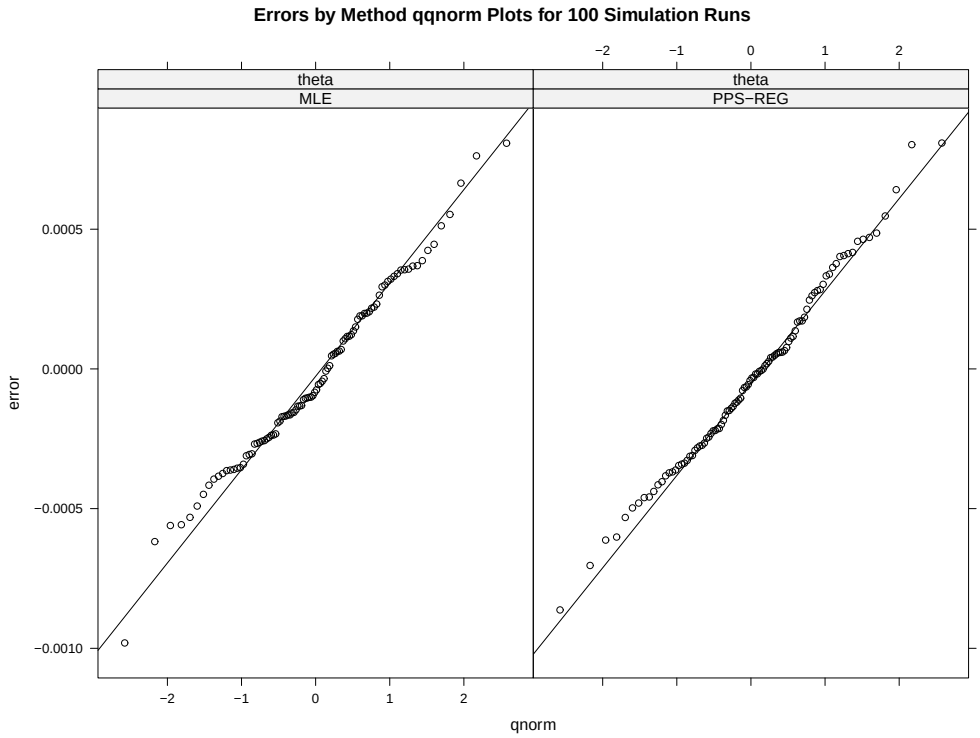

Fig. 4.6. MA(q = 1) Normal QQ-Plots of Estimation Error of θ by Estimation Method, PPS-REG v. MLE

Both estimation methods across differing simulations yield $\hat{\theta}$ estimates that are normally distributed about the simulated θ value.

4.1.4 MA(q = 2)

PPS-REG parameter estimates were initialized to be $(\theta_1^{(0)}, \theta_2^{(0)}, \sigma^2_{(0)}) = (1.25, 0.5, e^{\overline{y_k}})$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

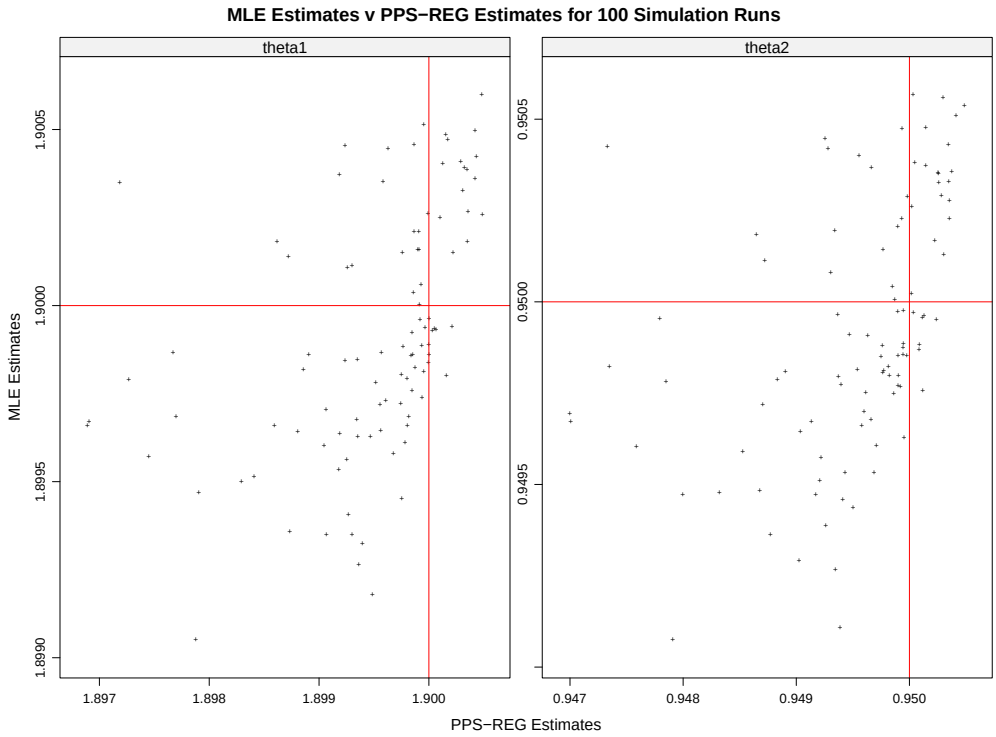

Fig. 4.7. MA(q = 2) Scatterplot of MLE estimates v. PPS-REG estimates by θ_1 and θ_2

The results are comparable and more importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of θ_1 and θ_2 indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

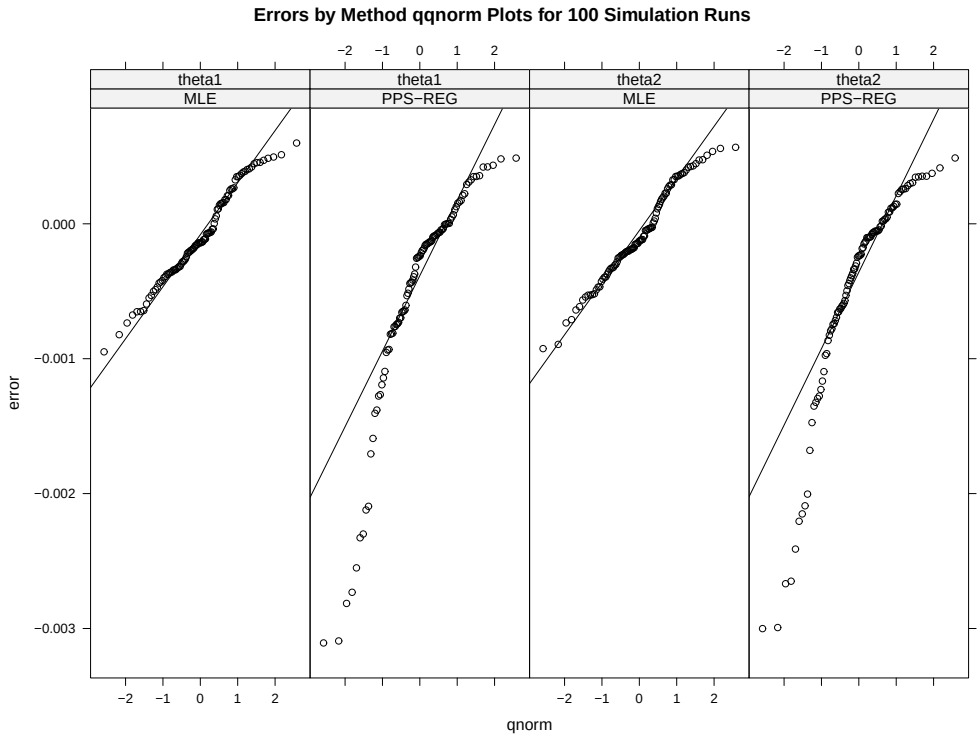

Fig. 4.8. MA(q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, θ_1 or θ_2

4.1.5 ARMA(p = 1, q = 1)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

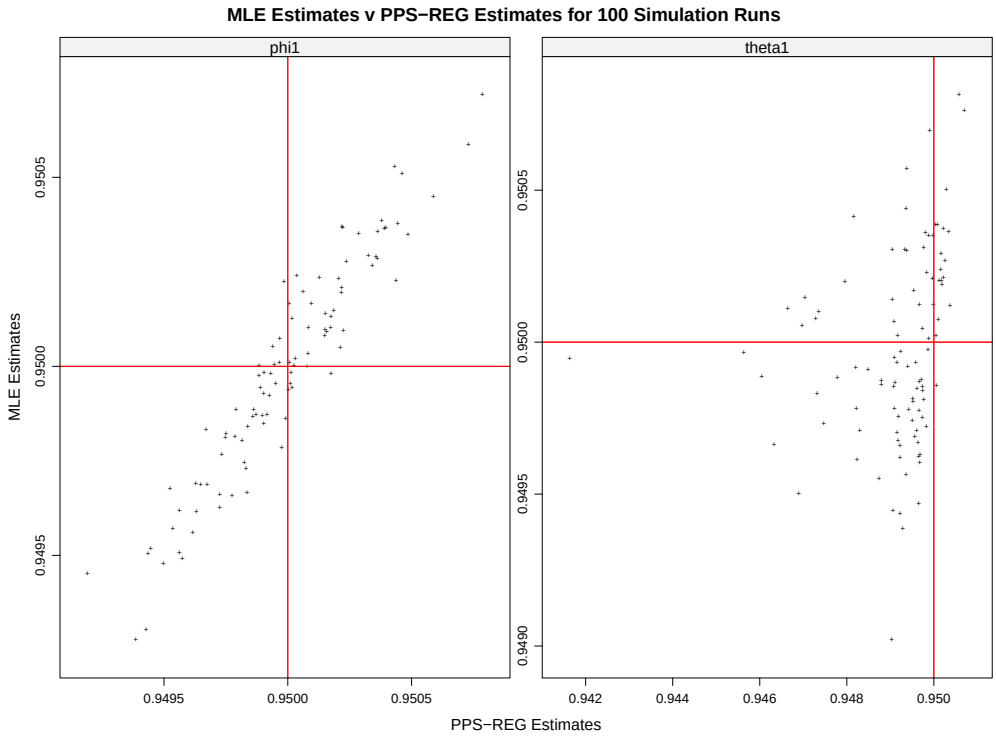

Fig. 4.9. ARMA(p = 1, q = 1) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ or θ

The results are comparable and more importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of θ and ϕ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

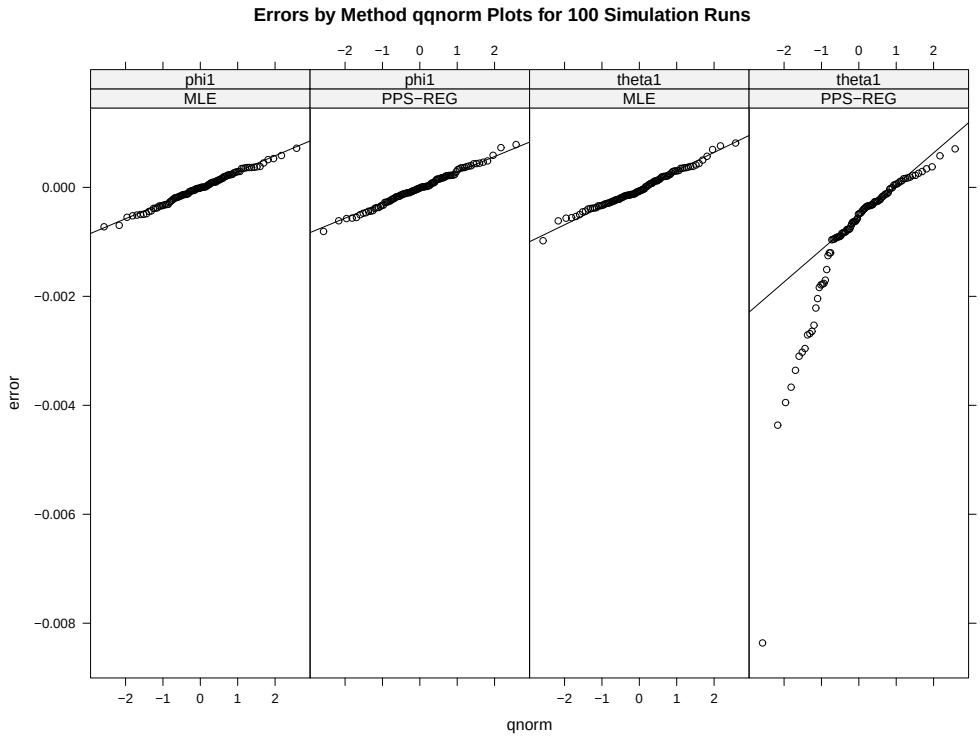

Fig. 4.10. ARMA11(p = 1, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, θ or ϕ

PPS-REG exhibits a heavy left tail for $\hat{\theta}$ on a little more than 10% of the simulation runs, but otherwise is normally distributed. The MLE estimates across differing simulations yield $\hat{\theta}$ and $\hat{\phi}$ estimates that are normally distributed about the simulated θ and ϕ values respectively.

AR Component Filtration

Parameter estimation difficulties are greater for MA parameters than AR parameters. The effect of AR parameters, once estimated, can be filtered out from the data. The same MA model can be fit to the filtered data to improve θ parameter estimates. This estimation approach is similar to that of Marple [9]

The $ARMA(p=1,q=1)$ time series model is:

$$x_t = \phi x_{t-1} + \theta\varepsilon_{t-1} + \varepsilon_t$$

The $MA(q=1)$ filtered data time series model is:

$$x_t^* = x_t - \hat{\phi}x_{t-1} = (\phi - \hat{\phi})x_{t-1} + \theta\varepsilon_{t-1} + \varepsilon_t = \varepsilon_\phi x_{t-1} + \theta\varepsilon_{t-1} + \varepsilon_t \approx \theta\varepsilon_{t-1} + \varepsilon_t$$

The PPS-REG log power spectrum estimation on the $MA(q=1)$ filtered data yielded the following log averaged periodogram versus frequency results:

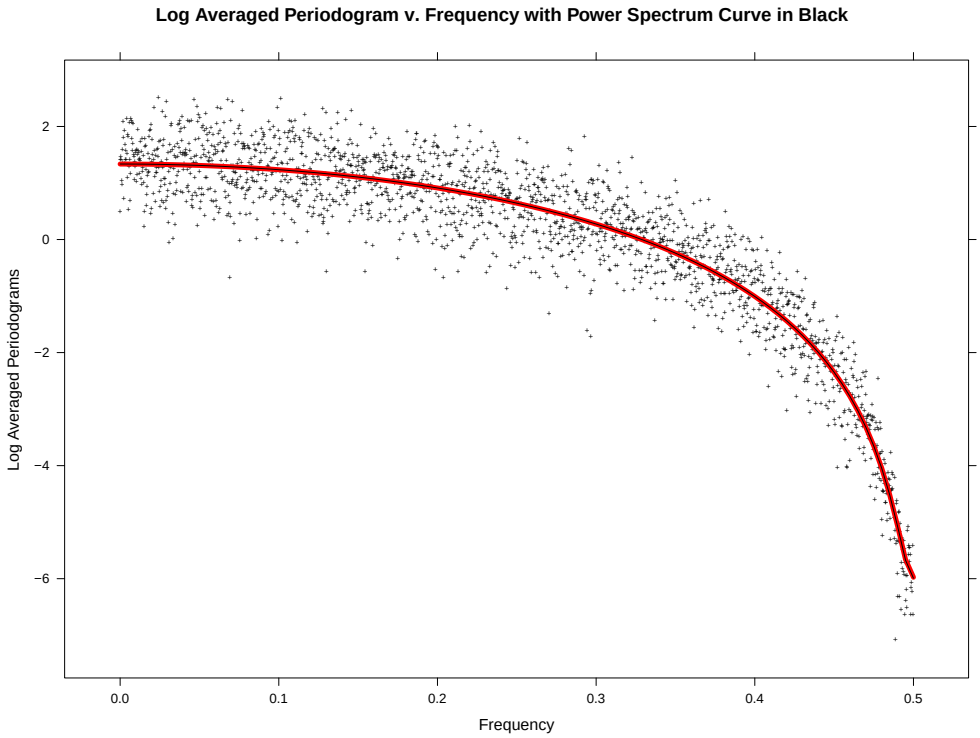

Fig. 4.11. MA(q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

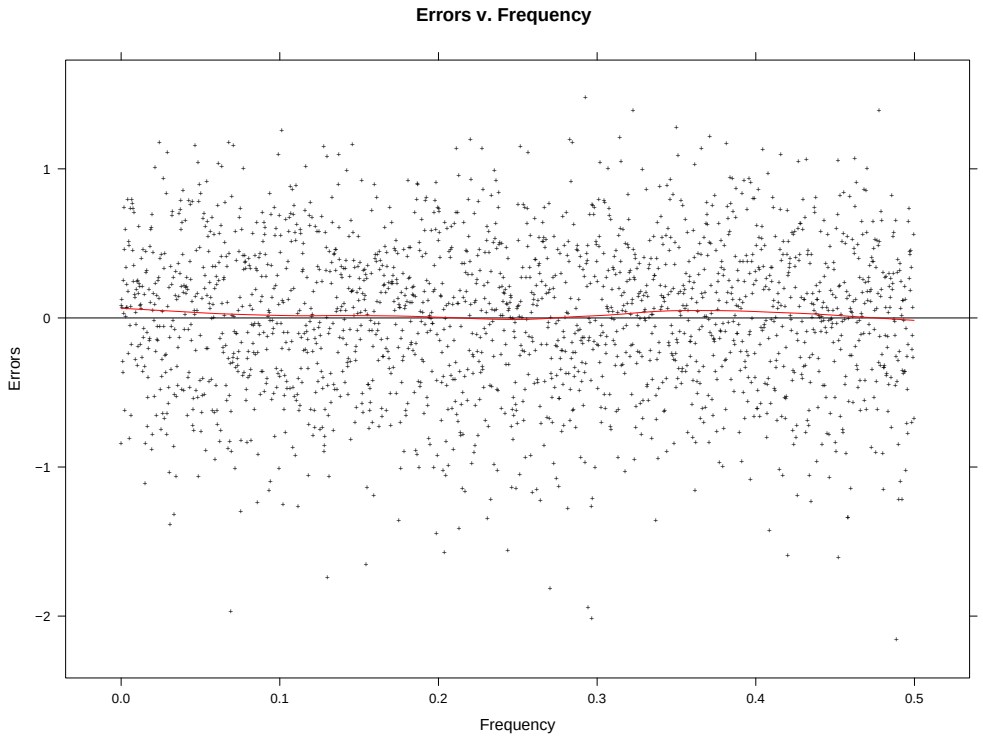

Fig. 4.12. MA(q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

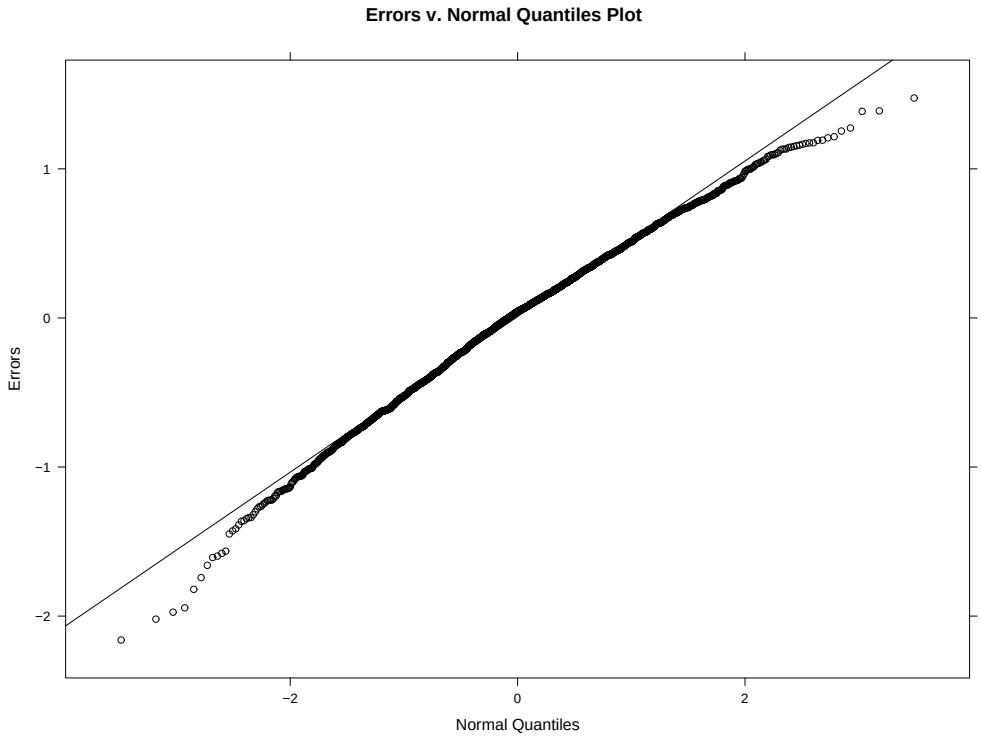

Fig. 4.13. MA(q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

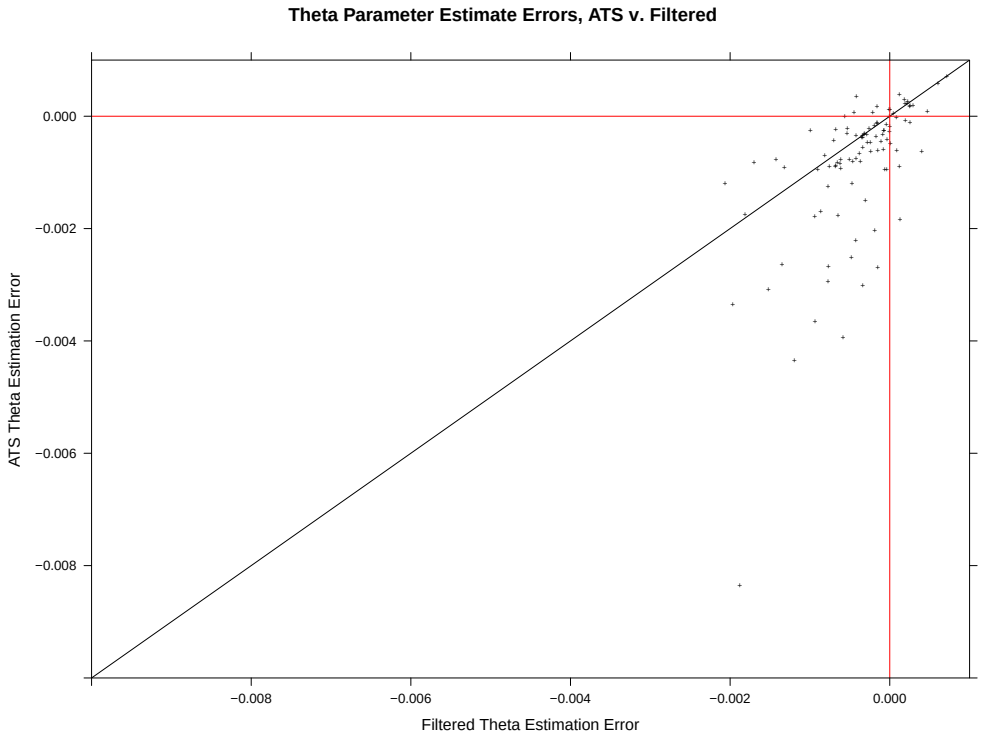

Fig. 4.14. MA(q = 1) Scatterplot of PPS-REG estimates v. PPS-REG Filtered estimates

The worst PPS-REG errors are improved by filtering and more importantly there is a positive linear relationship between the PPS-REG error and the PPS-REG filtered error of θ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

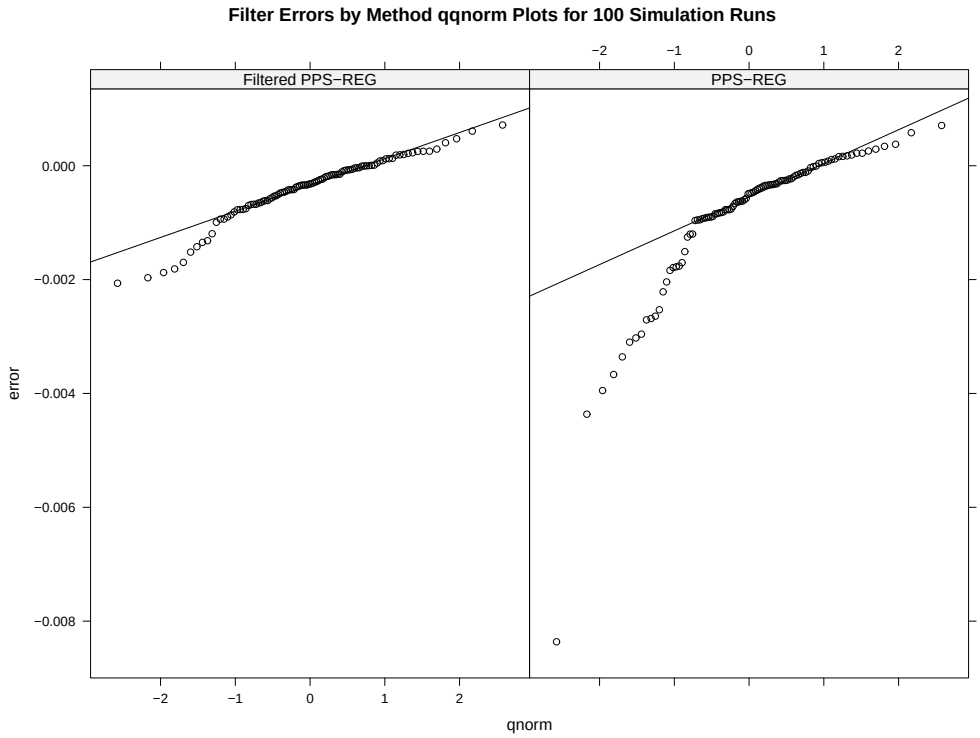

Fig. 4.15. MA1(q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

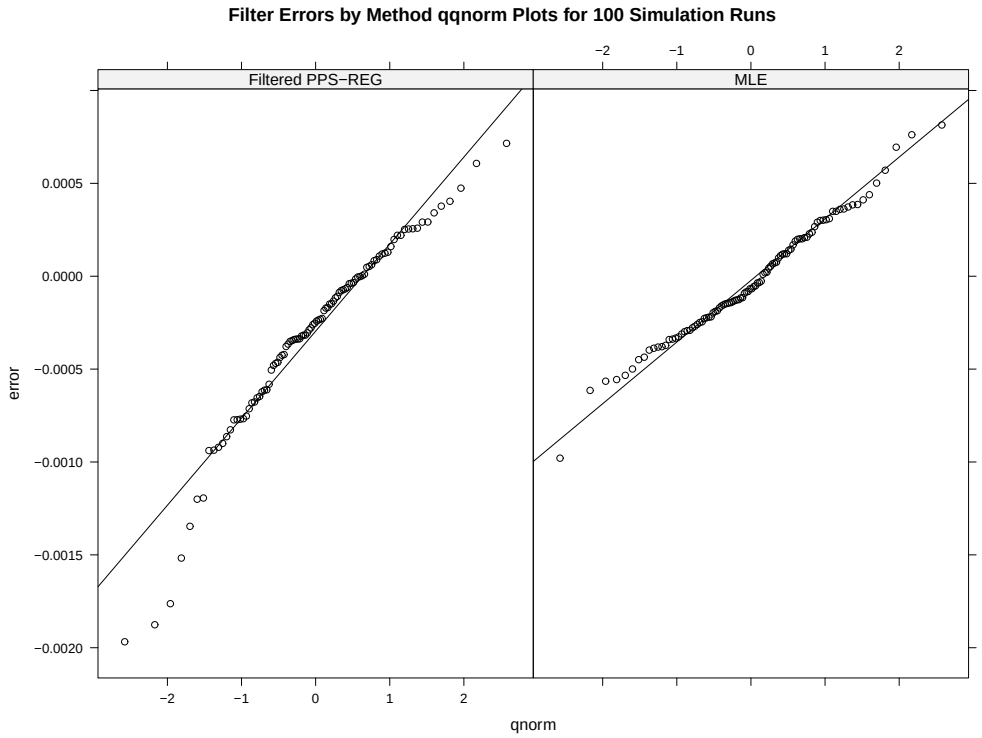

Fig. 4.16. MA1(q = 1) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE

Filtered PPS-REG exhibits a left tail for $\hat{\theta}$, but otherwise is normally distributed and is an improvement from a heavier PPS-REG left tail.

4.1.6 ARMA(p = 2, q = 1)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for all simulation runs. In runs where Conditional Sum of Squares was unsuccessful for the full model, PPS-REG parameter estimates were initialized with Conditional Sum of Squares estimates for an $AR(p = 2)$ model with $\theta^{(0)} = 0$. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

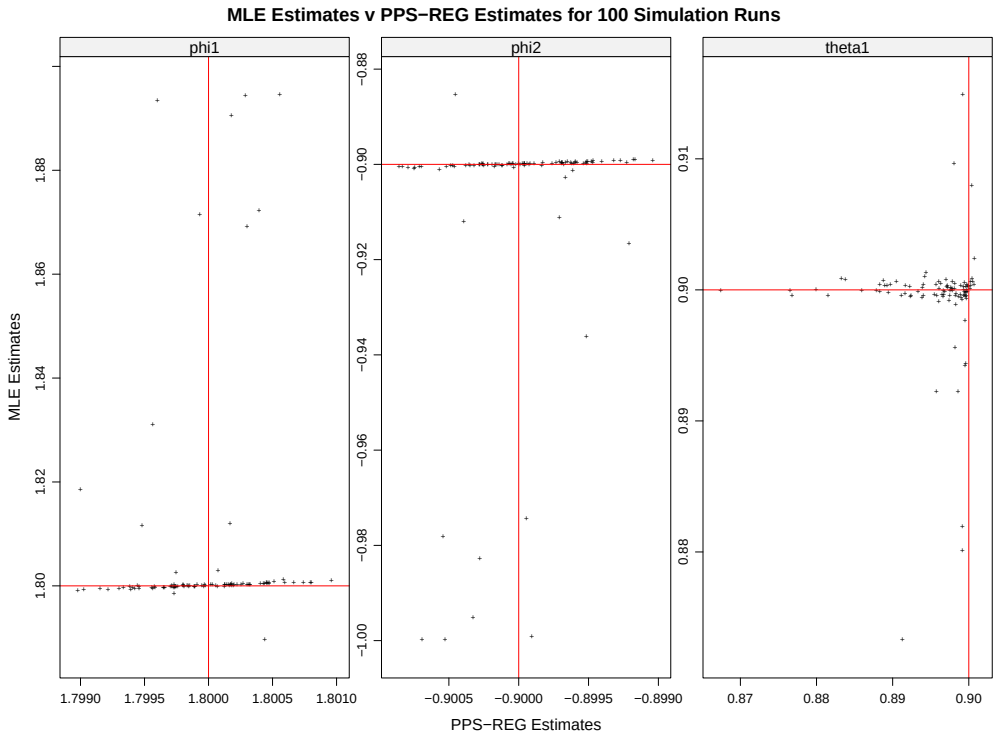

Fig. 4.17. ARMA(p = 2, q = 1) Scatterplot of MLE estimates v. PPS-REG estimates by ϕ_1, ϕ_2, or θ

MLE was consistently accurately, but had some simulations runs that varied from simulated ϕ_1 and ϕ_2 parameter values while PPS-REG did not possess any significant estimation problems for any of the parameters. The default iterative estimation method in *arima()* is "CSS-ML", which initiates parameter estimates with the Conditional Sum of Squares (CSS) estimate, before iterating to obtain the MLE. This failed on some test runs, so for consistency, elected to choose the method "ML" for all simulation runs. Running this iterative method with an initial parameter estimate of the origin is what yielded results for the $ARMA(p = 2, q = 1)$ model.

To assess the normality of our estimates by method, QQ plots were constructed:

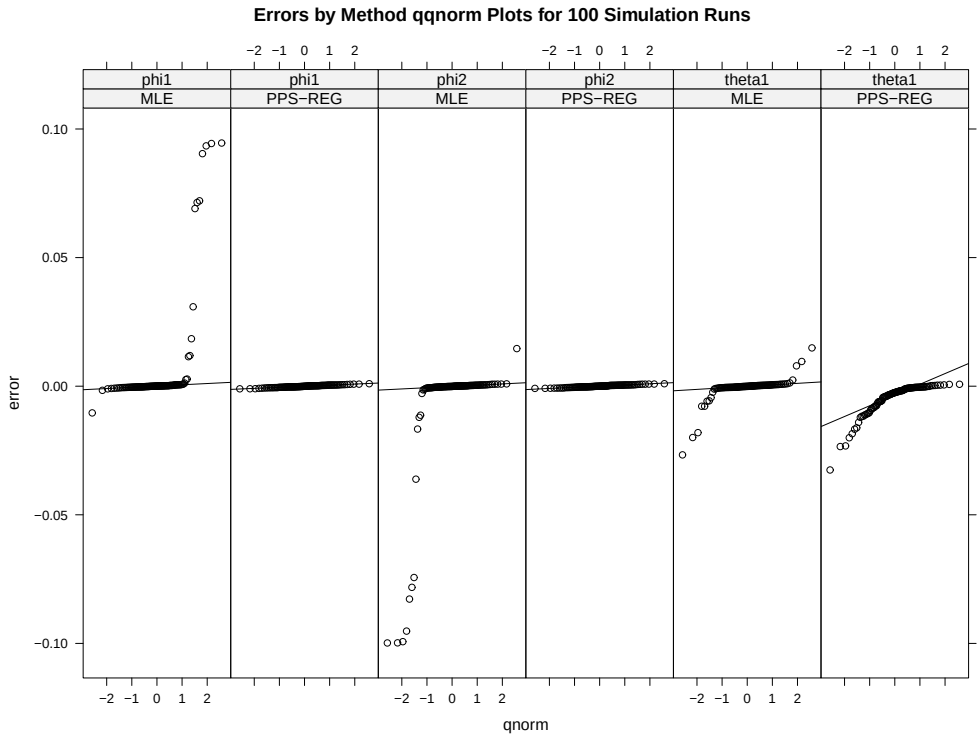

Fig. 4.18. ARMA(p = 2, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and parameter, ϕ_1, ϕ_2, or θ

PPS-REG parameter estimates are normally distributed, while MLE parameter estimates possess deviations from normality only due to a little more than 10% of the simulation runs deviating from simulated ϕ_1 and ϕ_2 parameter values when using the "ML" estimation method in the *arima()* routine.

AutoRegressive (AR) Component Filtration

The $ARMA(p=2,q=1)$ time series model is:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \theta\varepsilon_{t-1} + \varepsilon_t$$

The $MA(q=1)$ filtered data time series model is:

$$x_t^* = x_t - \hat{\phi_1} x_{t-1} - \hat{\phi_2} x_{t-2} = (\phi_1 - \hat{\phi_1}) x_{t-1} + (\phi_2 - \hat{\phi_2}) x_{t-2} + \theta\varepsilon_{t-1} + \varepsilon_t$$

$$= \varepsilon_{\phi_1} x_{t-1} + \varepsilon_{\phi_2} x_{t-2} + \theta\varepsilon_{t-1} + \varepsilon_t \approx \theta\varepsilon_{t-1} + \varepsilon_t$$

The PPS-REG log power spectrum estimation on the $MA(q=1)$ filtered data yielded the following log averaged periodogram versus frequency results:

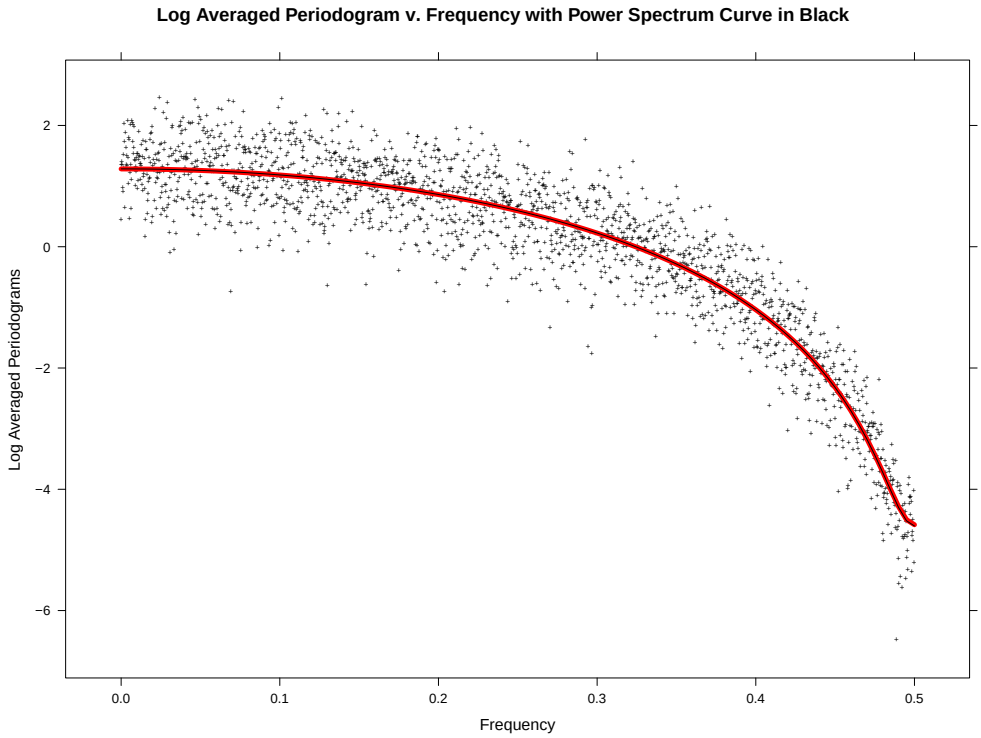

Fig. 4.19. MA(q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

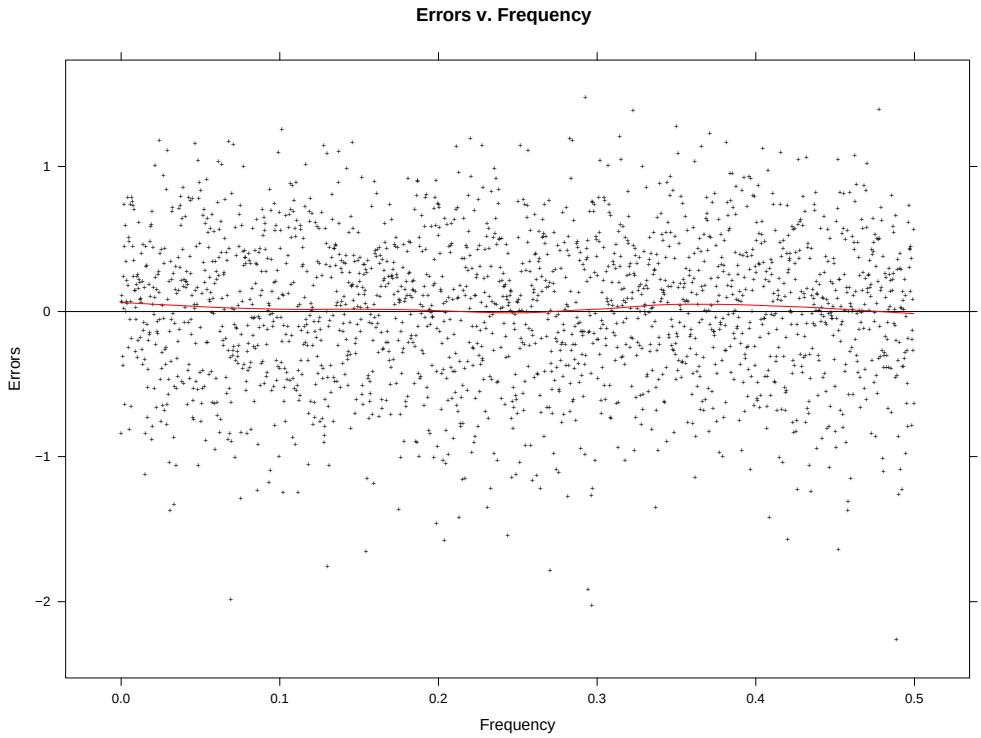

Fig. 4.20. MA(q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

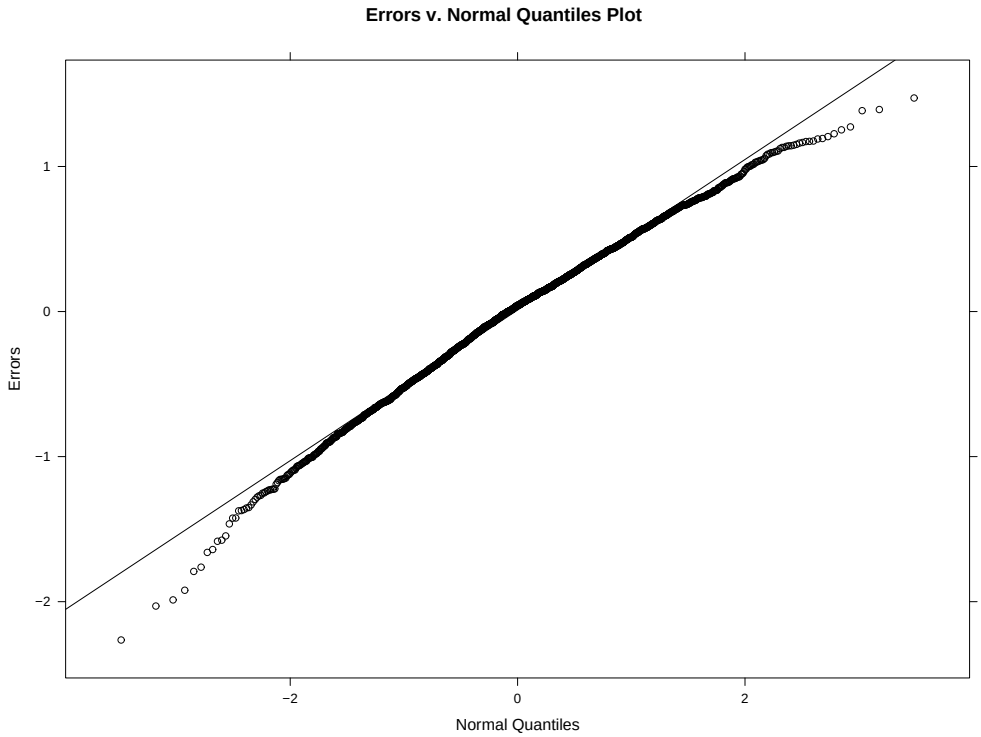

Fig. 4.21. MA(q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

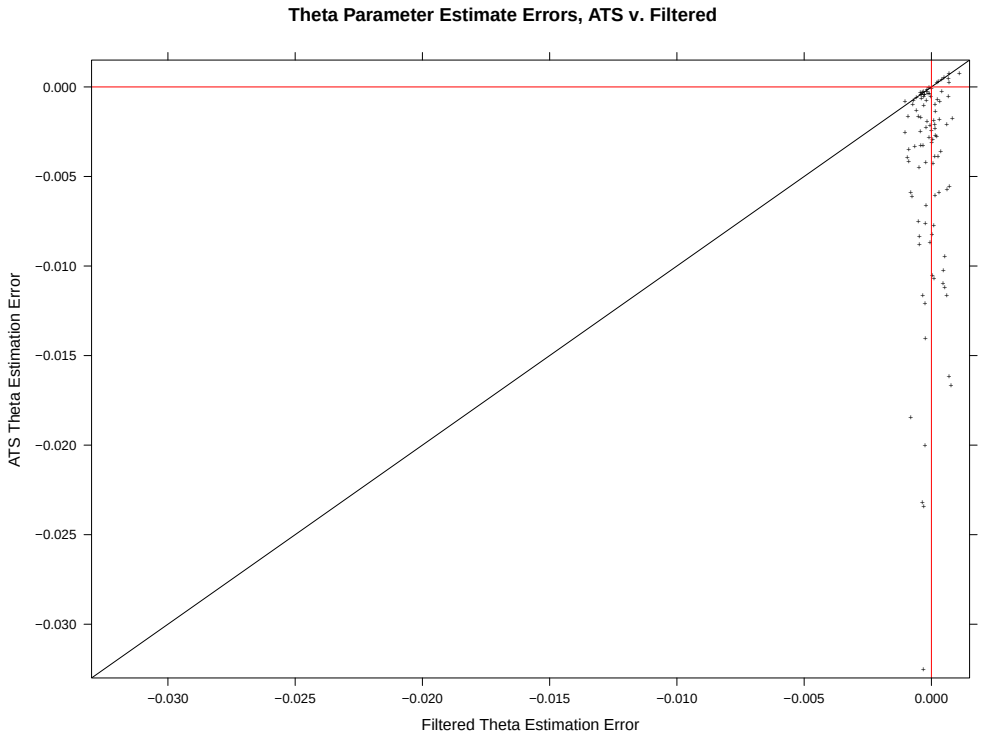

Fig. 4.22. MA(q = 1) Scatterplot of PPS-REG estimates v. PPS-REG Filtered estimates

The worst PPS-REG errors are improved by filtering.

To assess the normality of our estimates by method, QQ plots were constructed:

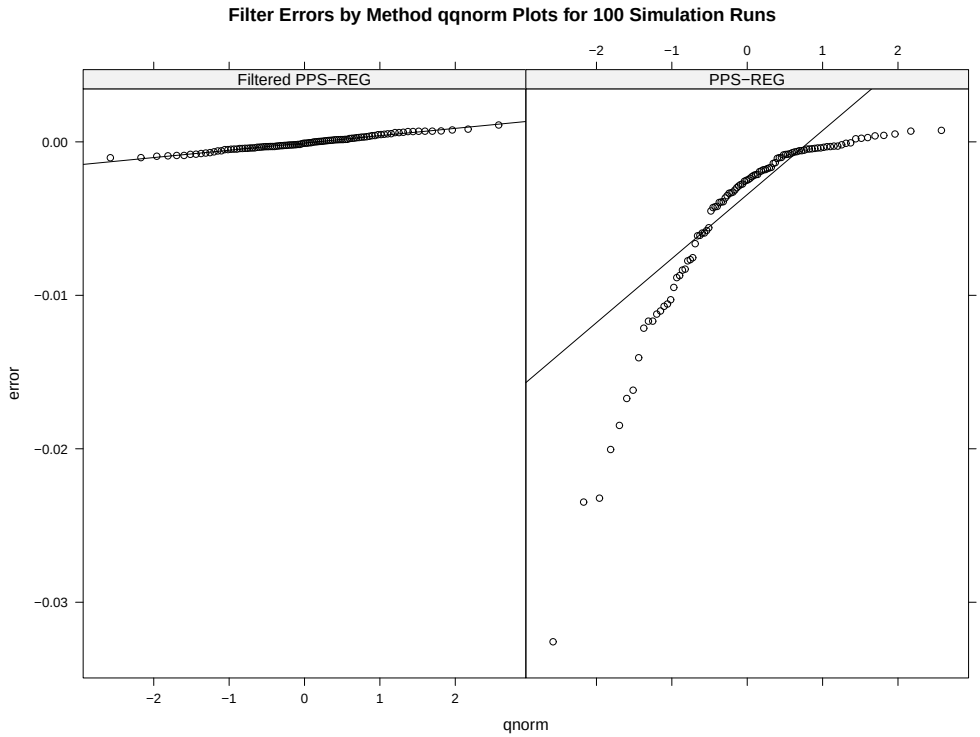

Fig. 4.23. MA(q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

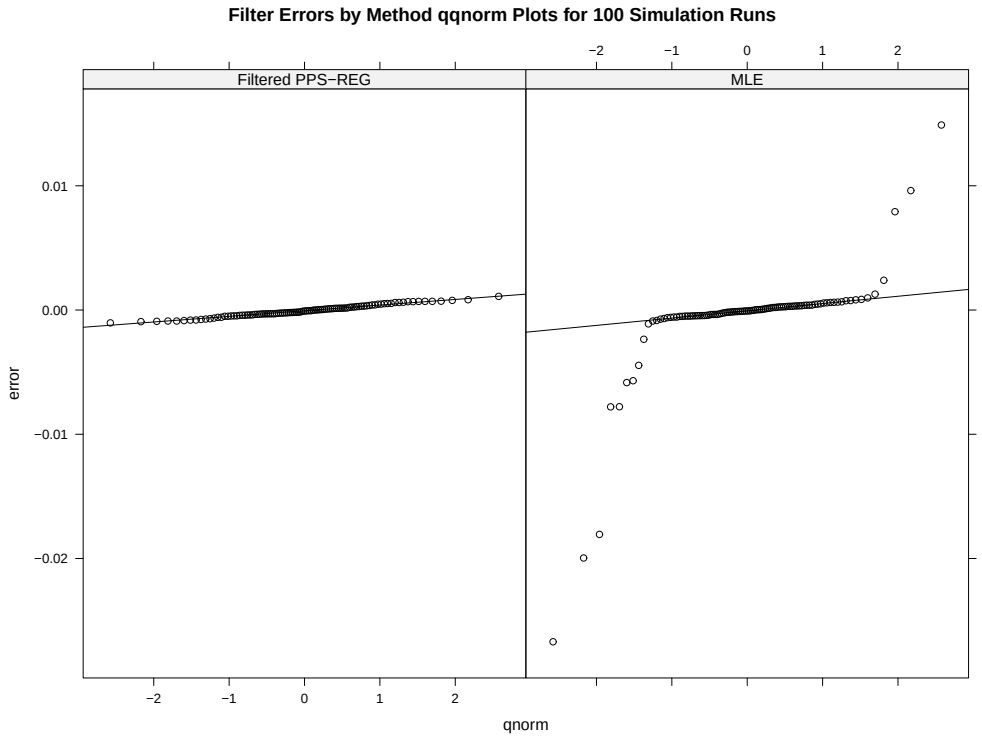

Fig. 4.24. MA(q = 1) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE

Filtered PPS-REG exhibits a left tail for $\hat{\theta}$, but otherwise is normally distributed and is an improvement from a heavier PPS-REG left tail.

4.1.7 ARMA(p = 1, q = 2)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

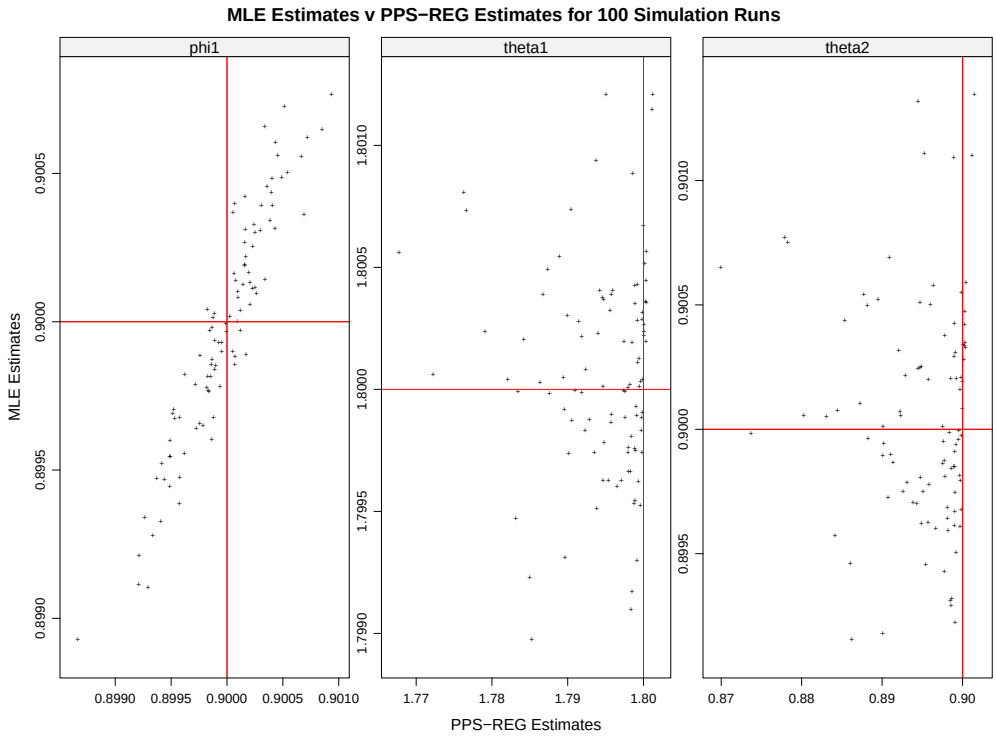

Fig. 4.25. ARMA(p = 1, q = 2) Scatterplot of MLE estimates v. PPS-REG estimates by ϕ, θ_1, or θ_2

The results are comparable for the ϕ parameter, but exhibit more variability and a slight bias towards the origin for θ_1 and θ_2 for PPS-REG estimation compared to MLE estimates.

To assess the normality of our estimates by method, QQ plots were constructed:

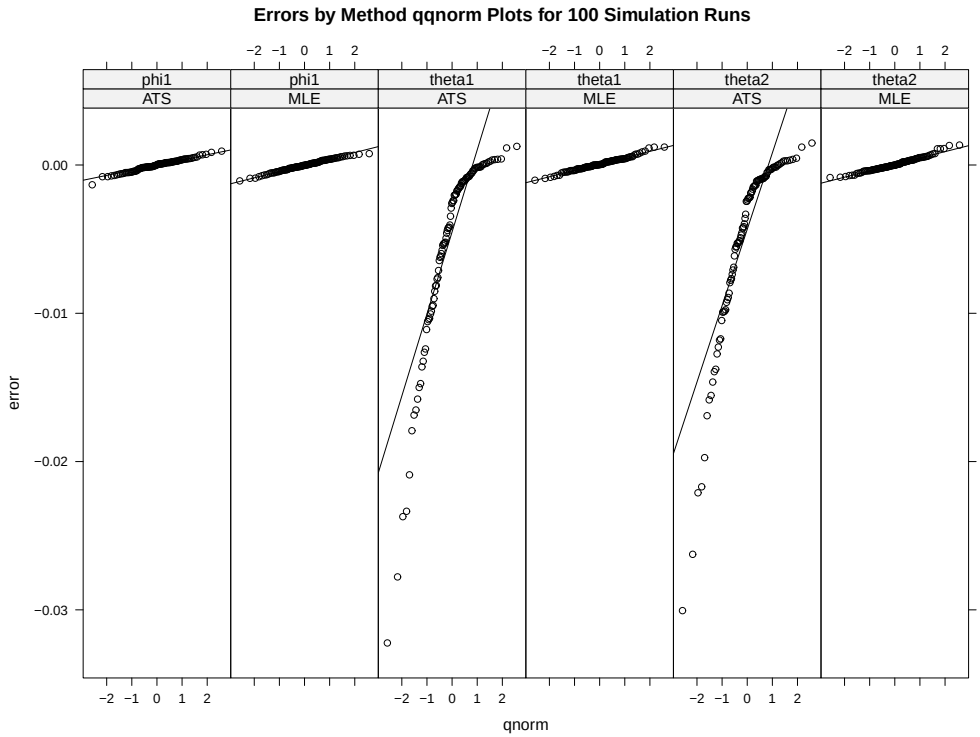

Fig. 4.26. ARMA(p = 1, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and parameter, ϕ, θ_1, or θ_2

PPS-REG exhibits a heavy left tail for $\hat{\theta}_1$ and $\hat{\theta}_2$ on a little more than 10% of the simulation runs, but otherwise is normally distributed. The MLE estimates across differing simulations yield parameter estimates that are normally distributed about the simulated values.

AutoRegressive (AR) Component Filtration

The $ARMA(p=1,q=2)$ time series model is:

$$x_t = \phi x_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

The $MA(q=2)$ filtered data time series model is:

$$x_t^* = x_t - \hat{\phi} x_{t-1} = (\phi - \hat{\phi}) x_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t = \varepsilon_\phi x_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

$$\approx \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

The PPS-REG log power spectrum estimation on the $MA(q=1)$ filtered data yielded the following log averaged periodogram versus frequency results:

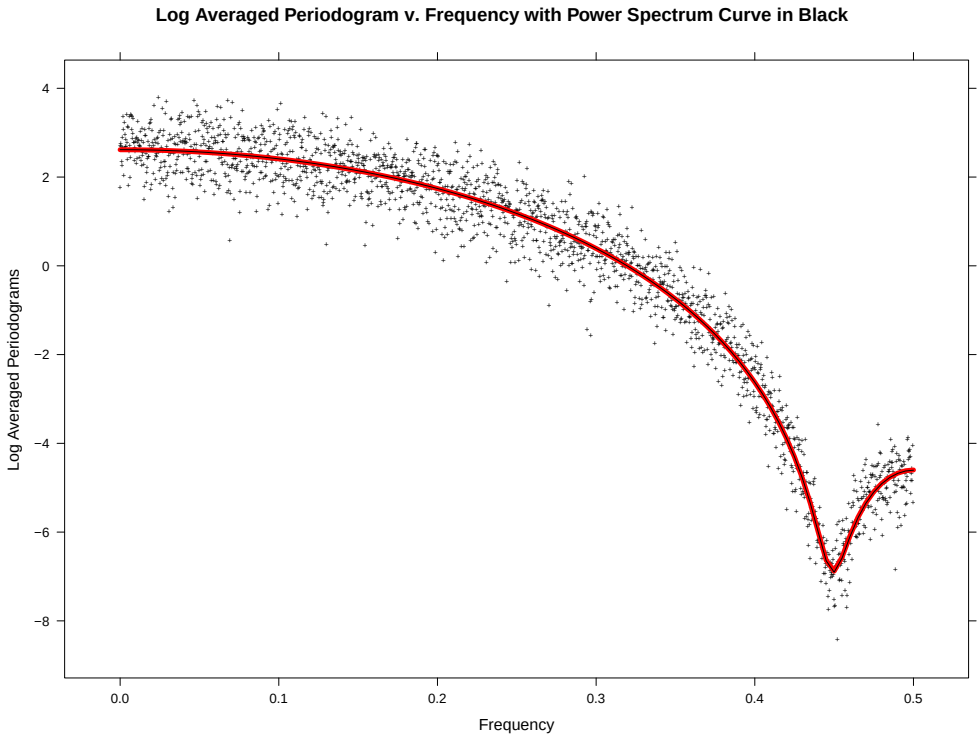

Fig. 4.27. MA(q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

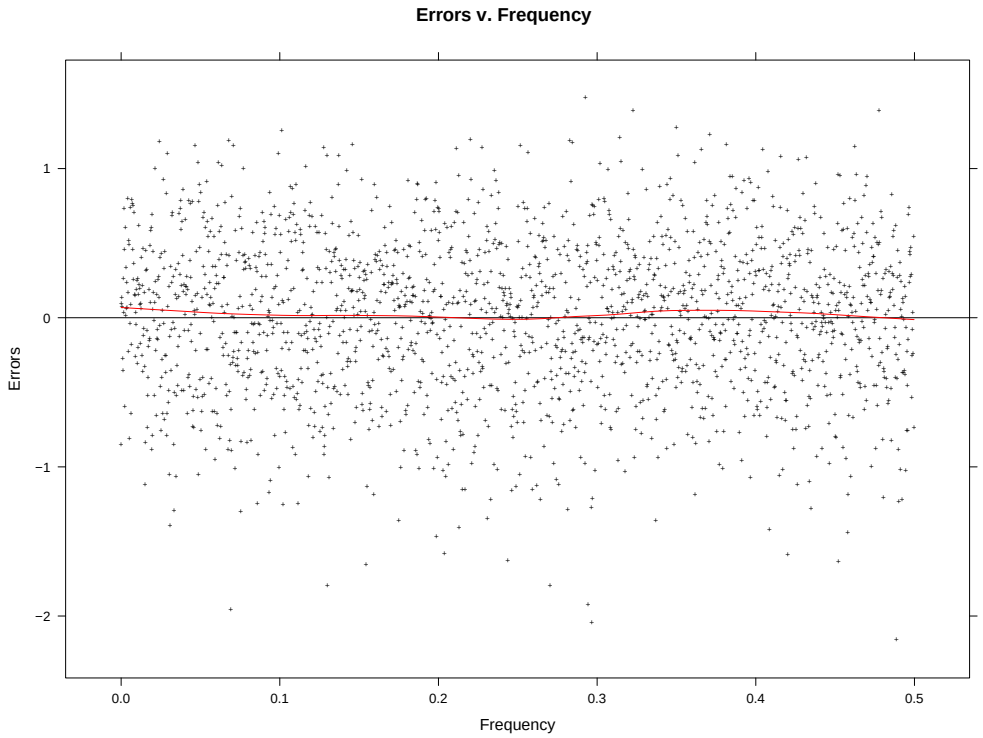

Fig. 4.28. MA(q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

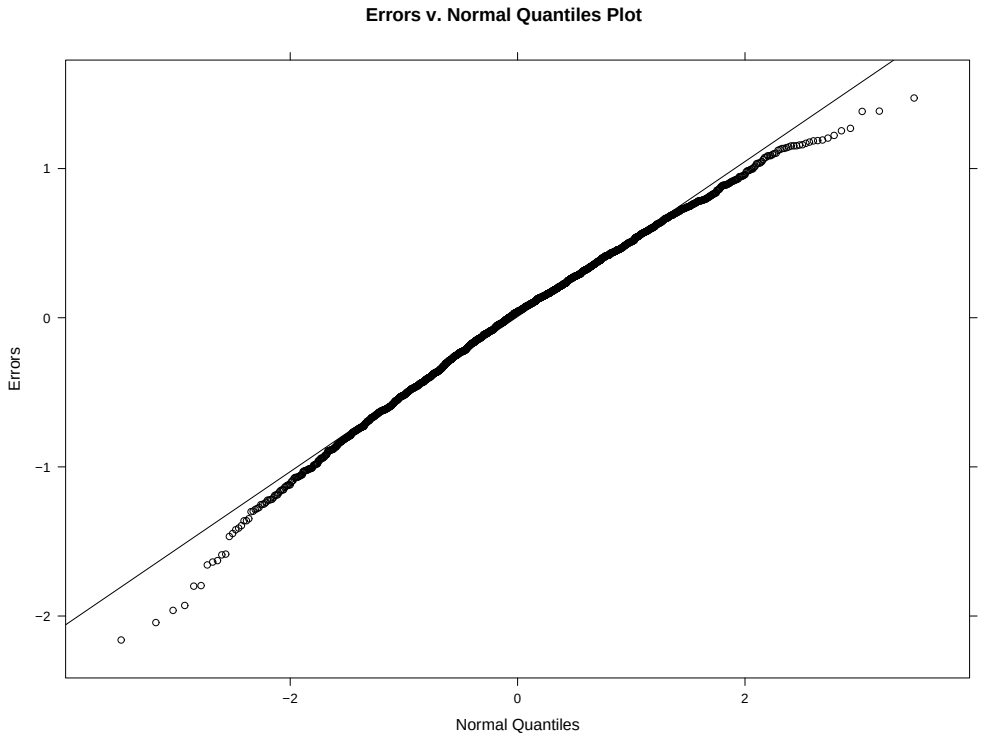

Fig. 4.29. MA(q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

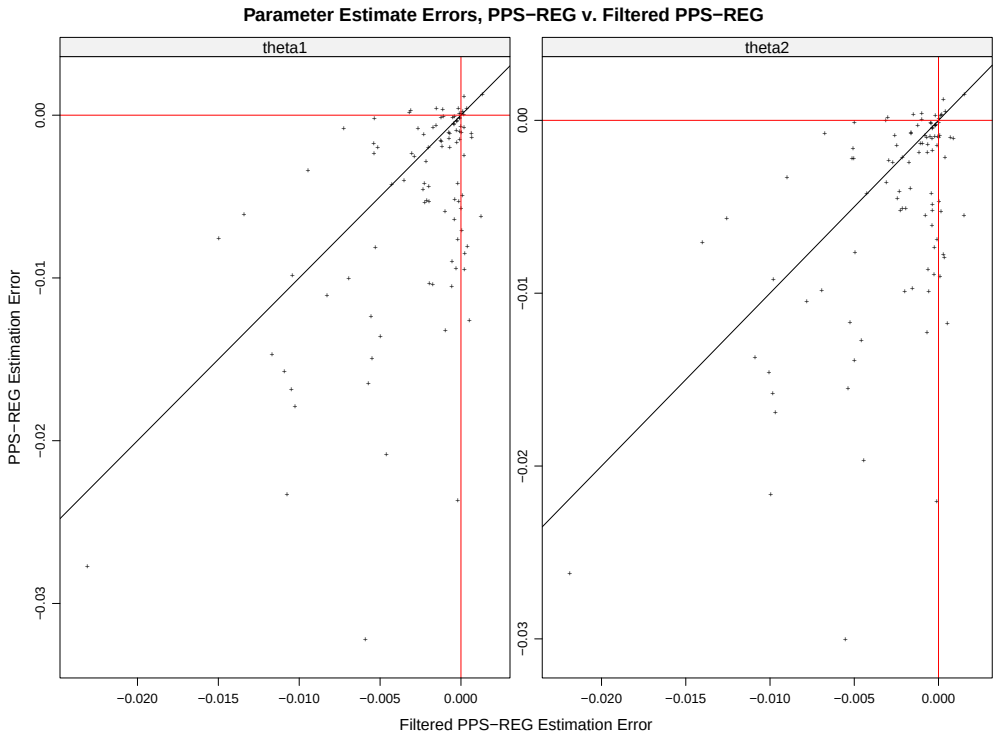

Fig. 4.30. MA(q = 2) Scatterplot of PPS-REG Estimation Error v. Filtered Estimation Error by parameter, θ_1 or θ_2

The worst PPS-REG errors are improved by filtering.

To assess the normality of our estimates by method, QQ plots were constructed:

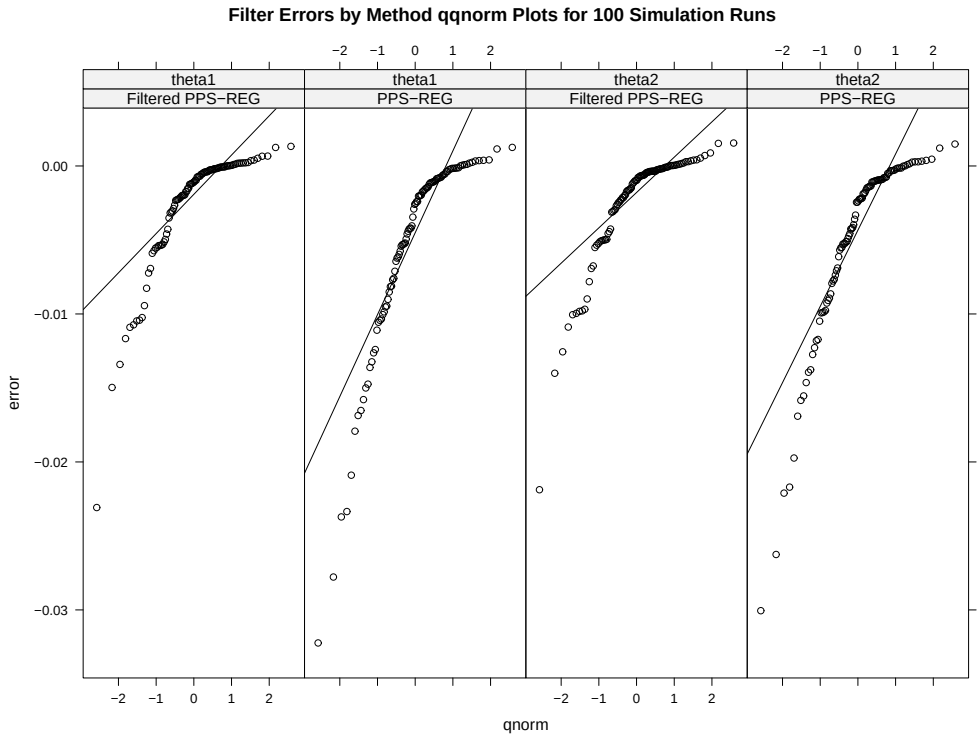

Fig. 4.31. MA(q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered, and Parameter, θ_1 or θ_2

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

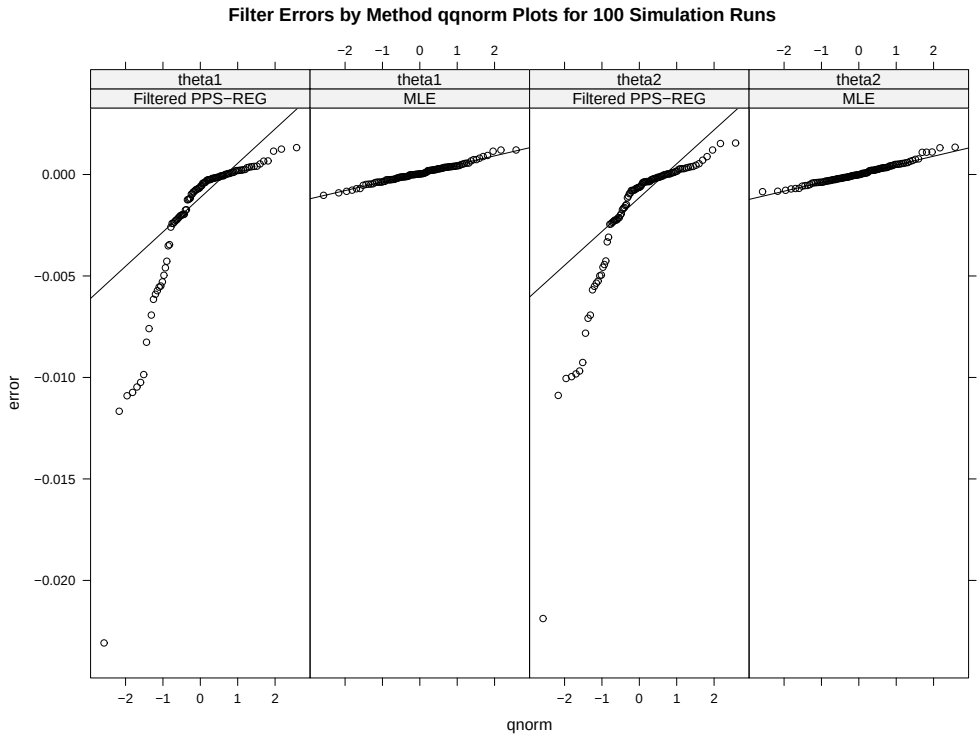

Fig. 4.32. MA(q = 2) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE, and Parameter, θ_1 or θ_2

Filtered PPS-REG exhibits a left tail for $\hat{\theta}_1$ and $\hat{\theta}_2$, but is an improvement from a heavier PPS-REG left tail.

4.1.8 ARMA(p = 2, q = 2)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

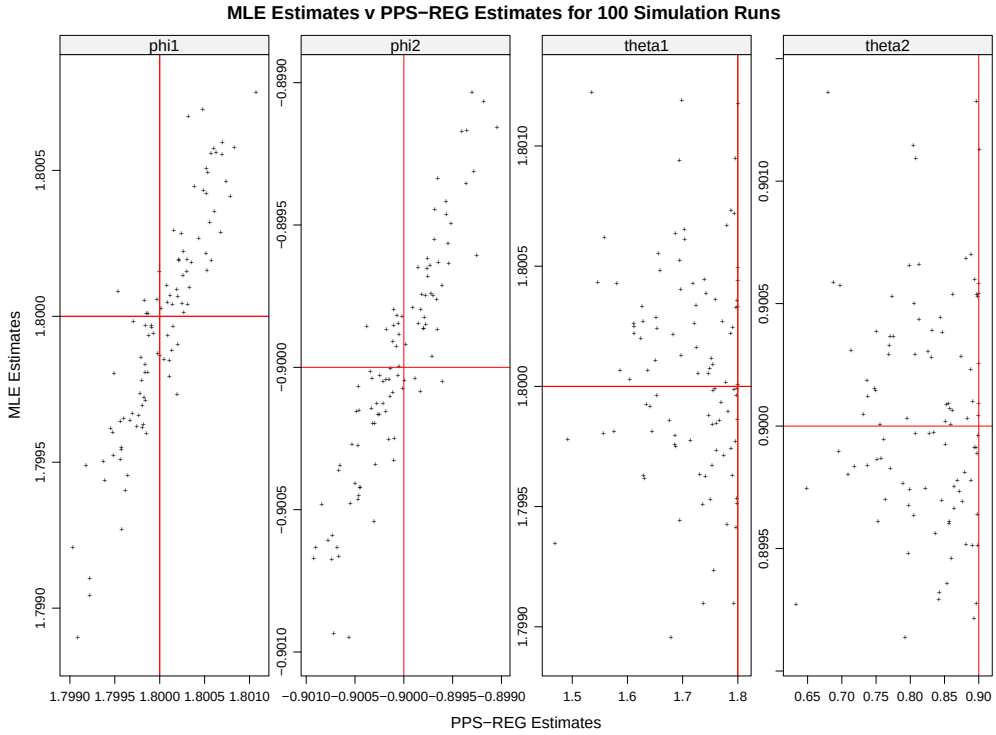

Fig. 4.33. ARMA(p = 2, q = 2) Scatterplot of MLE estimates v. PPS-REG estimates by ϕ_1, ϕ_2, θ_1, or θ_2

The results are comparable for the ϕ_1 and ϕ_2 parameters, but exhibit more variability and a slight bias towards the origin for θ_1 and θ_2 for PPS-REG estimation compared to MLE estimates.

To assess the normality of our estimates by method, QQ plots were constructed:

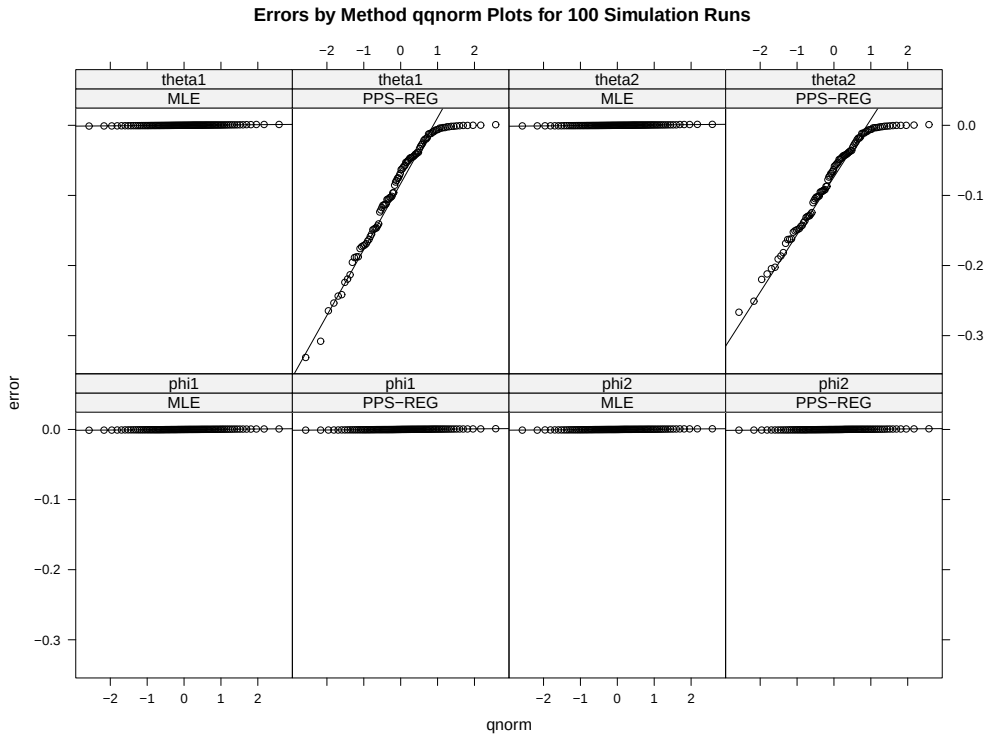

Fig. 4.34. ARMA(p = 2, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and parameter, ϕ_1, ϕ_2, θ_1, or θ_2

PPS-REG exhibits a heavy left tail for $\hat{\theta}_1$ and $\hat{\theta}_2$ on a little more than 10% of the simulation runs, but otherwise is normally distributed. The MLE estimates across differing simulations yield parameter estimates that are normally distributed about the simulated values.

AutoRegressive (AR) Component Filtration

The $ARMA(p=2,q=2)$ time series model is:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

The $MA(q=2)$ filtered data time series model is:

$$x_t^* = x_t - \hat{\phi}_1 x_{t-1} - \hat{\phi}_2 x_{t-2} = (\phi_1 - \hat{\phi}_1) x_{t-1} + (\phi_2 - \hat{\phi}_2) x_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

$$= \varepsilon_{\phi_1} x_{t-1} + \varepsilon_{\phi_2} x_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t \approx \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

The PPS-REG log power spectrum estimation on the $MA(q=2)$ filtered data yielded the following log averaged periodogram versus frequency results:

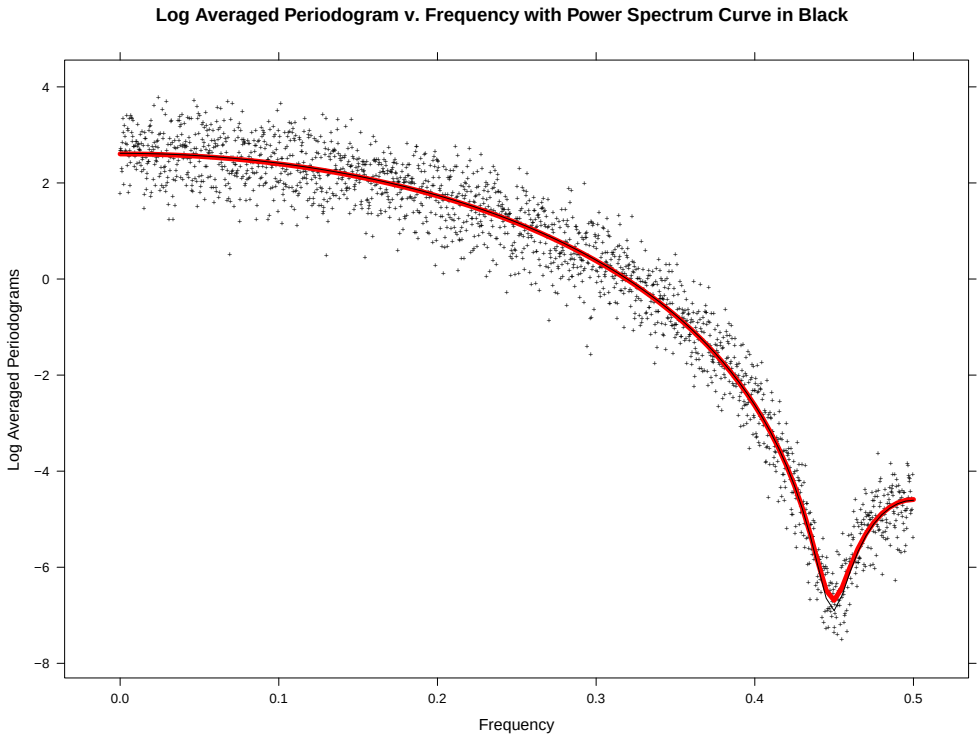

Fig. 4.35. MA(q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

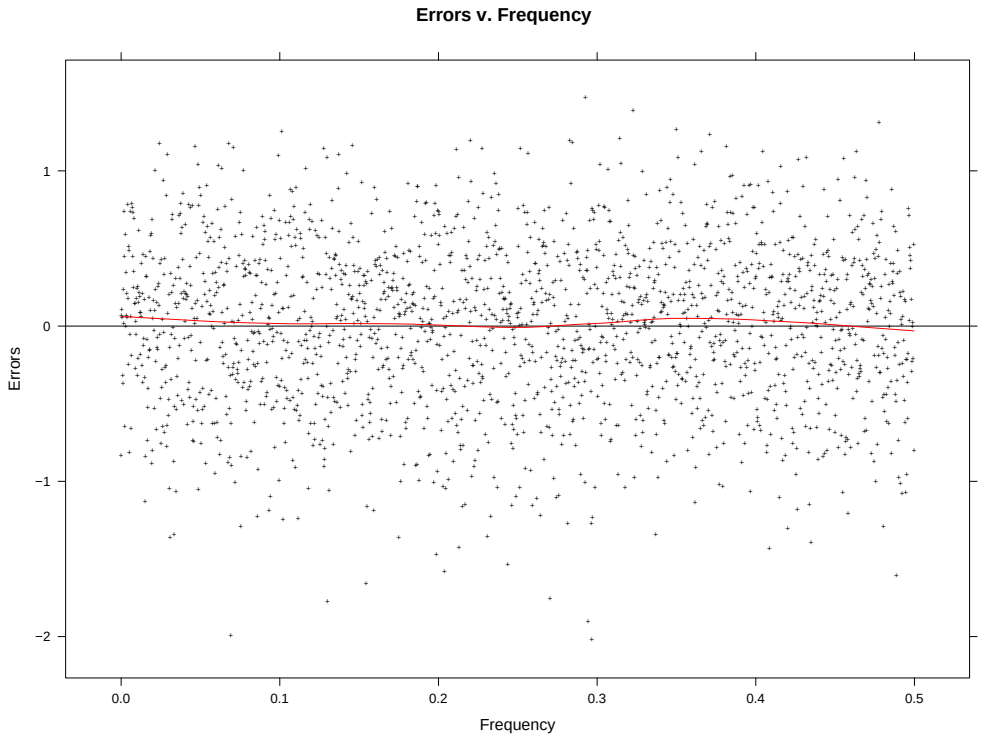

Fig. 4.36. MA(q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

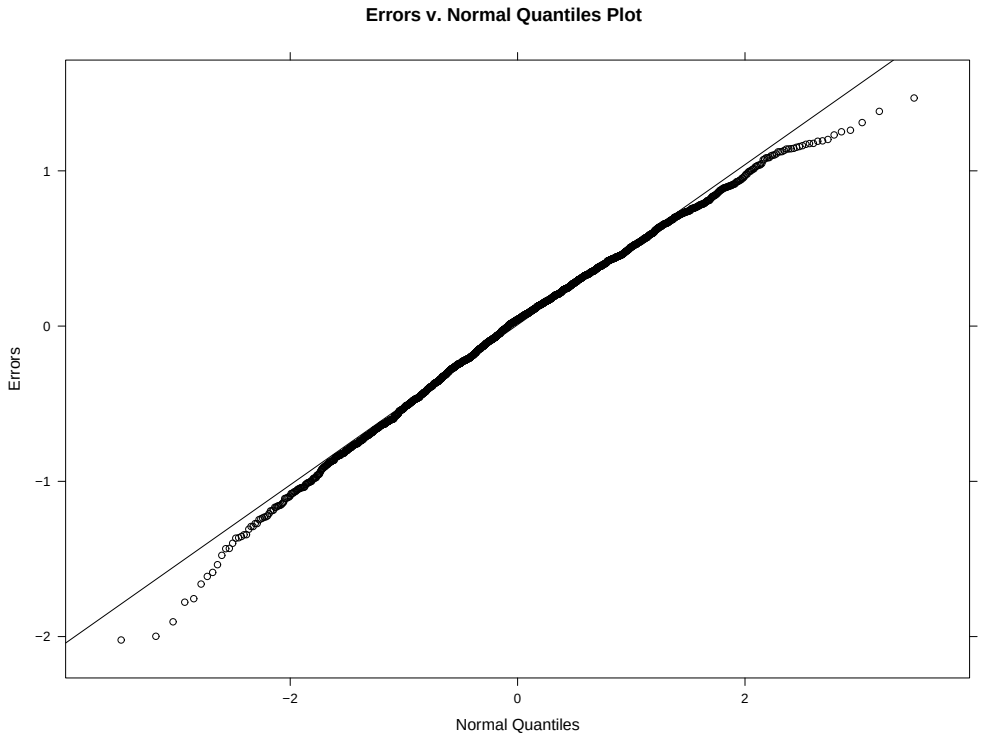

Fig. 4.37. MA(q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

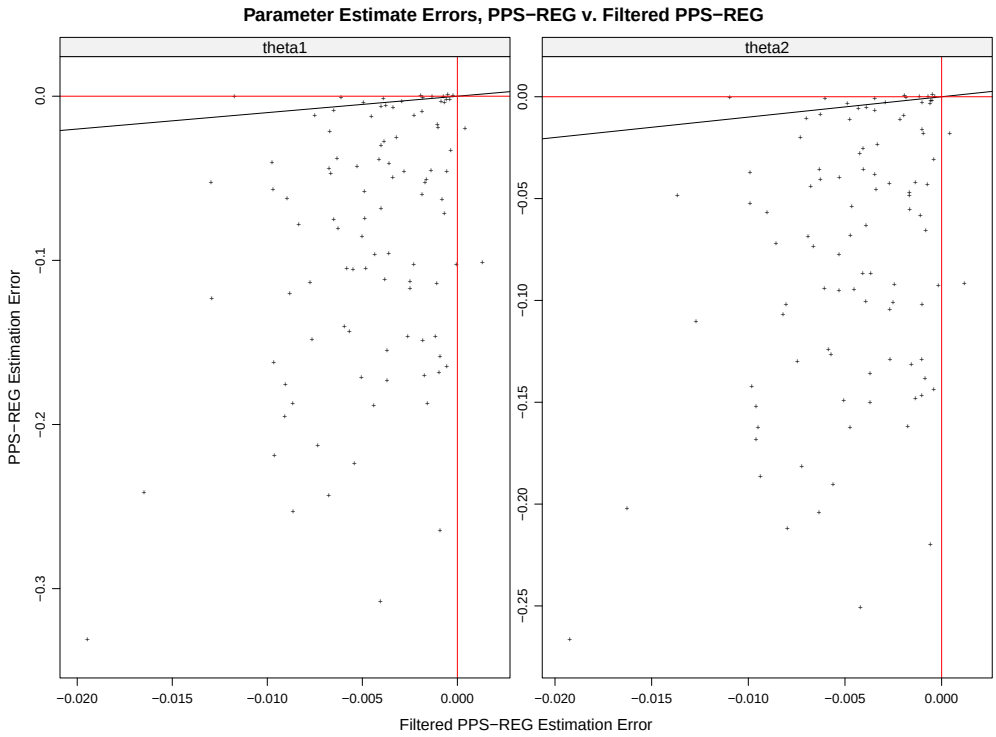

Fig. 4.38. MA(q = 2) Scatterplot of PPS-REG Estimation Error v. Filtered Estimation Error by parameter, θ_1 or θ_2

The worst PPS-REG errors are improved by filtering.

To assess the normality of our estimates by method, QQ plots were constructed:

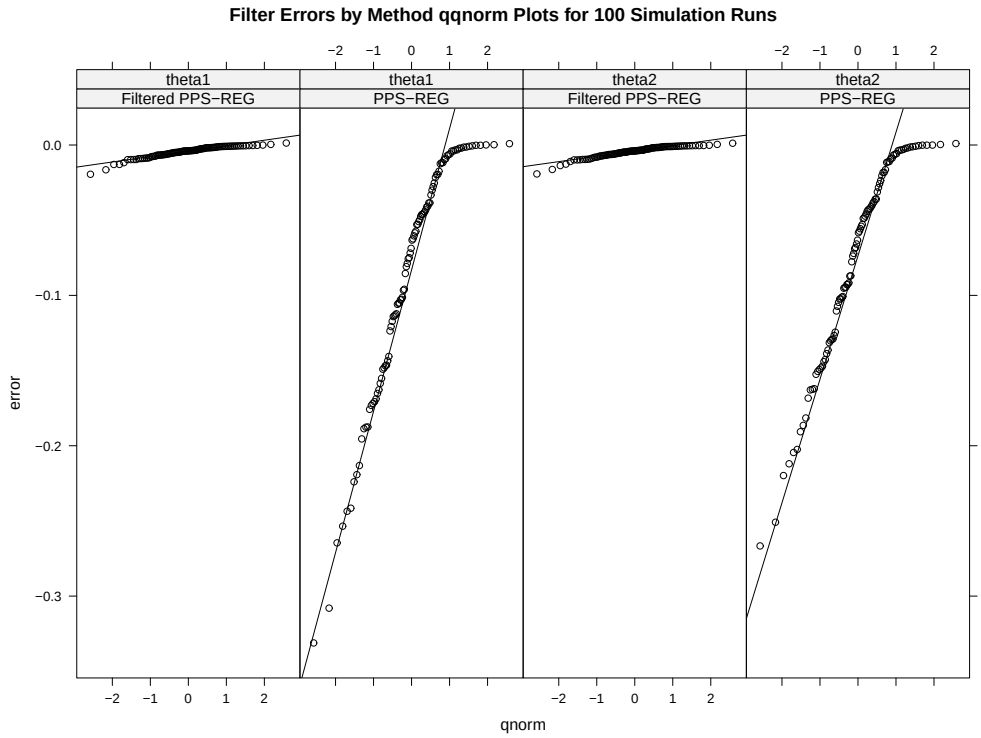

Fig. 4.39. MA(q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered, and Parameter, θ_1 or θ_2

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

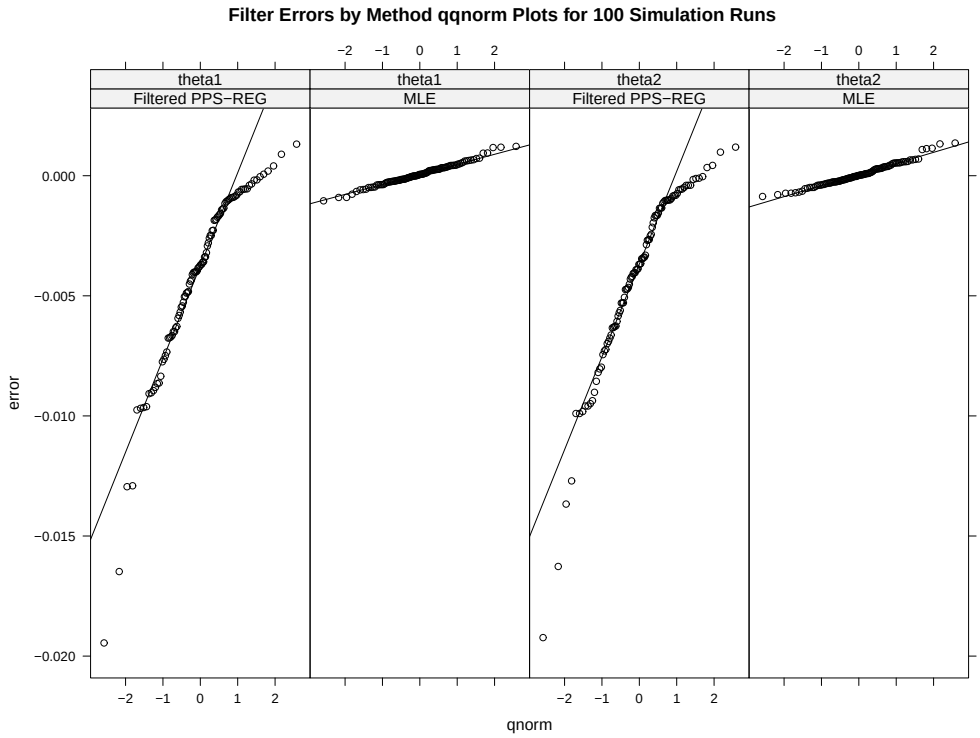

Fig. 4.40. MA(q = 2) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE, and Parameter, θ_1 or θ_2

Filtered PPS-REG exhibits a left tail for $\hat{\theta}_1$ and $\hat{\theta}_2$, but is an improvement from a heavier PPS-REG left tail.

4.2 FARMA Models

4.2.1 ARFIMA(p = 0, d, q = 0)

PPS-REG parameter estimates were initialized to be $(d^{(0)}, \sigma^2_{(0)}) = (0, e^{\overline{y_k}})$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

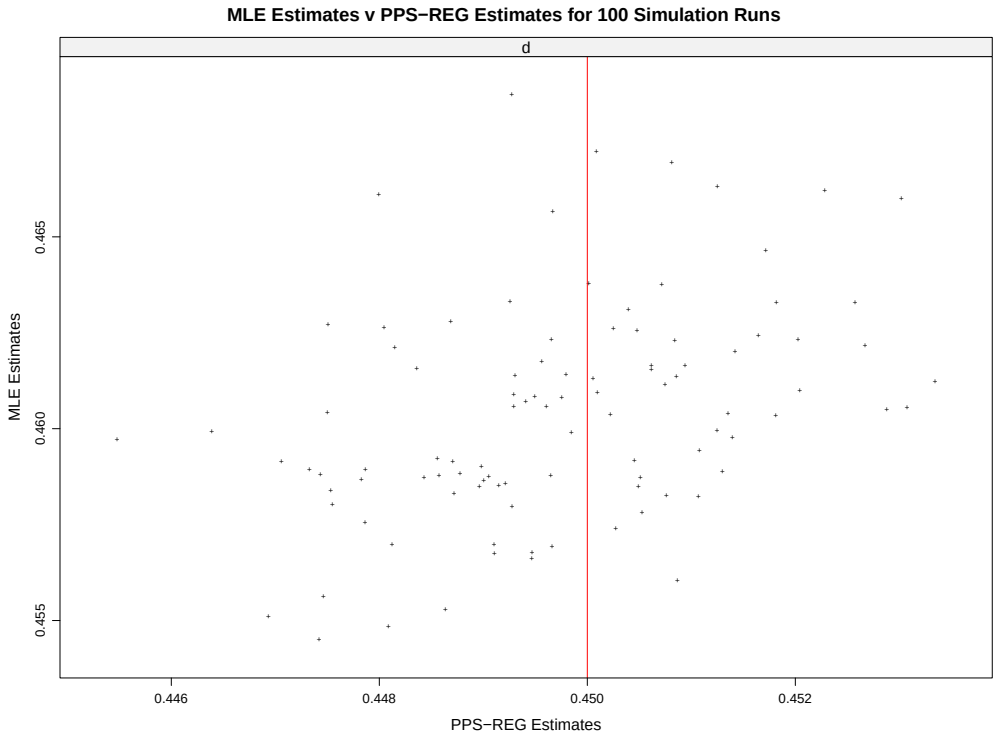

Fig. 4.41. FARMA(p = 0, d, q = 0) Scatterplot of MLE estimates of *d* v. PPS-REG estimates of *d*

The MLE estimation of *d* exhibits a slight overestimation bias while PPS-REG does not exhibit an evident bias. More importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of *d* indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

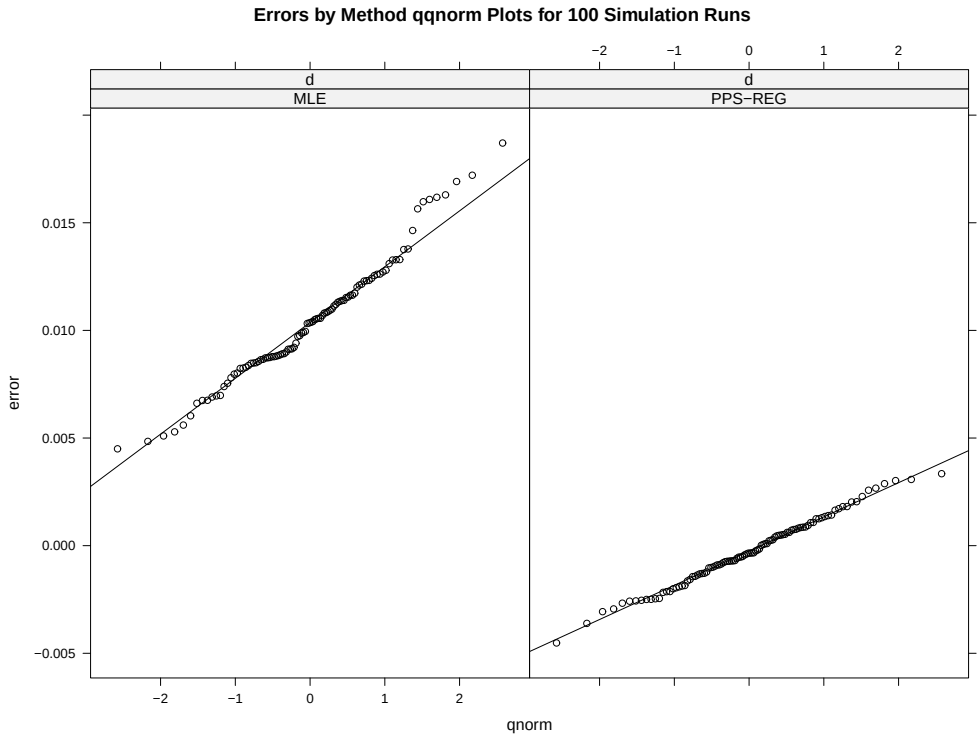

Fig. 4.42. FARMA(p = 0, d, q = 0) Normal QQ-Plots of Estimation Error of d by Estimation Method, PPS-REG v. MLE

Both estimation methods across differing simulations yield $\hat{d}$ estimates that are normally distributed about the simulated d value for PPS-REG and a slight overestimation bias for MLE.

4.2.2 ARFIMA(p = 1, d, q = 0)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

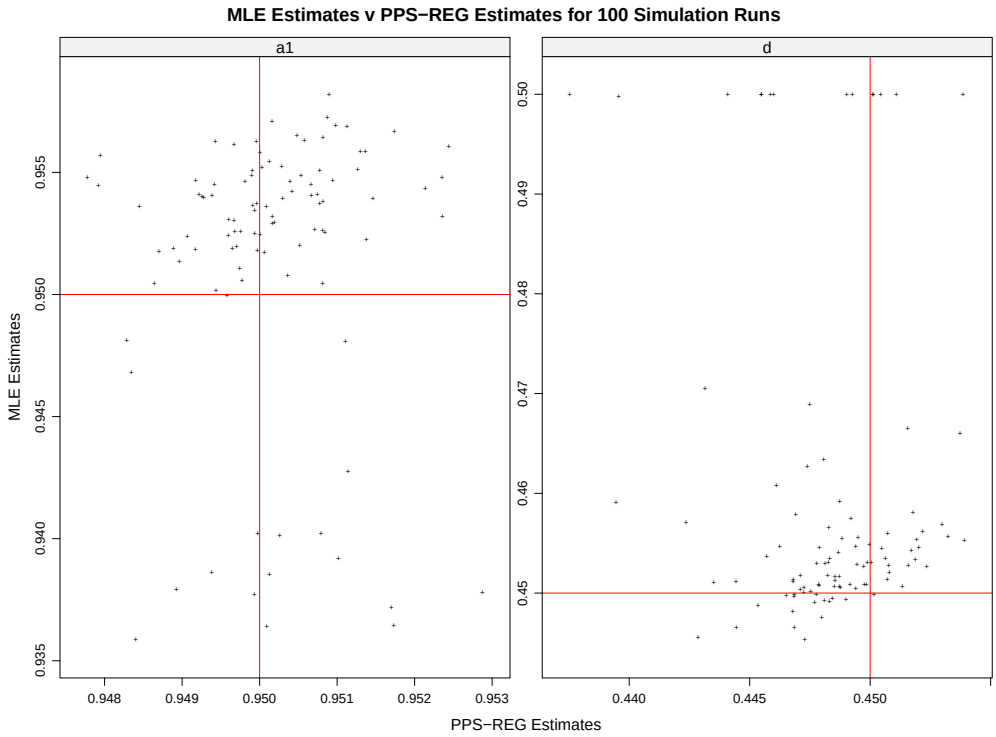

Fig. 4.43. FARMA(p = 1, d, q = 0) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ or d

The MLE estimation of d exhibits a slight overestimation bias while PPS-REG does not exhibit an evident bias. More importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of ϕ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

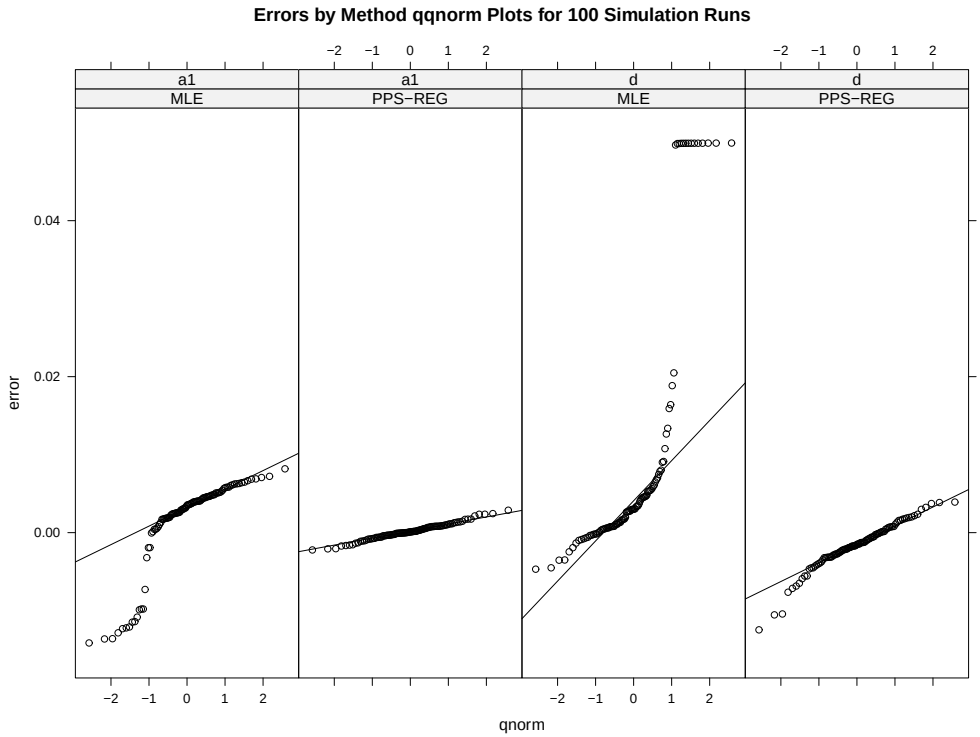

Fig. 4.44. FARMA(p = 1, d, q = 0) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ or d

Both estimation methods across differing simulations yield parameter estimates that are normally distributed about their simulated values, but MLE estimation of d possesses a slight overestimation bias.

4.2.3 ARFIMA(p = 2, d, q = 0)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

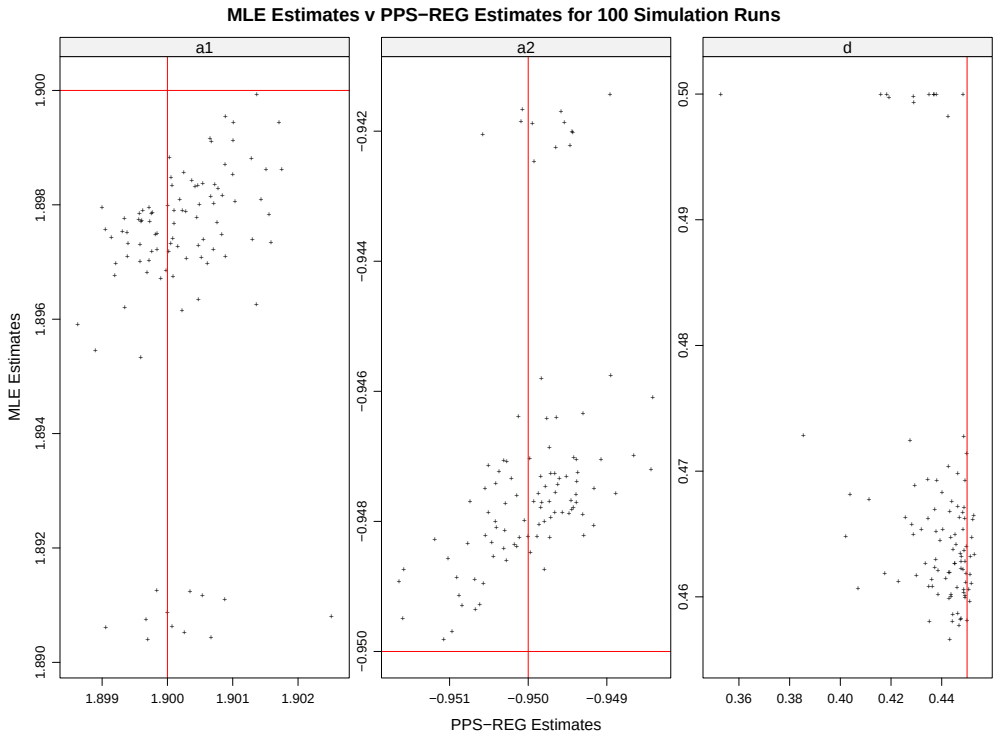

Fig. 4.45. FARMA(p = 2, d, q = 0) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ_1, ϕ_2, or d

MLE estimation of ϕ_1 and ϕ_2 exhibits a slight underestimation bias in magnitude while d exhibits a slight overestimation bias. PPS-REG does not exhibit an evident bias in ϕ_1 or ϕ_2, but exhibits higher underestimation bias in d and greater variability than MLE estimates of d.

To assess the normality of our estimates by method, QQ plots were constructed:

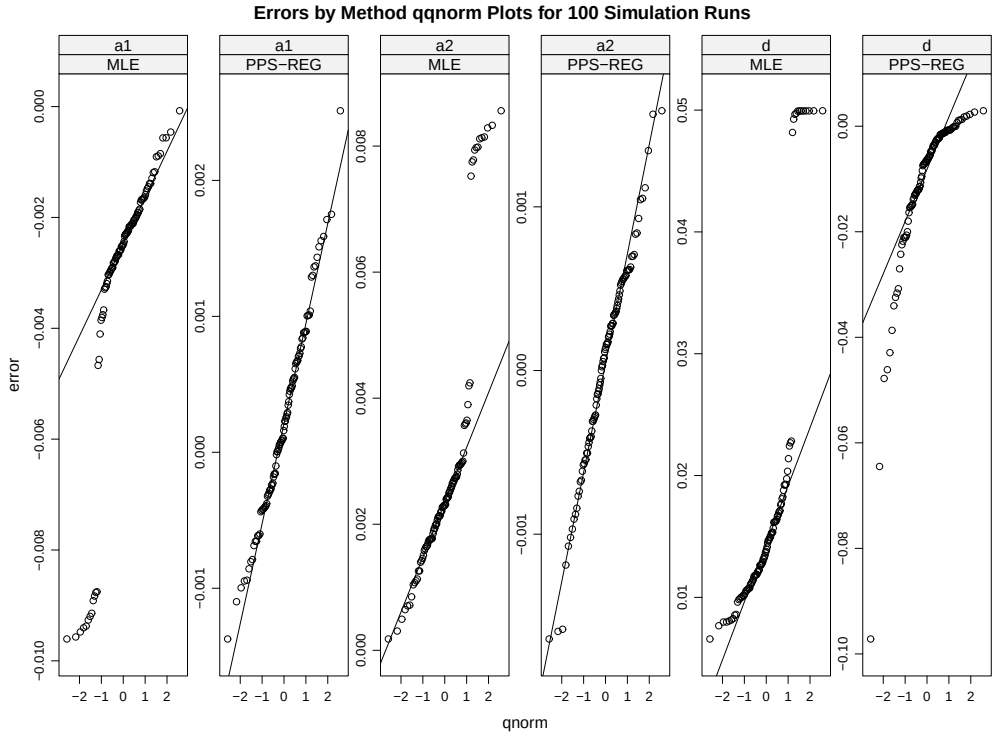

Fig. 4.46. FARMA(p = 1, d, q = 0) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ_1, ϕ_2, or d

PPS-REG estimation method across differing simulations yield parameter estimates of ϕ_1 and ϕ_2 that are normally distributed about their simulated values, but MLE estimation possesses heavy tails towards the value of 0. PPS-REG estimation of d exhibits a heavy left tail and underestimation bias, while MLE estimation of d exhibits overestimation bias and a heavy right tail.

4.2.4 ARFIMA(p = 0, d, q = 1)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

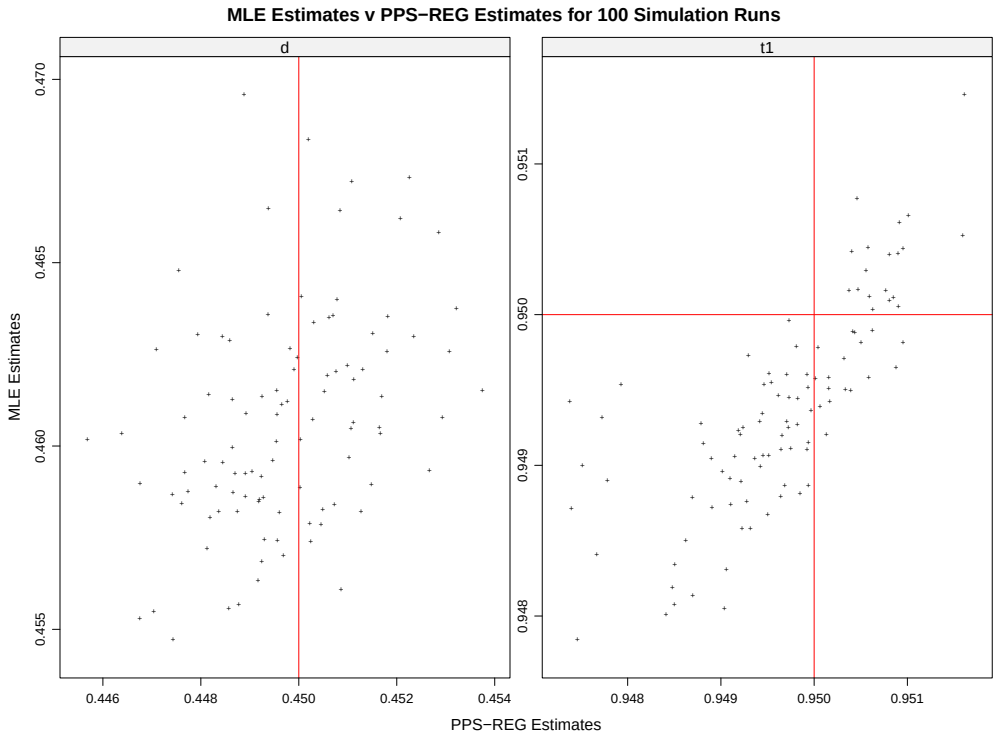

Fig. 4.47. FARMA(p = 0, d, q = 1) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, *d* or θ

The MLE estimation of *d* exhibits a slight overestimation bias while PPS-REG does not exhibit an evident bias. More importantly there is a positive linear relationship between the PPS-REG estimate and the MLE estimate of θ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

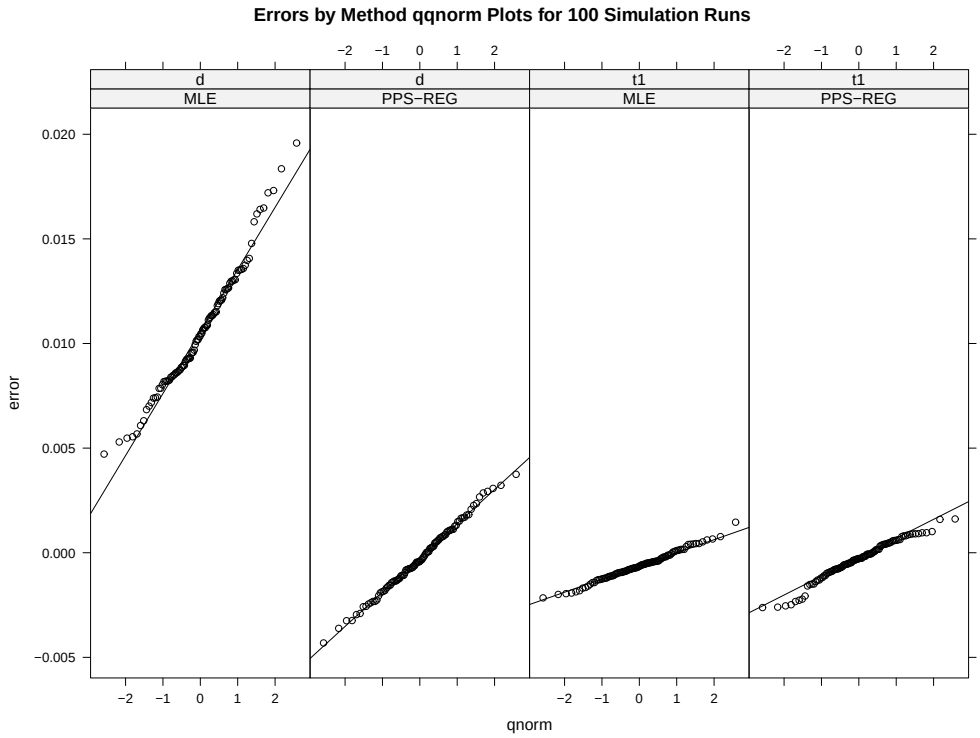

Fig. 4.48. FARMA(p = 0, d, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, d or θ

Both estimation methods across differing simulations yield parameter estimates that are normally distributed about their simulated values, but MLE estimation of d possesses a slight overestimation bias.

4.2.5 ARFIMA(p = 0, d, q = 2)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

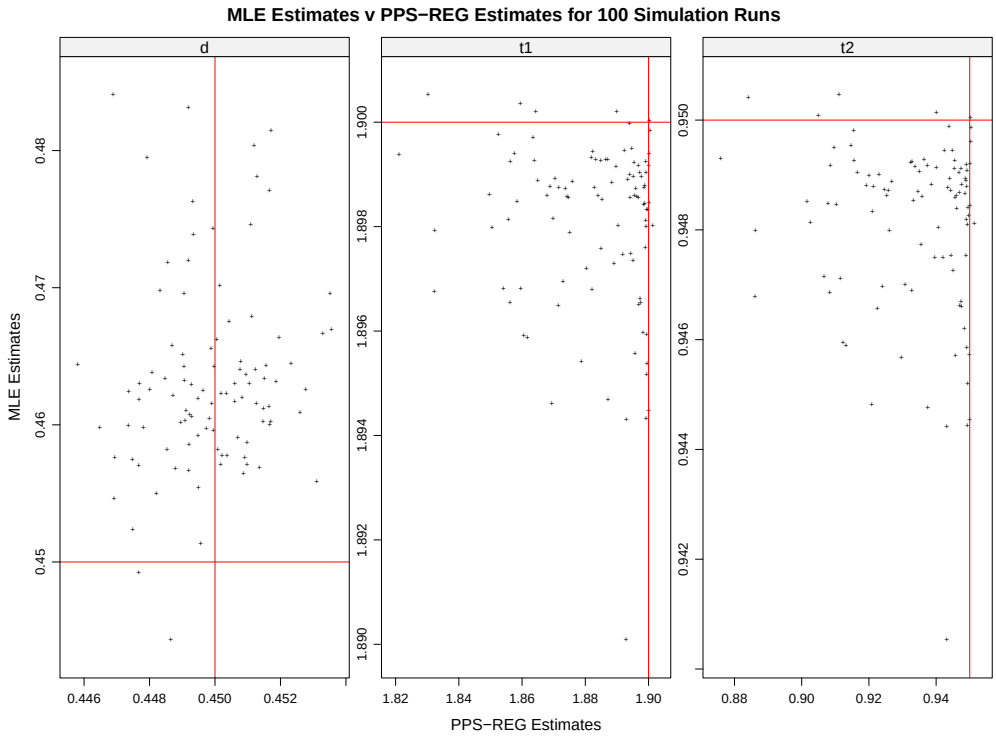

Fig. 4.49. FARMA(p = 0, d, q = 2) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, d, θ_1, or θ_2

MLE estimation of θ_1 and θ_2 exhibits a slight underestimation bias in magnitude while d exhibits a slight overestimation bias. PPS-REG exhibits a stronger underestimation bias in θ_1 and θ_2, but does not exhibit a bias in d.

To assess the normality of our estimates by method, QQ plots were constructed:

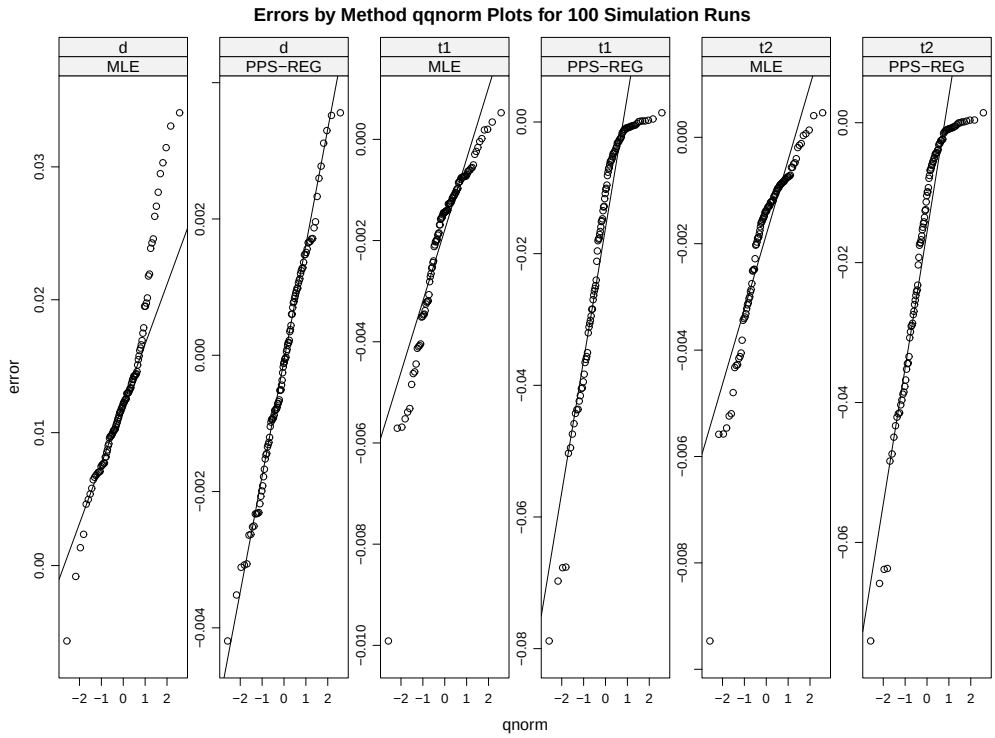

Fig. 4.50. FARMA(p = 0, d, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, d, θ_1, or θ_2

PPS-REG estimation method across differing simulations yield parameter estimates of θ_1 and θ_2 that possess heavy left tails, but MLE estimates are normally distributed about the simulated values of θ_1 and θ_2 respectively. PPS-REG estimation of d is normally distributed, while MLE estimation of d exhibits overestimation bias and a heavy right tail.

4.2.6 ARFIMA(p = 1, d, q = 1)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

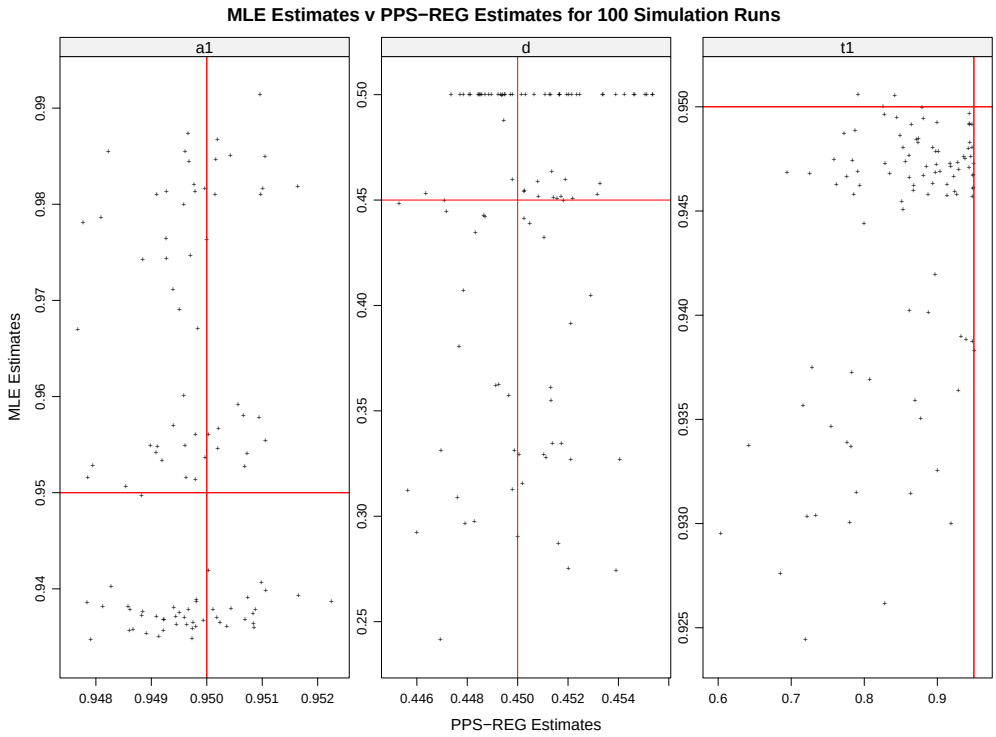

Fig. 4.51. FARMA(p = 1, d, q = 1) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ, d, or θ

PPS-REG estimates of d and ϕ possess far less variability than MLE estimates. MLE estimates of θ possess far less variability than PPS-REG estimates.

To assess the normality of our estimates by method, QQ plots were constructed:

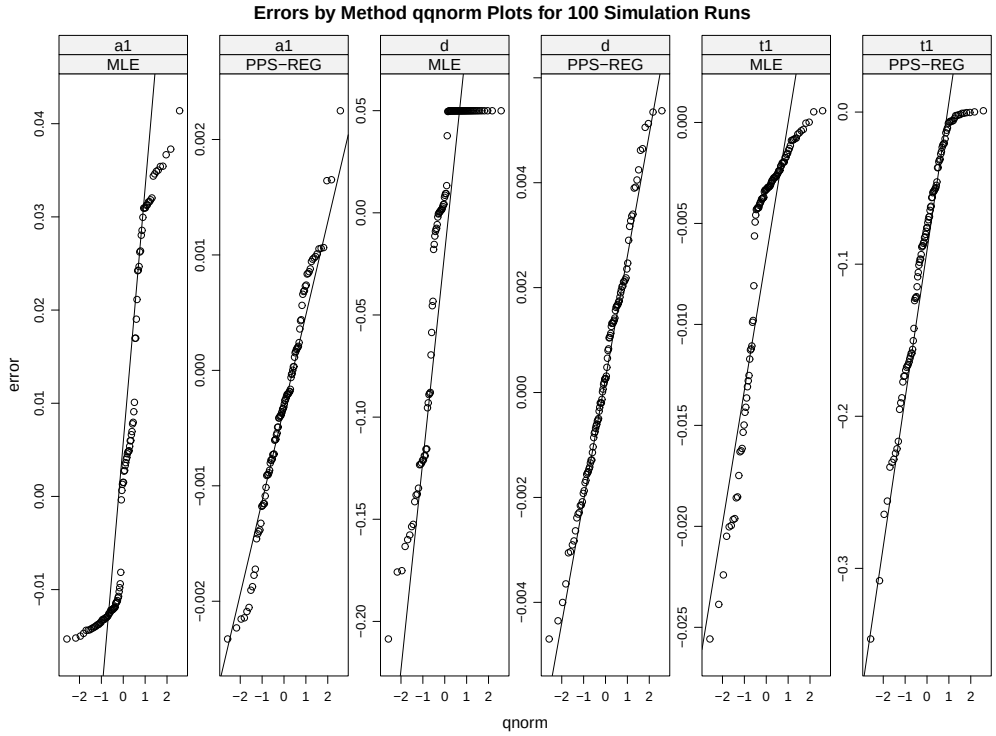

Fig. 4.52. FARMA(p = 1, d, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ, d, or θ

PPS-REG estimates are normally distributed for all parameters about their simulated values, except for $\hat{\theta}$, which is normally distributed with a bias towards 0. MLE estimates are normally distributed for all parameters about their simulated values, except for $\hat{d}$, which is normally distributed with a bias towards 0.

AutoRegressive (AR) Component Filtration

The $FARMA(p=1,d,q=1)$ time series model is:

$$(1-\phi B)(1-B)^d x_t = (1-B)^d x_t - \phi(1-B)^d x_{t-1} = (1-B)^d(x_t - \phi x_{t-1}) = (1+\theta B)\varepsilon_t$$

The $FARMA(p=0,d,q=1)$ filtered data time series model is:

$$(1-B)^d x_t^* = (1-B)^d(x_t - \hat{\phi} x_{t-1}) \approx (1+\theta B)\varepsilon_t$$

The PPS-REG log power spectrum estimation on the $FARMA(p=0,d,q=1)$ filtered data yielded the following log averaged periodogram versus frequency results:

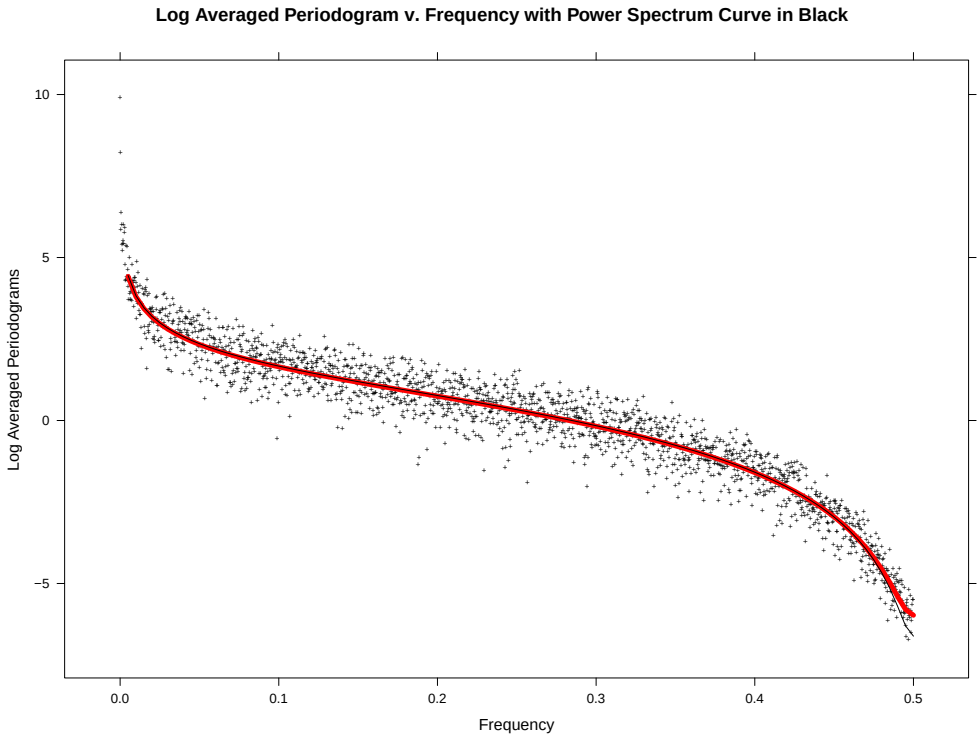

Fig. 4.53. FARMA(p = 0, d, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

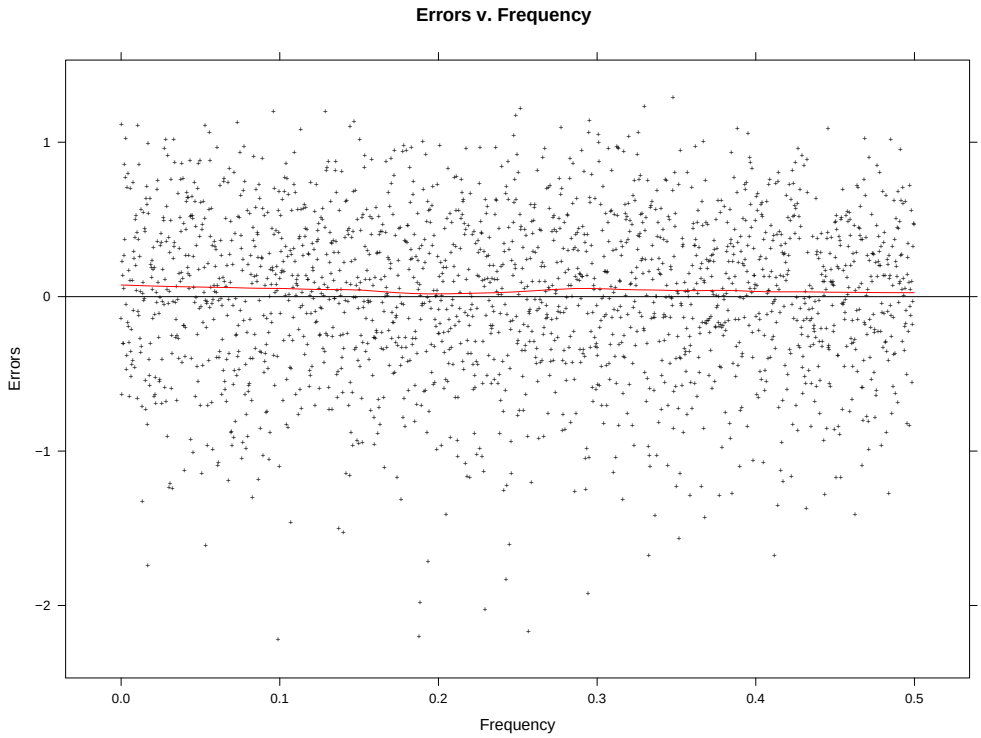

Fig. 4.54. FARMA(p = 0, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

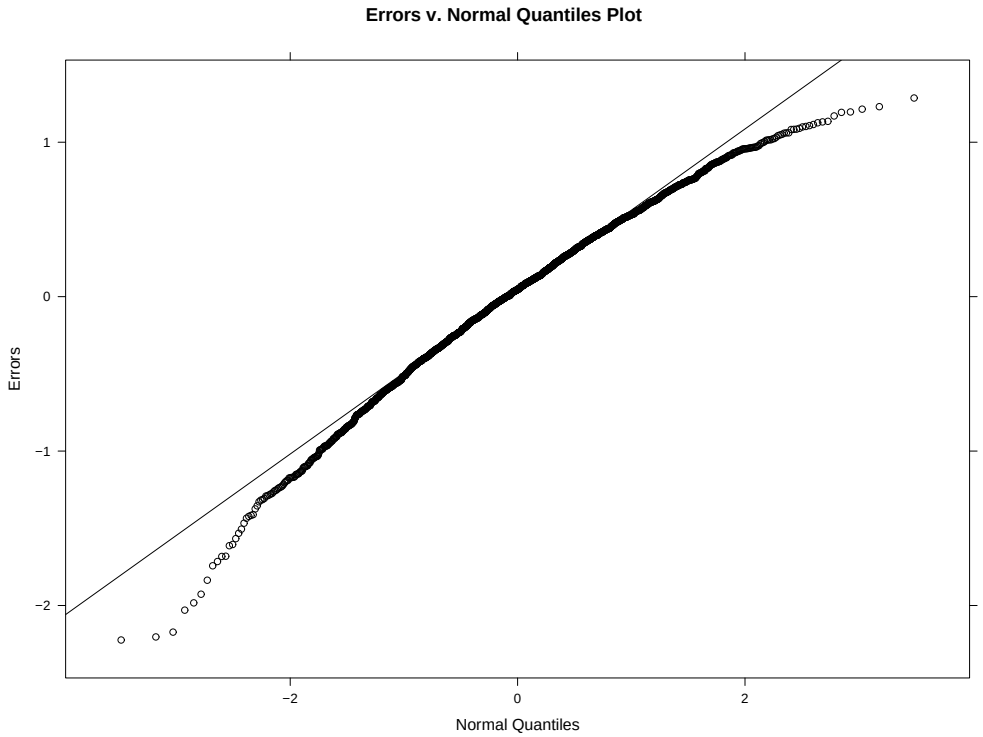

Fig. 4.55. FARMA(p = 0, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

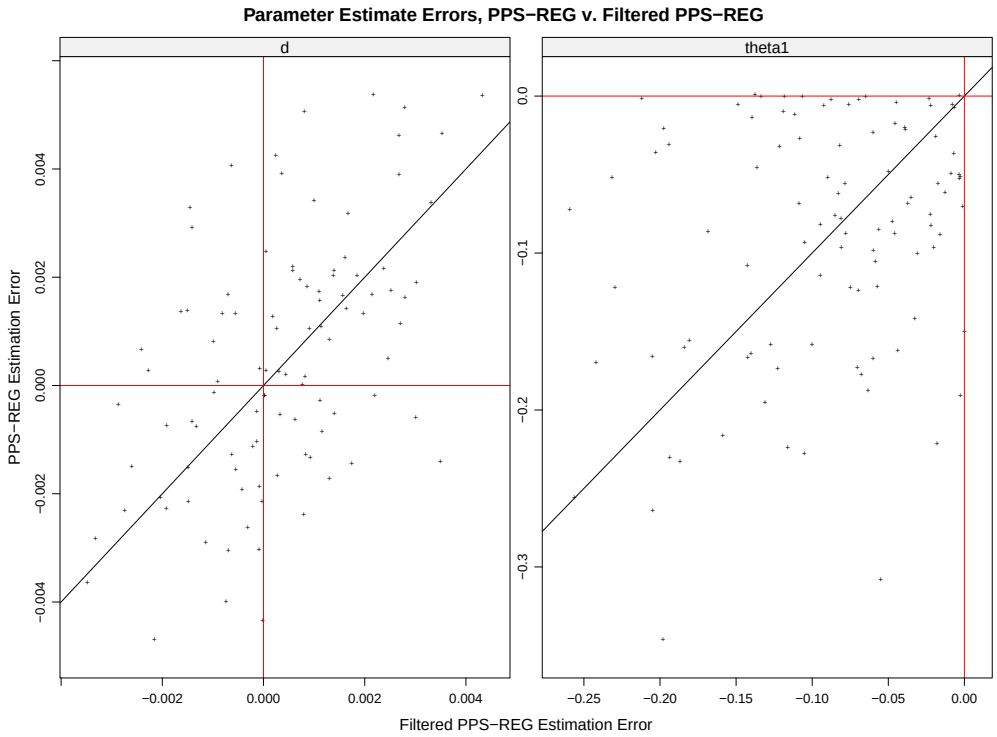

Fig. 4.56. FARMA(p = 0, d, q = 1) Scatterplot of PPS-REG Estimates v. PPS-REG Filtered Estimates by Parameter, *d* or θ

The worst PPS-REG errors are improved by filtering and more importantly there is a positive linear relationship between the PPS-REG error and the PPS-REG filtered error of θ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

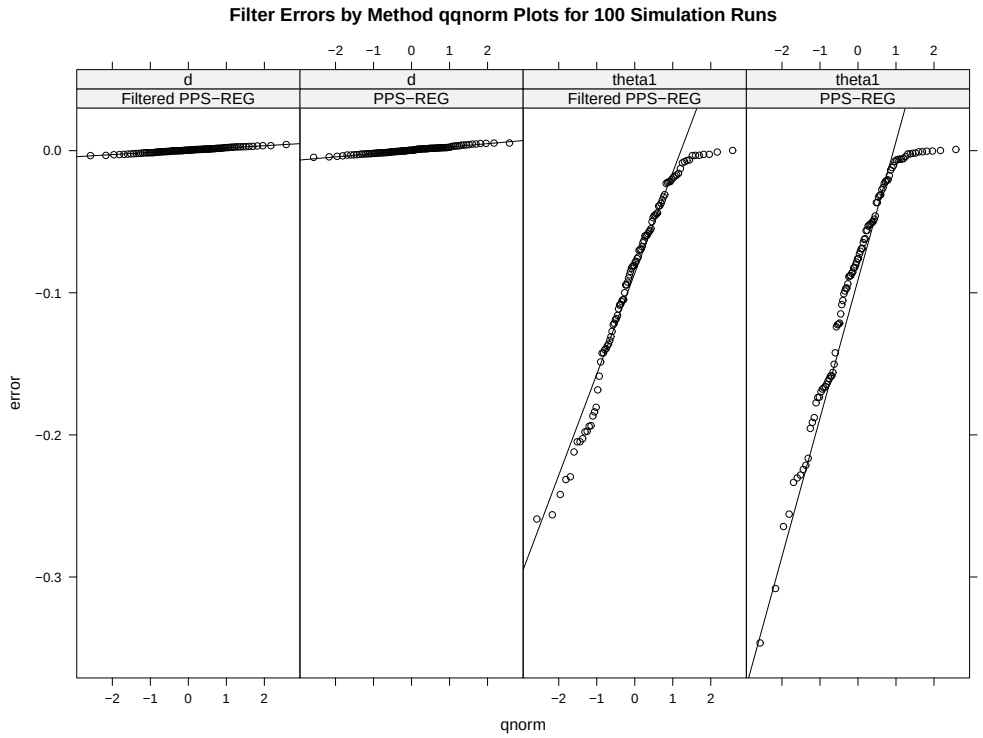

Fig. 4.57. FARMA(p = 0, d, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered, and Parameter, d or θ

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

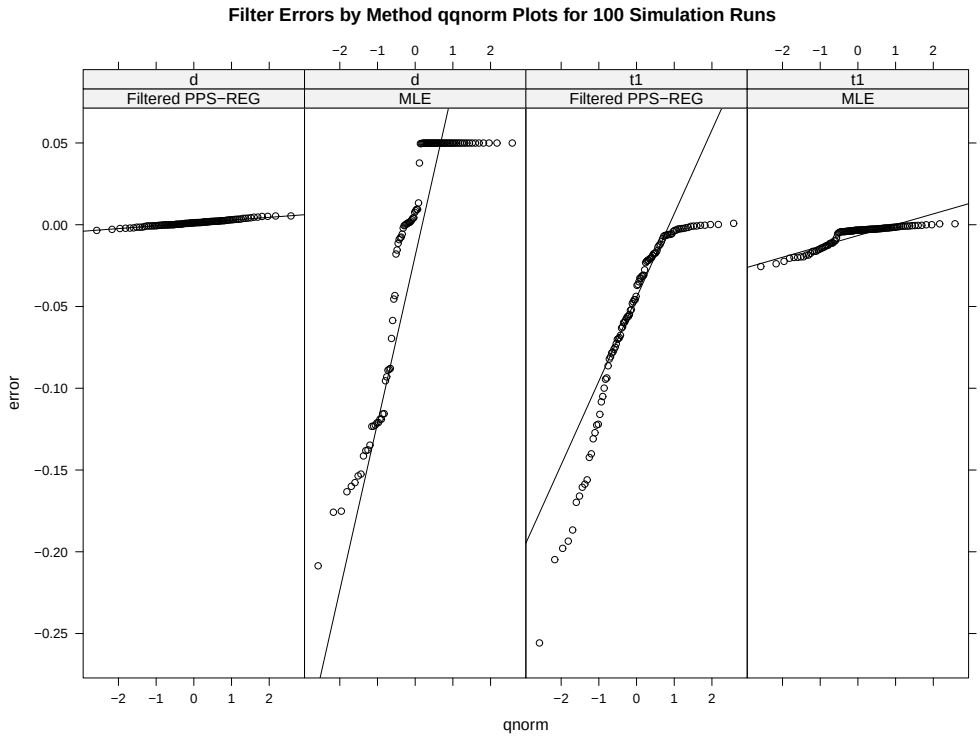

Fig. 4.58. FARMA(p = 0, d, q = 1) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE, and Parameter, d or θ

Filtered PPS-REG exhibits a left tail for $\hat{\theta}$, but otherwise is normally distributed and is an improvement from a heavier PPS-REG left tail.

4.2.7 ARFIMA(p = 2, d, q = 1)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

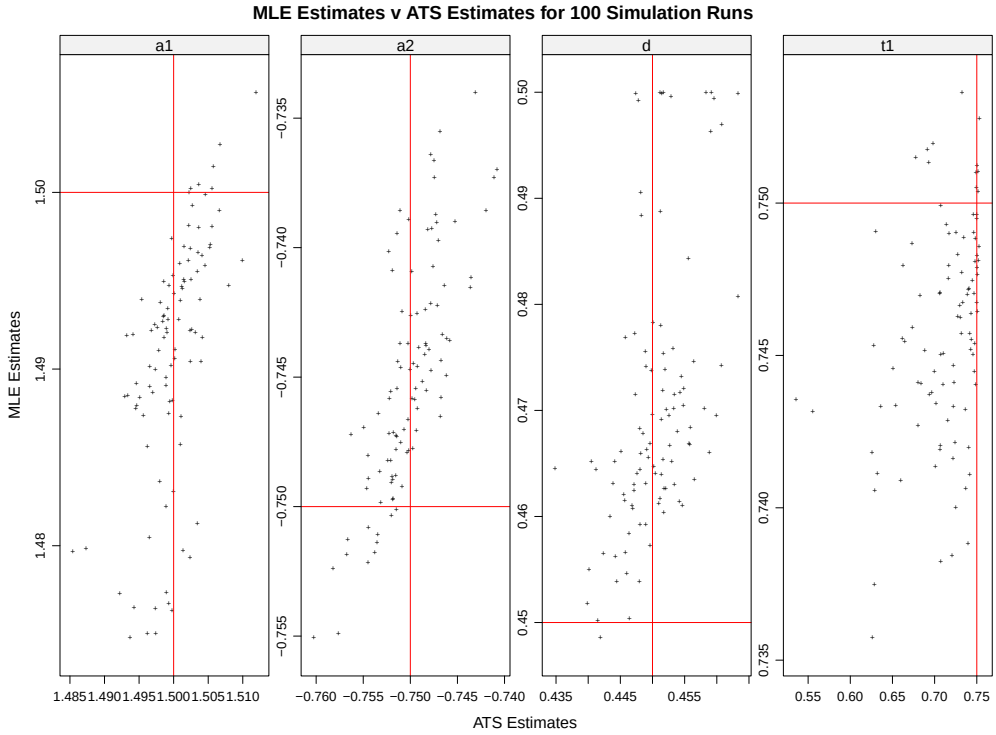

Fig. 4.59. FARMA(p = 2, d, q = 1) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ_1, ϕ_2, d, or θ

PPS-REG estimates of ϕ_1, ϕ_2, and d possess far less variability than MLE estimates. MLE estimates of θ possess far less variability than PPS-REG estimates.

To assess the normality of our estimates by method, QQ plots were constructed:

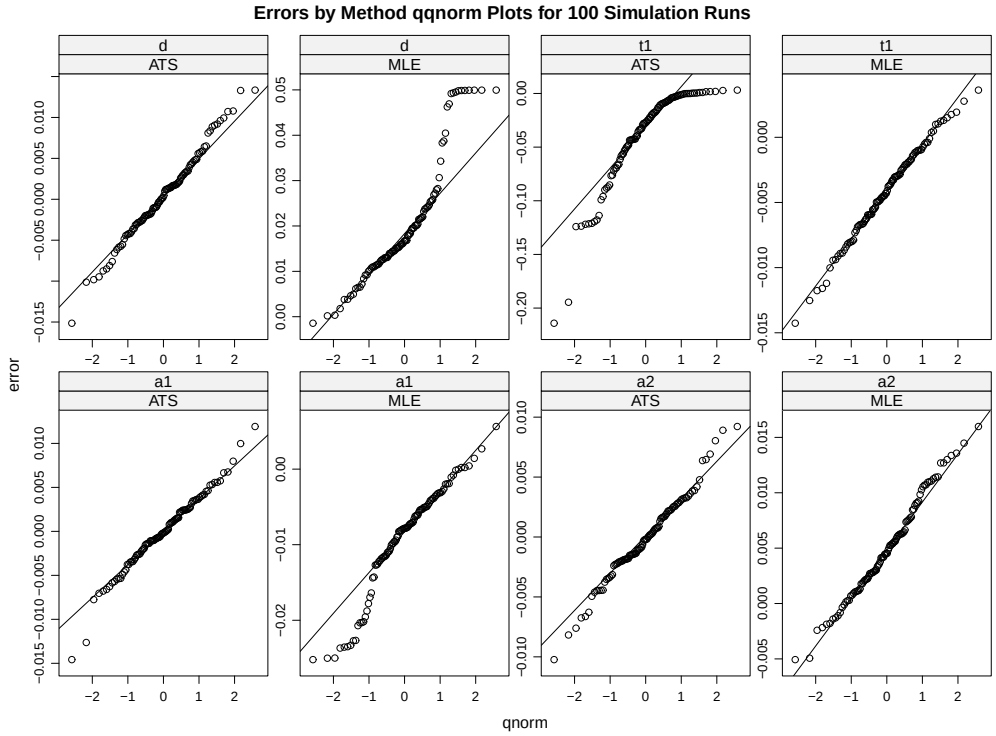

Fig. 4.60. FARMA(p = 2, d, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ_1, ϕ_2 d, or θ

PPS-REG estimates are normally distributed for all parameters about their simulated values, except for $\hat{\theta}$, which is normally distributed with a bias towards 0. MLE estimates are normally distributed for all parameters about their simulated values, except for $\hat{d}$, which is normally distributed with a bias towards 0.5.

AutoRegressive (AR) Component Filtration

The $FARMA(p=2,d,q=1)$ time series model is:

$$(1-\phi_1 B-\phi_2 B^2)(1-B)^d x_t = (1-B)^d(x_t-\phi_1 x_{t-1}-\phi_2 x_{t-2}) = (1+\theta B)\varepsilon_t$$

The $FARMA(p=0,d,q=1)$ filtered data time series model is:

$$(1-B)^d x_t^* = (1-B)^d(x_t-\hat{\phi}_1 x_{t-1}-\hat{\phi}_2 x_{t-2}) \approx (1+\theta B)\varepsilon_t$$

The PPS-REG log power spectrum estimation on the $FARMA(p=0,d,q=1)$ filtered data yielded the following log averaged periodogram versus frequency results:

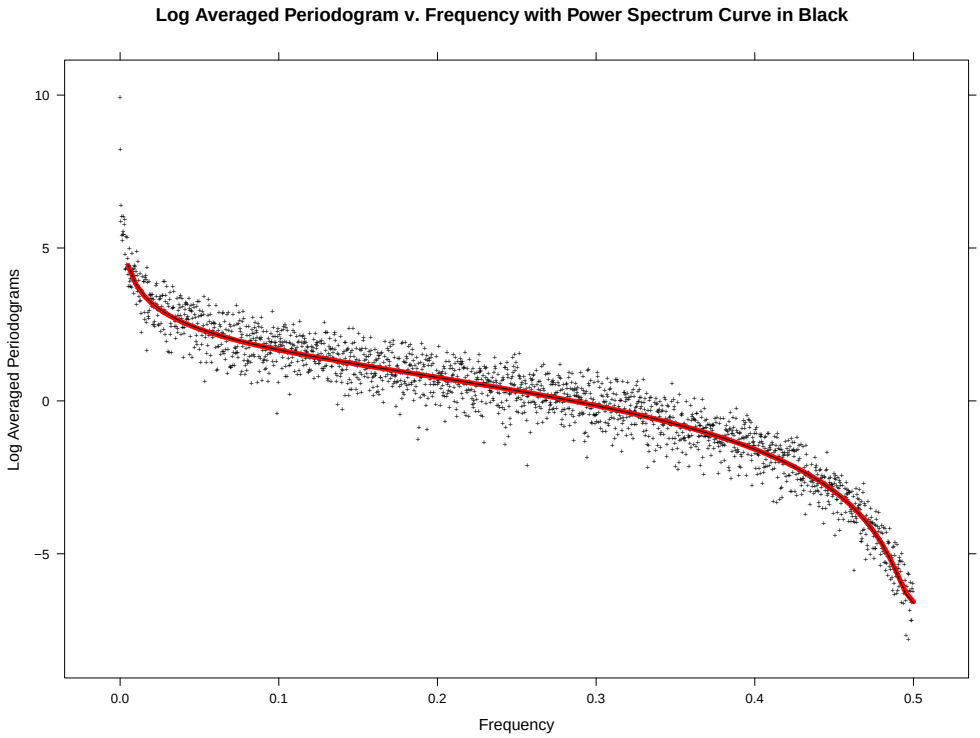

Fig. 4.61. FARMA(p = 0, d, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

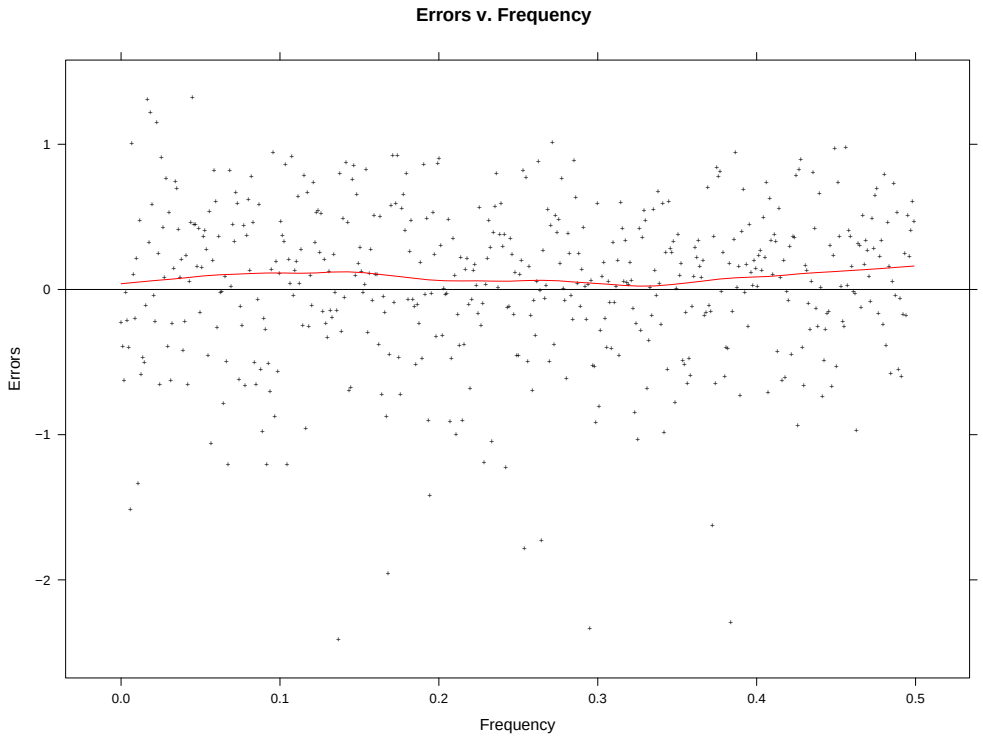

Fig. 4.62. FARMA(p = 0, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

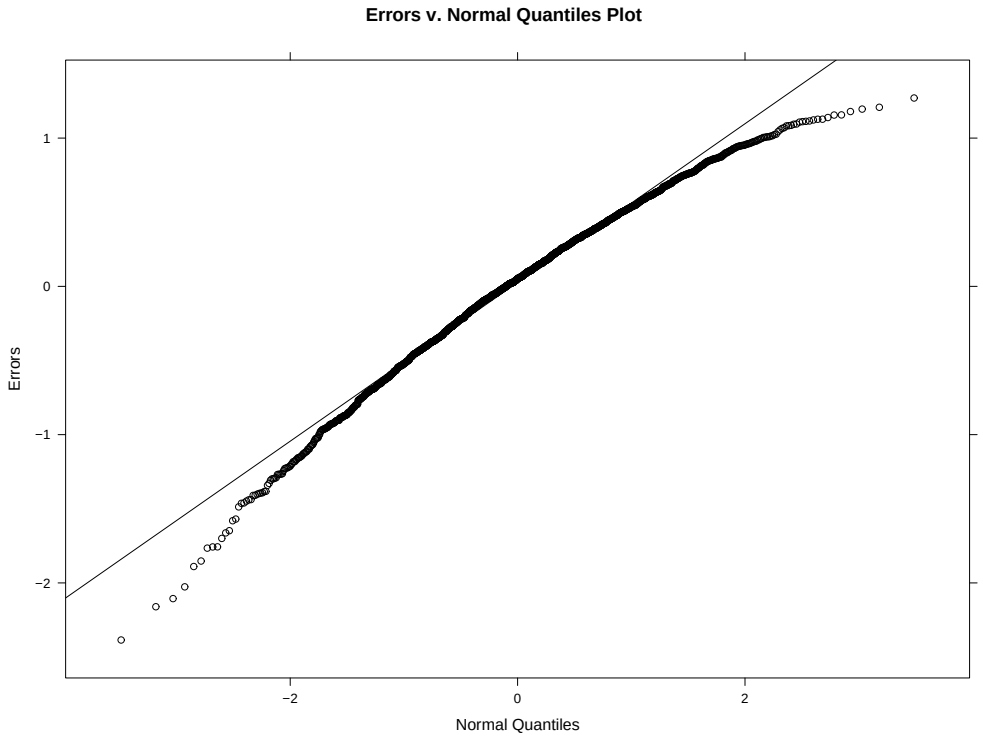

Fig. 4.63. FARMA(p = 0, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

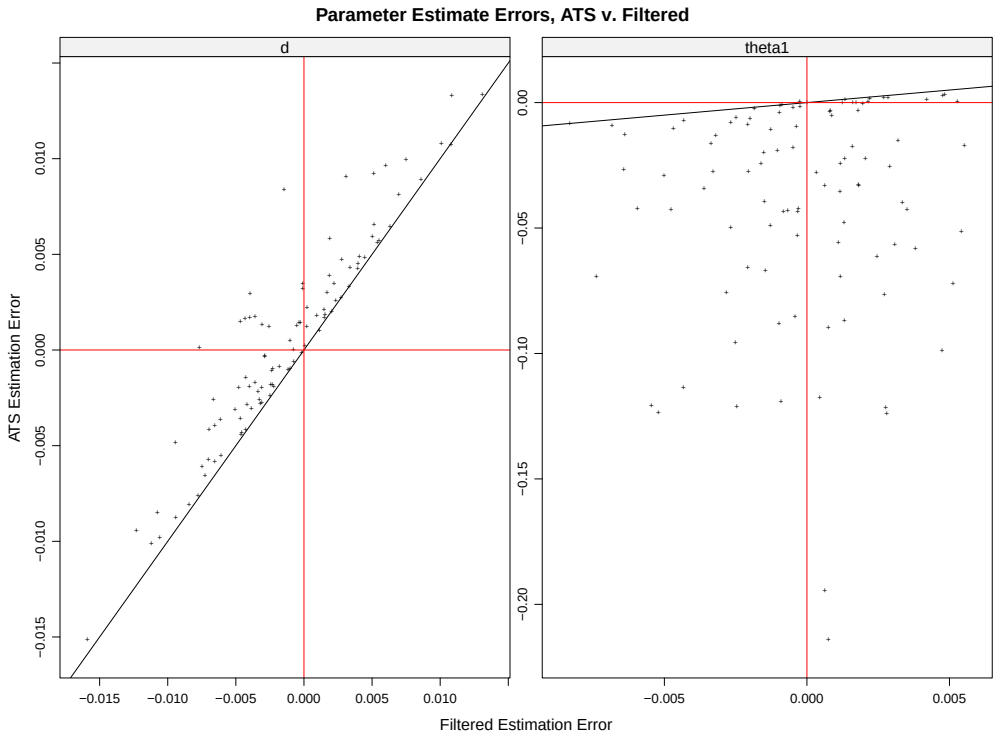

Fig. 4.64. FARMA(p = 0, d, q = 1) Scatterplot of PPS-REG Estimates v. PPS-REG Filtered Estimates by Parameter, *d* or θ

The worst PPS-REG errors are improved by filtering and more importantly there is a positive linear relationship between the PPS-REG error and the PPS-REG filtered error of θ indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

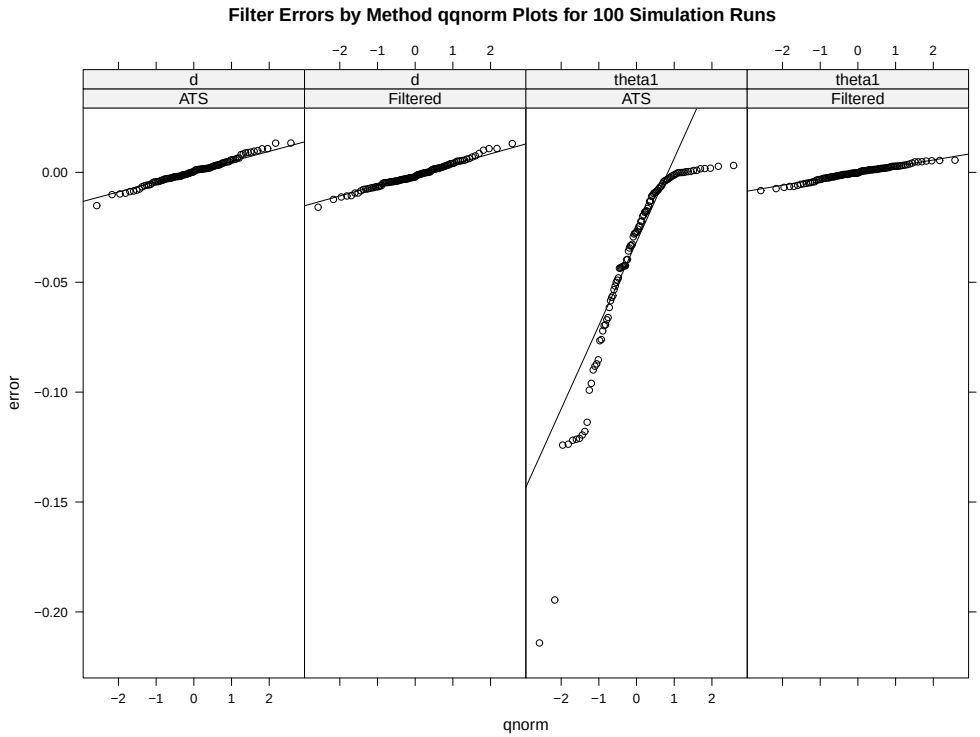

Fig. 4.65. FARMA(p = 0, d, q = 1) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered, and Parameter, *d* or θ

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

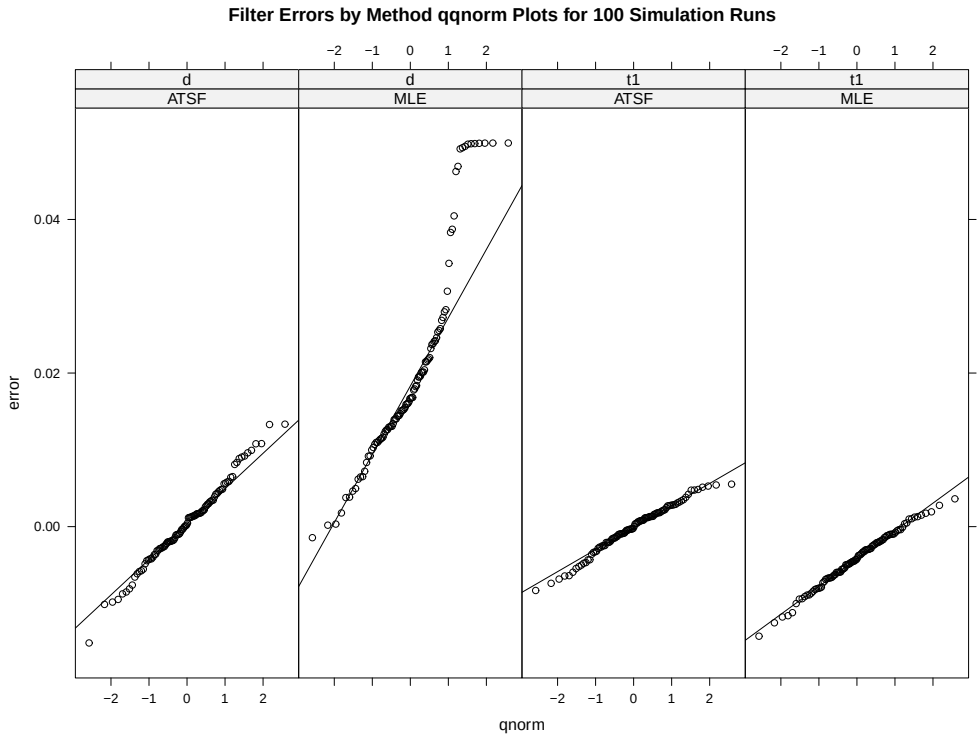

Fig. 4.66. FARMA(p = 0, d, q = 1) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE, and Parameter, d or θ

Filtered PPS-REG exhibits a left tail for $\hat{\theta}$, but otherwise is normally distributed and is an improvement from a heavier PPS-REG left tail.

4.2.8 ARFIMA(p = 1, d, q = 2)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

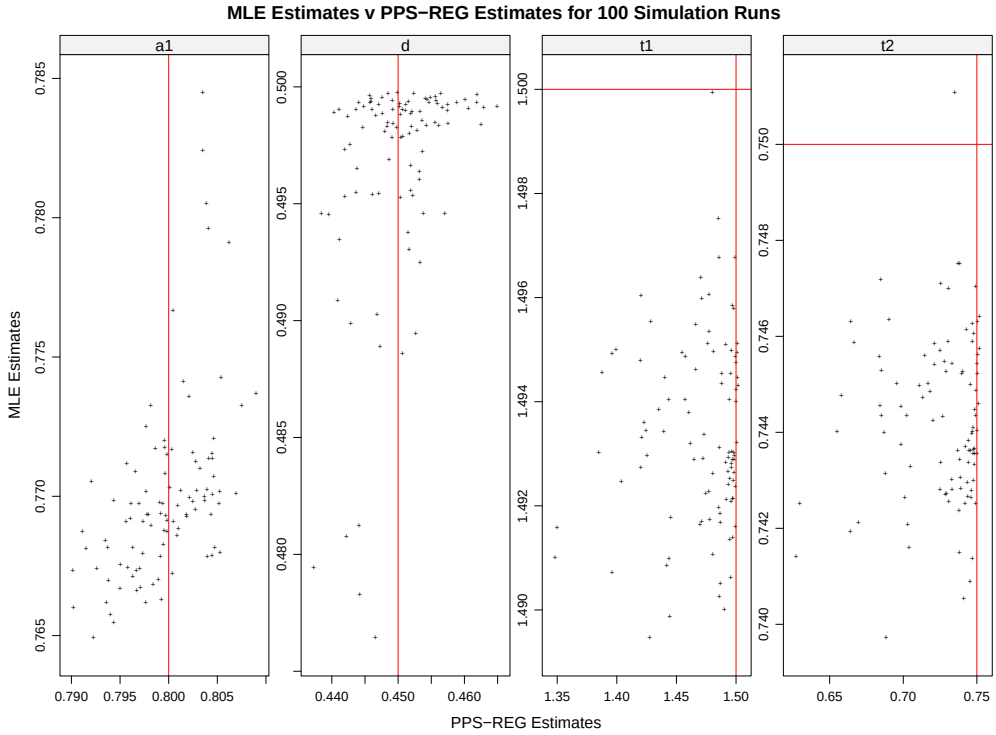

Fig. 4.67. FARMA(p = 1, d, q = 2) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ, d, θ_1, or θ_2

PPS-REG estimates of d and ϕ possess far less variability than MLE estimates. MLE estimates of θ_1 and θ_2 possess far less variability than PPS-REG estimates.

To assess the normality of our estimates by method, QQ plots were constructed:

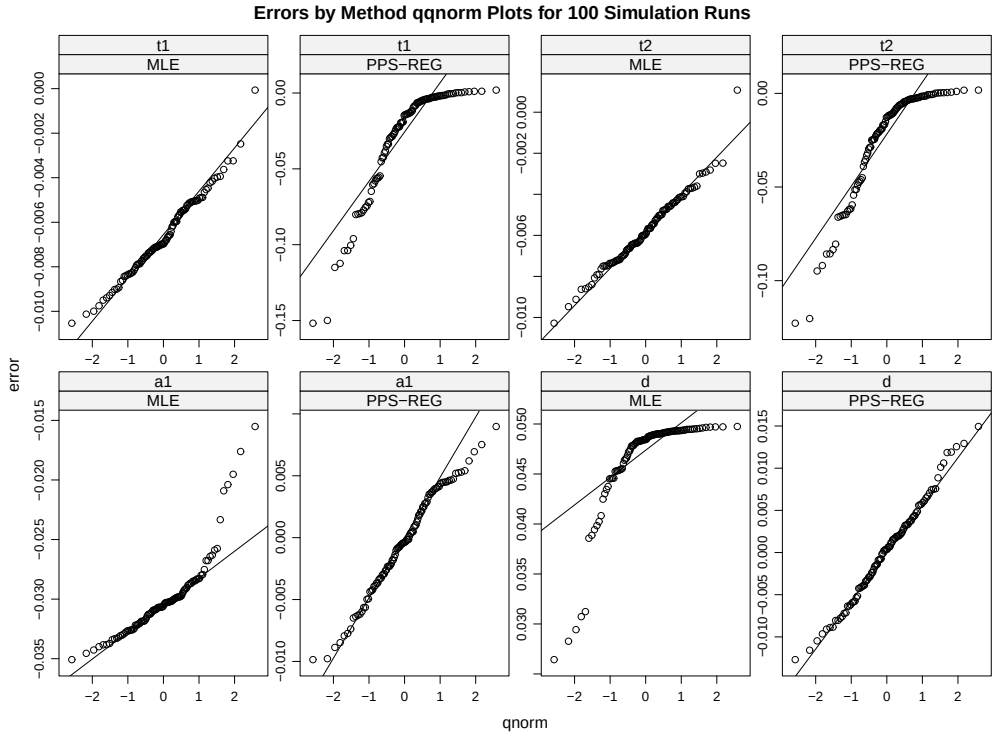

Fig. 4.68. FARMA(p = 1, d, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ, d, θ_1, or θ_2

PPS-REG estimates are normally distributed for all parameters about their simulated values, except for $\hat{\theta}_1$ and $\hat{\theta}_2$, which is normally distributed with a bias towards 0. MLE estimates are normally distributed for all parameters about their simulated values, except for $\hat{d}$, which is normally distributed with a bias towards 0.

AutoRegressive (AR) Component Filtration

The $FARMA(p=1,d,q=2)$ time series model is:

$$(1-\phi B)(1-B)^d x_t = (1-B)^d x_t - \phi(1-B)^d x_{t-1} = (1-B)^d (x_t - \phi x_{t-1})$$

$$= (1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The $FARMA(p=0,d,q=2)$ filtered data time series model is:

$$(1-B)^d x_t^* = (1-B)^d (x_t - \hat{\phi} x_{t-1}) \approx 1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The PPS-REG log power spectrum estimation on the $FARMA(p=0,d,q=2)$ filtered data yielded the following log averaged periodogram versus frequency results:

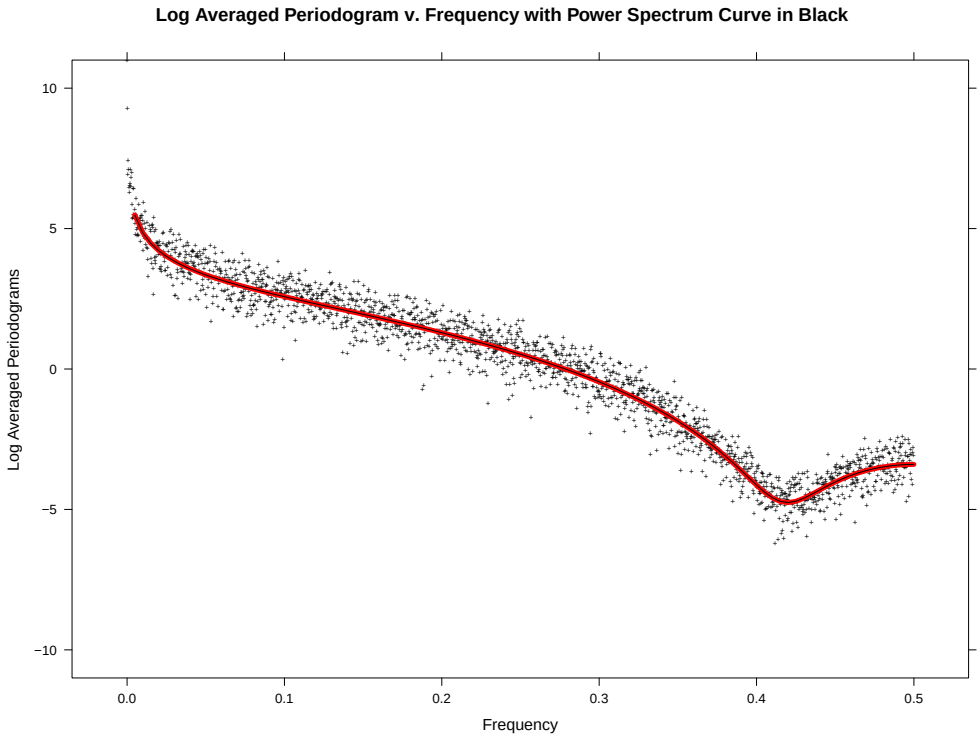

Fig. 4.69. FARMA(p = 0, d, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

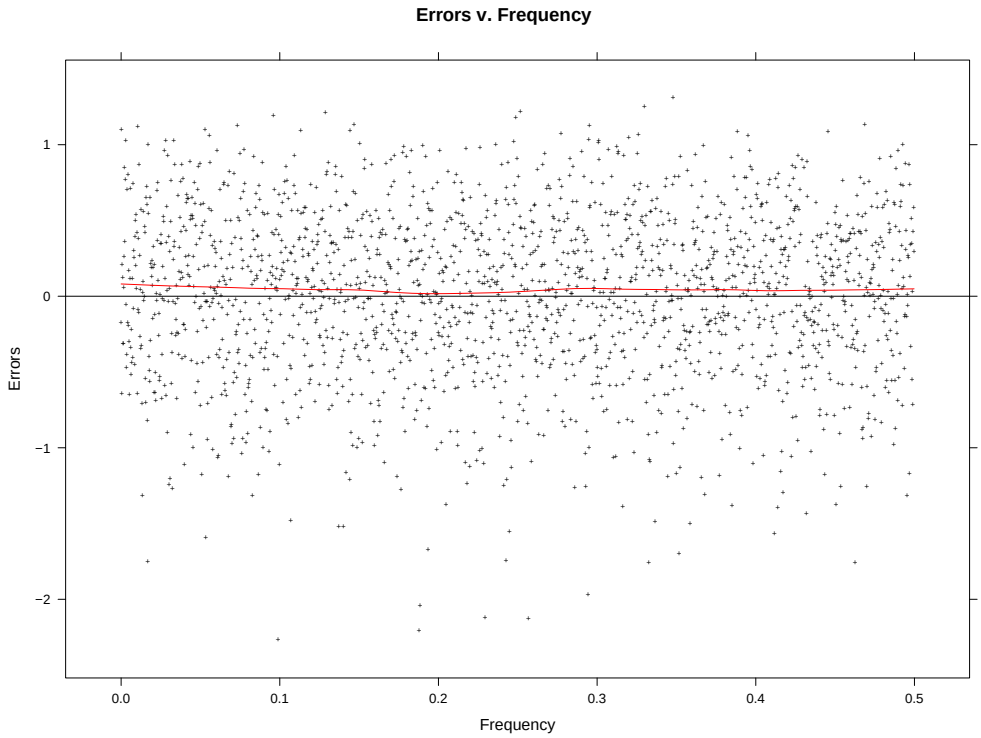

Fig. 4.70. FARMA(p = 0, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

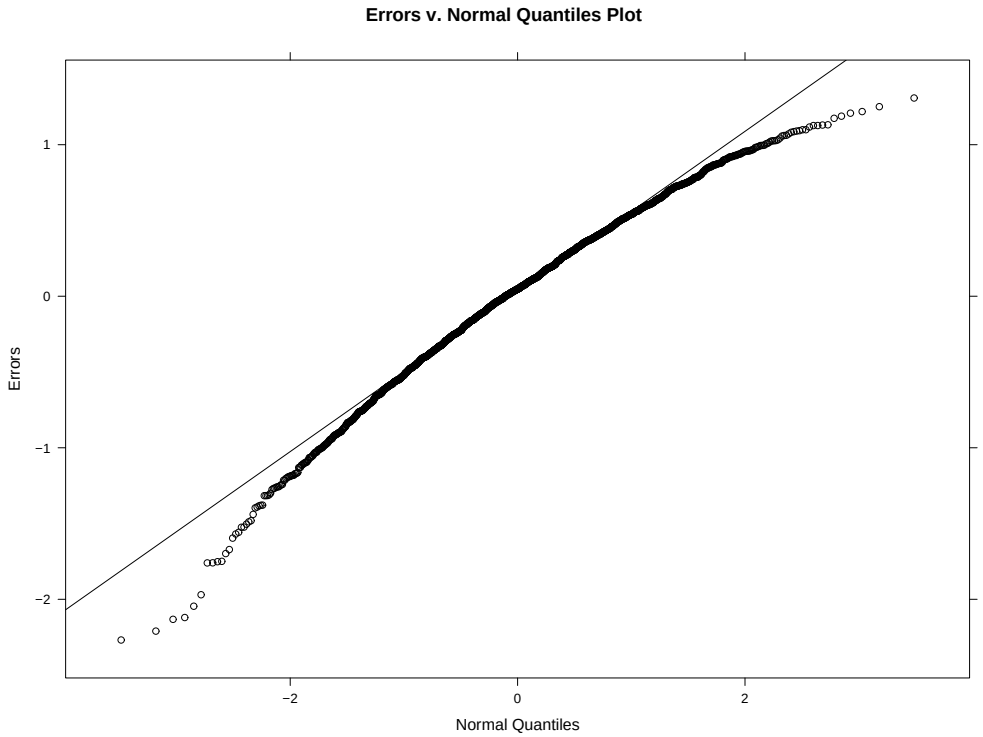

Fig. 4.71. FARMA(p = 0, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

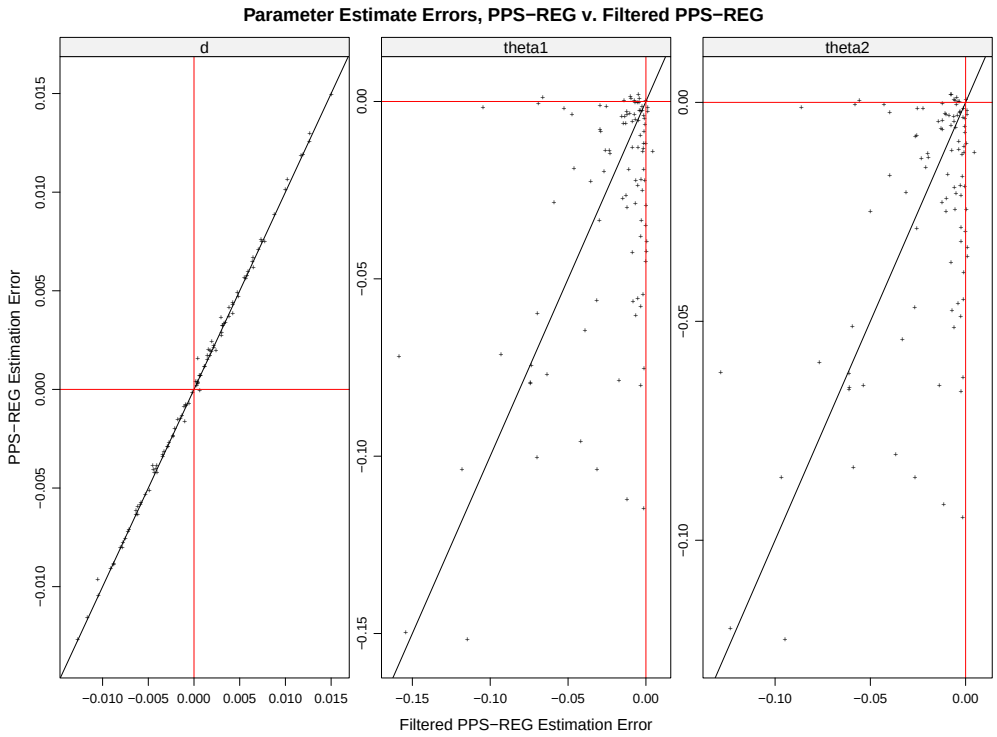

Fig. 4.72. FARMA(p = 0, d, q = 2) Scatterplot of PPS-REG Estimates v. PPS-REG Filtered Estimates by Parameter, d, θ_1 or θ_2

The worst PPS-REG errors are improved by filtering and more importantly there is a positive linear relationship between the PPS-REG error and the PPS-REG filtered error of θ_1 and θ_2 indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

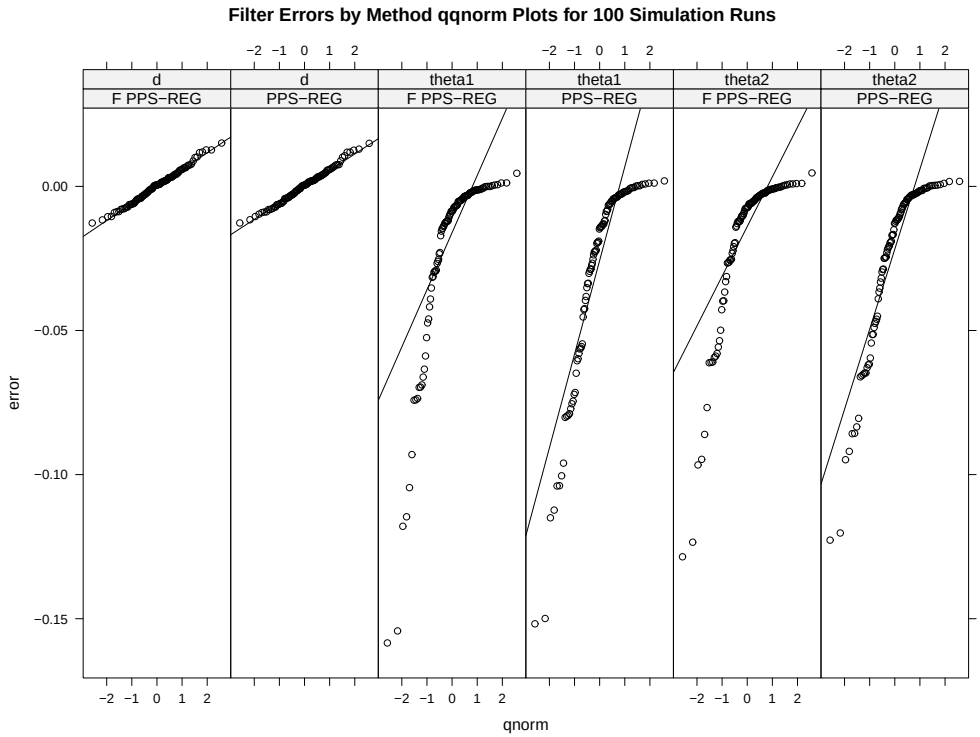

Fig. 4.73. FARMA(p = 0, d, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered, and Parameter, d, θ_1, or θ_2

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

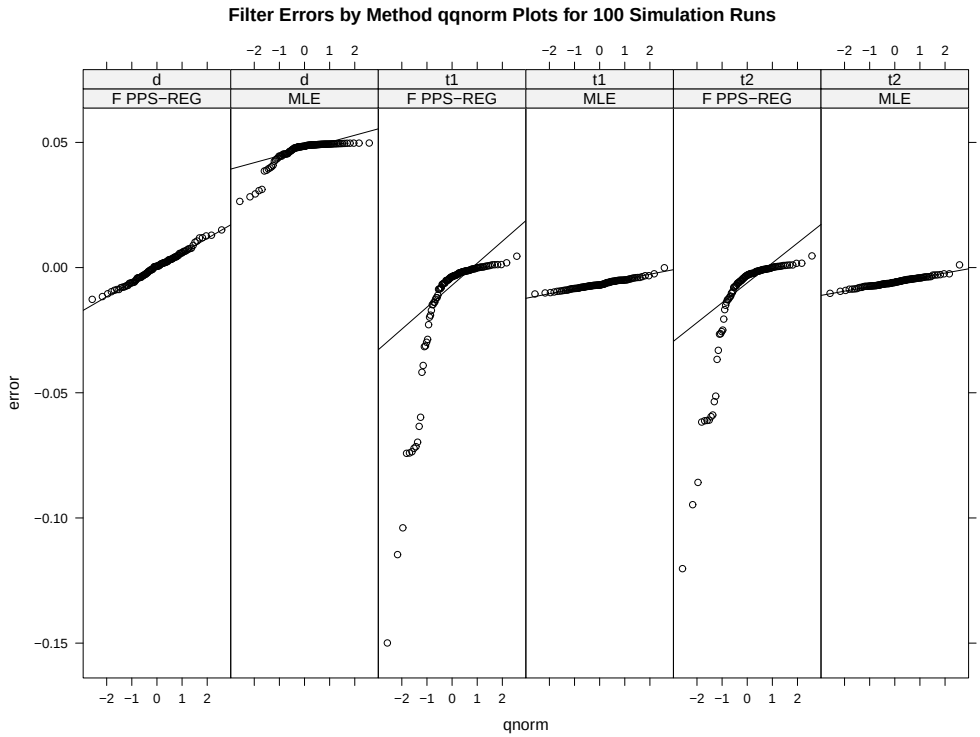

Fig. 4.74. FARMA(p = 0, d, q = 2) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE, and Parameter, d, θ_1, or θ_2

Filtered PPS-REG exhibits a left tail for $\hat{\theta}_1$ and $\hat{\theta}_2$, but otherwise is normally distributed and is an improvement from a heavier PPS-REG left tail.

4.2.9 ARFIMA(p = 2, d, q = 2)

PPS-REG parameter estimates were initialized with Conditional Sum of Squares for ARMA parameters and $d^{(0)} = 0.25$ for all simulation runs. The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

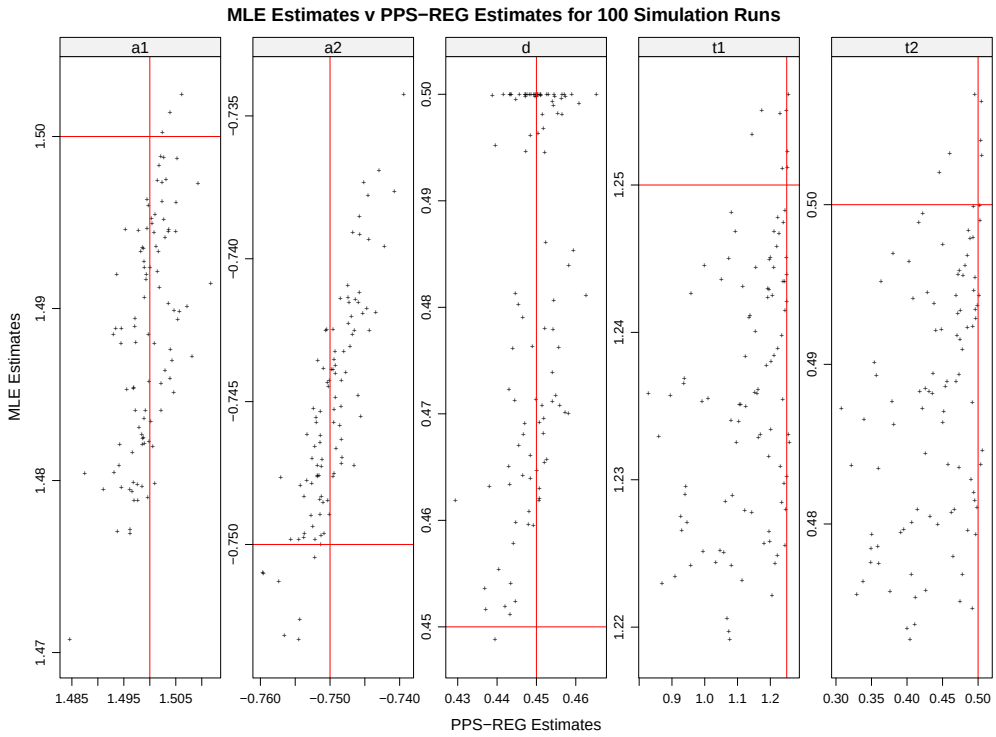

Fig. 4.75. FARMA(p = 2, d, q = 2) Scatterplot of MLE estimates v. PPS-REG estimates by Parameter, ϕ_1, ϕ_2, d, θ_1, or θ_2

PPS-REG estimates of d, ϕ_1, and ϕ_2 possess far less variability than MLE estimates. MLE estimates of θ_1 and θ_2 possess far less variability than PPS-REG estimates.

To assess the normality of our estimates by method, QQ plots were constructed:

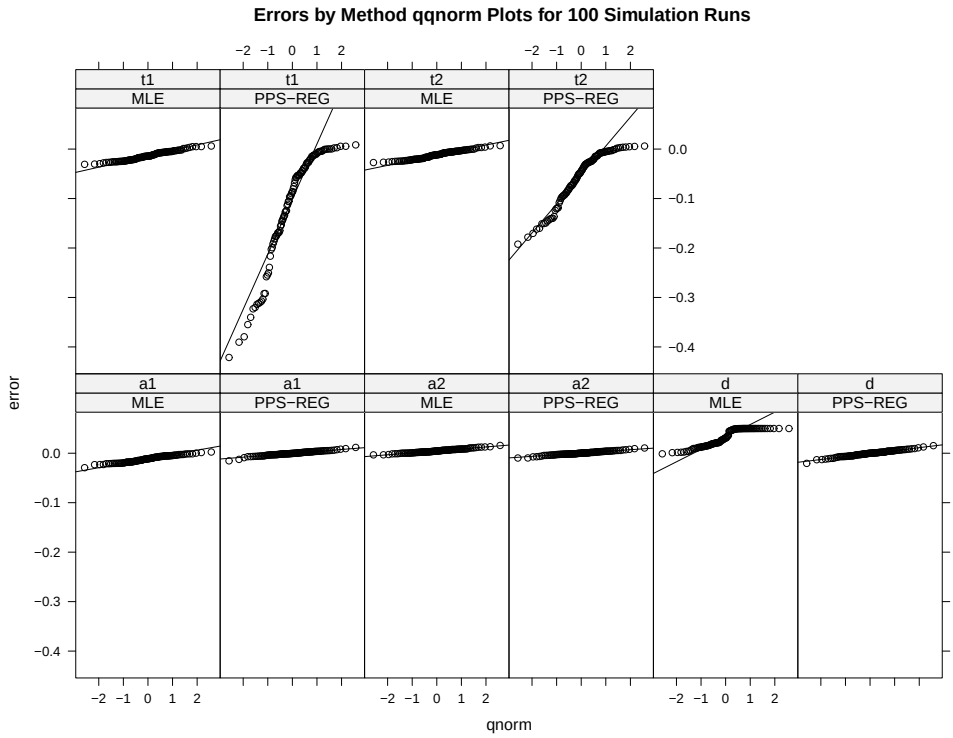

Fig. 4.76. FARMA(p = 2, d, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. MLE, and Parameter, ϕ_1, ϕ_2, d, θ_1, or θ_2

PPS-REG estimates are normally distributed for all parameters about their simulated values, except for $\hat{\theta}_1$ and $\hat{\theta}_2$, which is normally distributed with a bias towards 0. MLE estimates are normally distributed for all parameters about their simulated values, except for $\hat{d}$, which is normally distributed with a bias towards 0.

AutoRegressive (AR) Component Filtration

The $FARMA(p=2,d,q=2)$ time series model is:

$$(1-\phi_1 B-\phi_2 B^2)(1-B)^d x_t = (1-B)^d(x_t-\phi_1 x_{t-1}-\phi_2 x_{t-2})$$

$$= (1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The $FARMA(p=0,d,q=2)$ filtered data time series model is:

$$(1-B)^d x_t^* = (1-B)^d(x_t-\hat{\phi}_1 x_{t-1}-\hat{\phi}_2 x_{t-2}) \approx (1+\theta_1 B+\theta_2 B^2)\varepsilon_t$$

The PPS-REG log power spectrum estimation on the $FARMA(p=0,d,q=2)$ filtered data yielded the following log averaged periodogram versus frequency results:

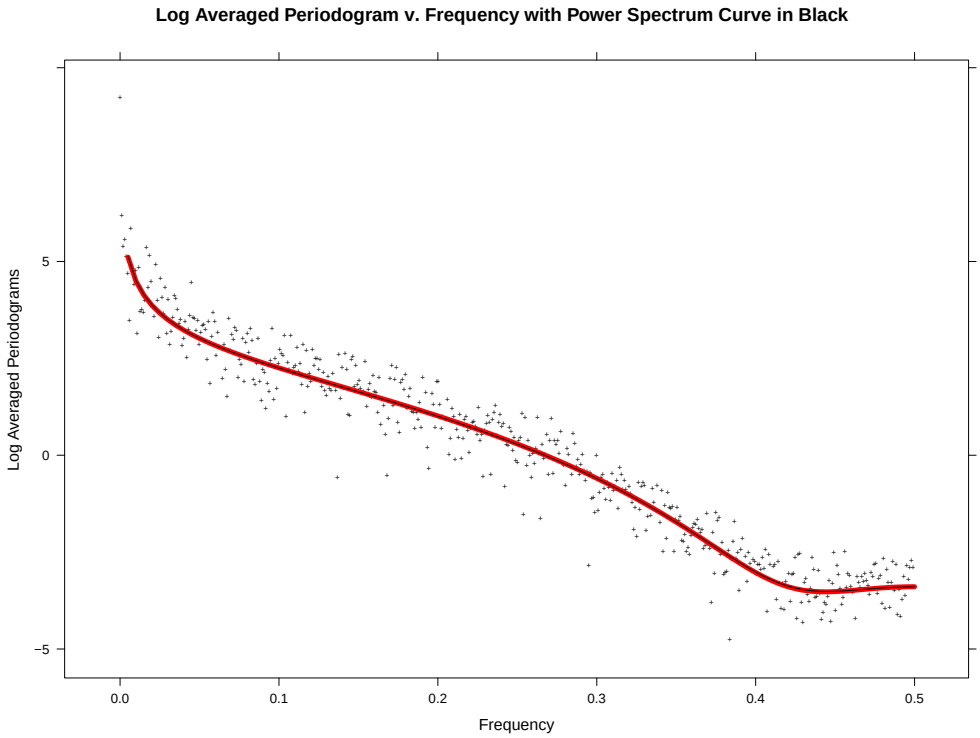

Fig. 4.77. FARMA(p = 0, d, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

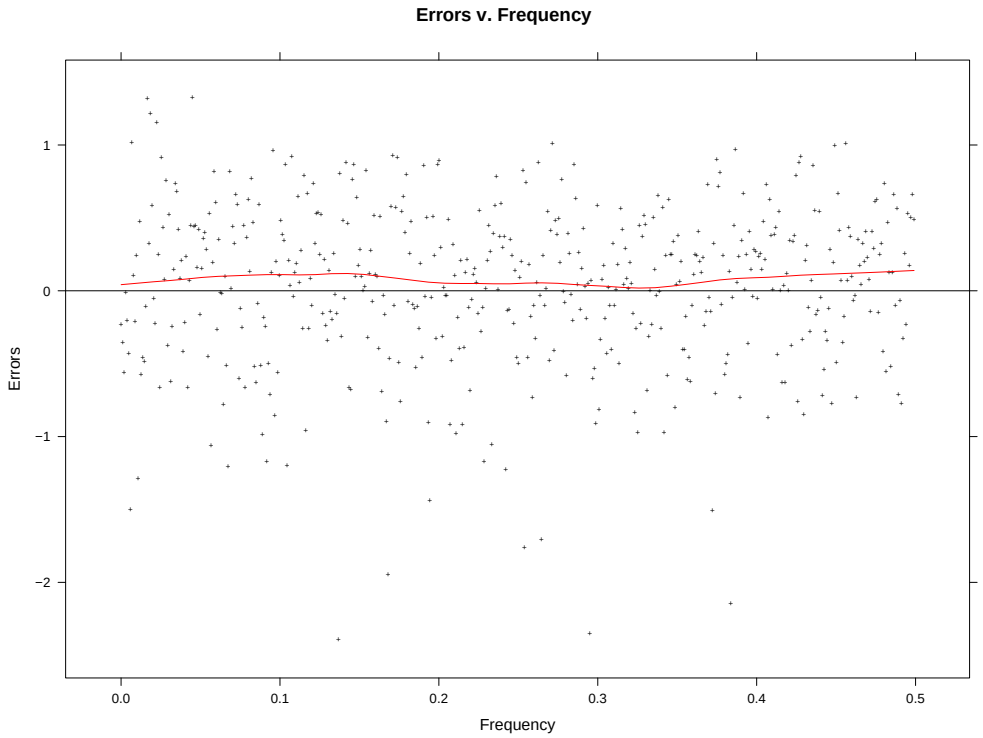

Fig. 4.78. FARMA(p = 0, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

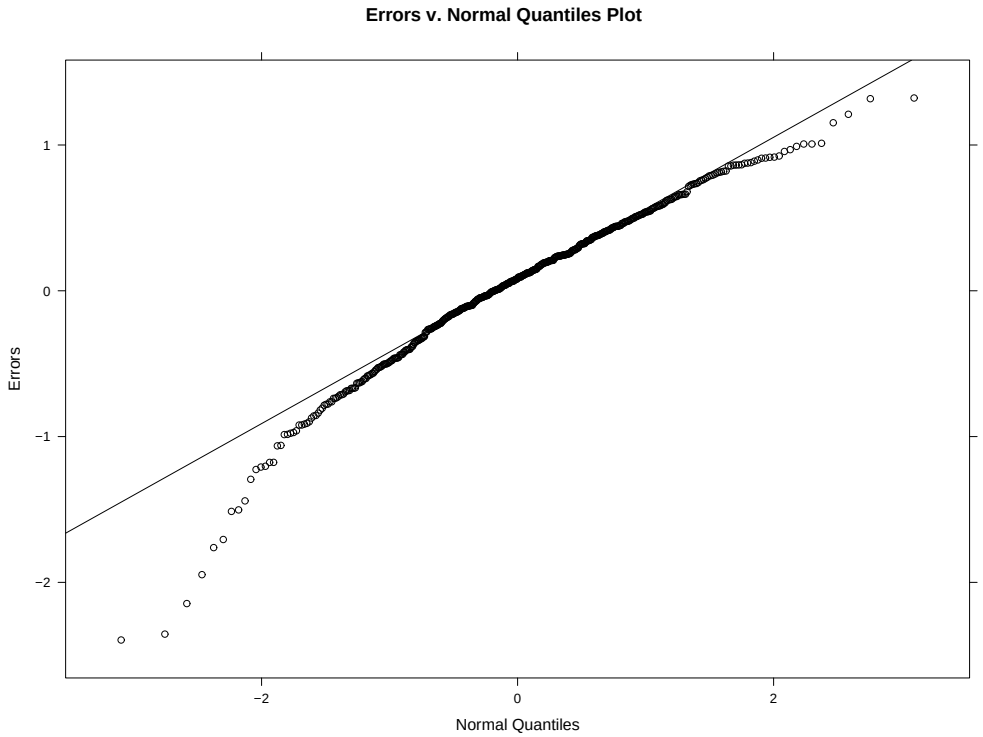

Fig. 4.79. FARMA(p = 0, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

The PPS-REG results summarized and compared to the summarized MLE results yielded the following scatterplot:

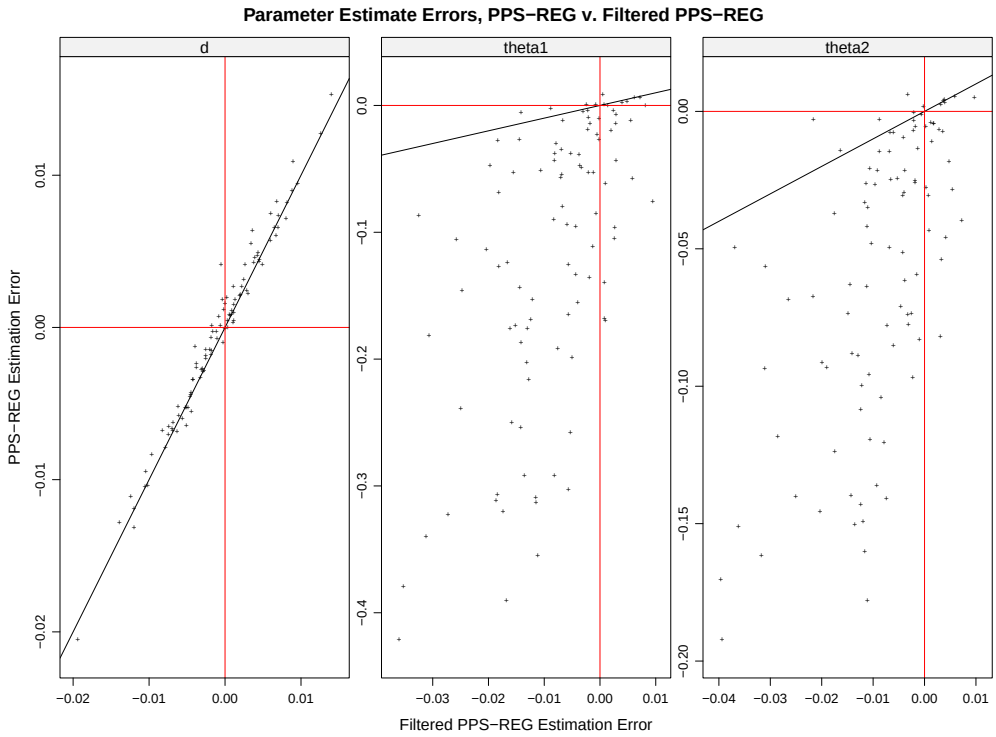

Fig. 4.80. FARMA(p = 0, d, q = 2) Scatterplot of PPS-REG Estimates v. PPS-REG Filtered Estimates by Parameter, d, θ_1 or θ_2

The worst PPS-REG errors are improved by filtering and more importantly there is a positive linear relationship between the PPS-REG error and the PPS-REG filtered error of θ_1 and θ_2 indicating consistent results between methods.

To assess the normality of our estimates by method, QQ plots were constructed:

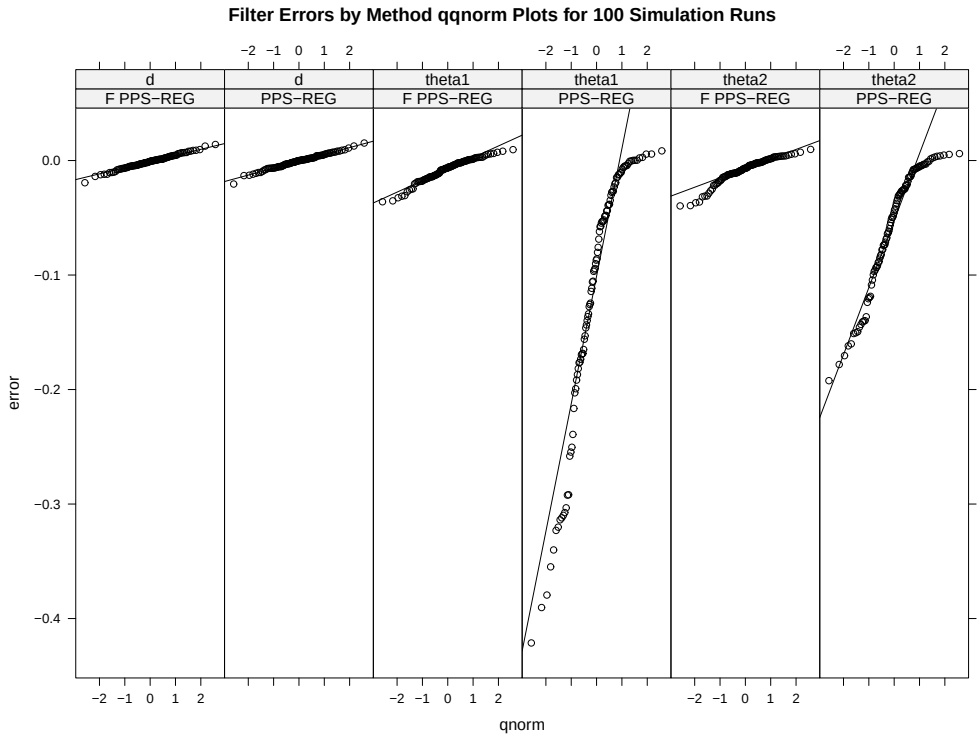

Fig. 4.81. FARMA(p = 0, d, q = 2) Normal QQ-Plots of Estimation Error by Estimation Method, PPS-REG v. PPS-REG Filtered, and Parameter, d, θ_1, or θ_2

There is improvement in the heavy left tail and the worst PPS-REG errors by filtering.

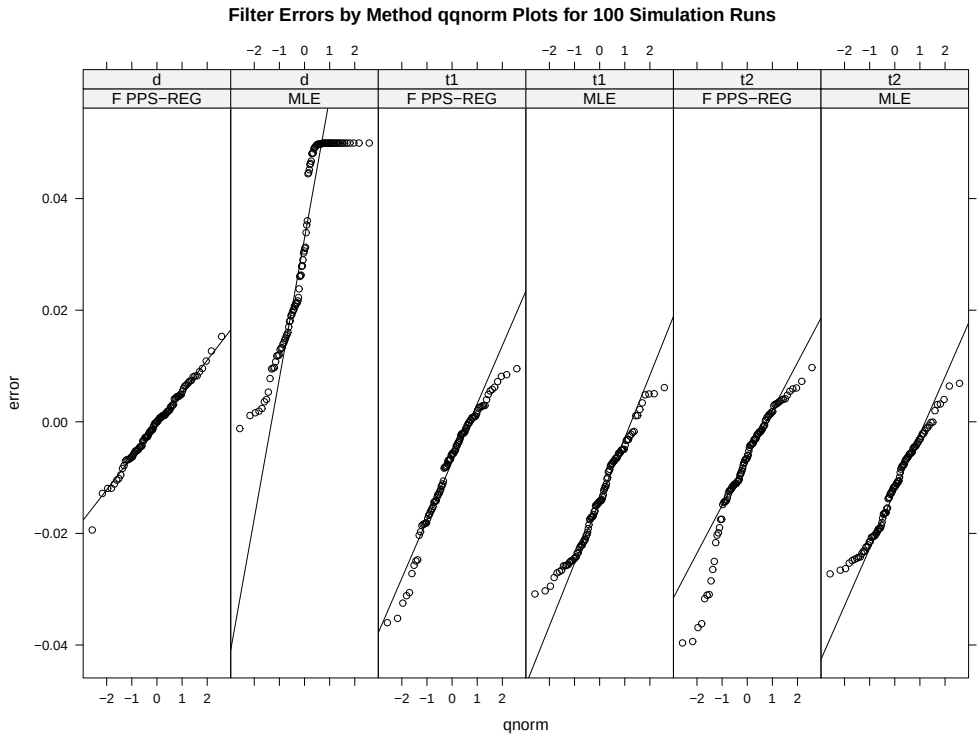

Fig. 4.82. FARMA(p = 0, d, q = 2) Normal QQ-Plots of Error by Estimation Method, Filtered PPS-REG v. MLE, and Parameter, d, θ_1, or θ_2

Filtered PPS-REG exhibits a left tail for $\hat{\theta}_1$ and $\hat{\theta}_2$, but otherwise is normally distributed and is an improvement from a heavier PPS-REG left tail.

5. APPENDIX

5.1 ARMA Models

5.1.1 AR(p = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

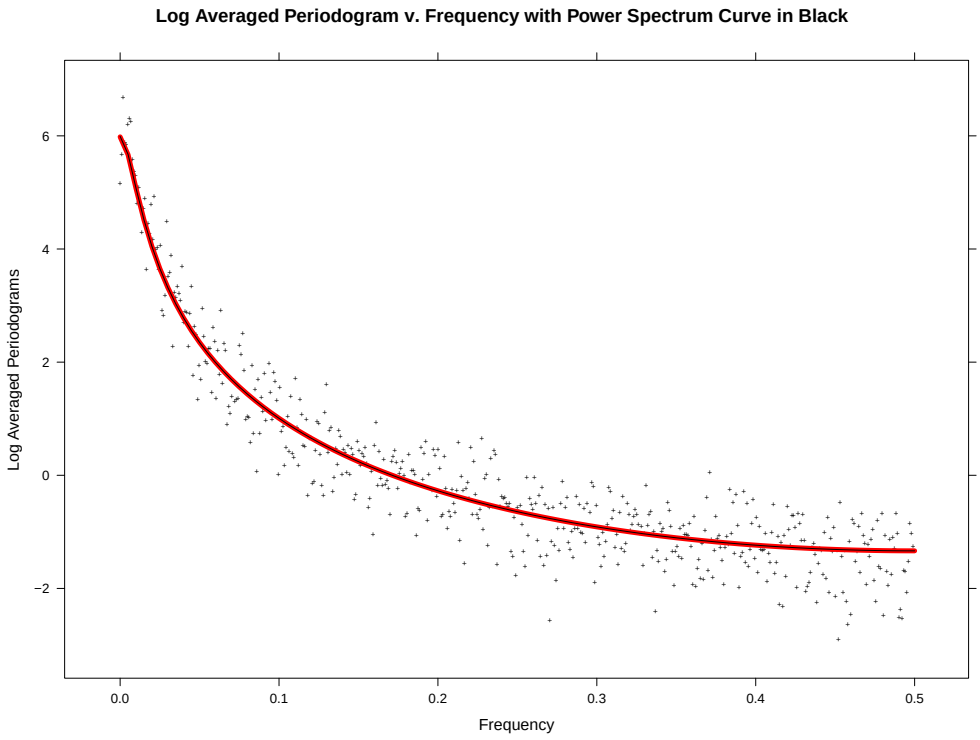

Fig. 5.1. AR(p = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

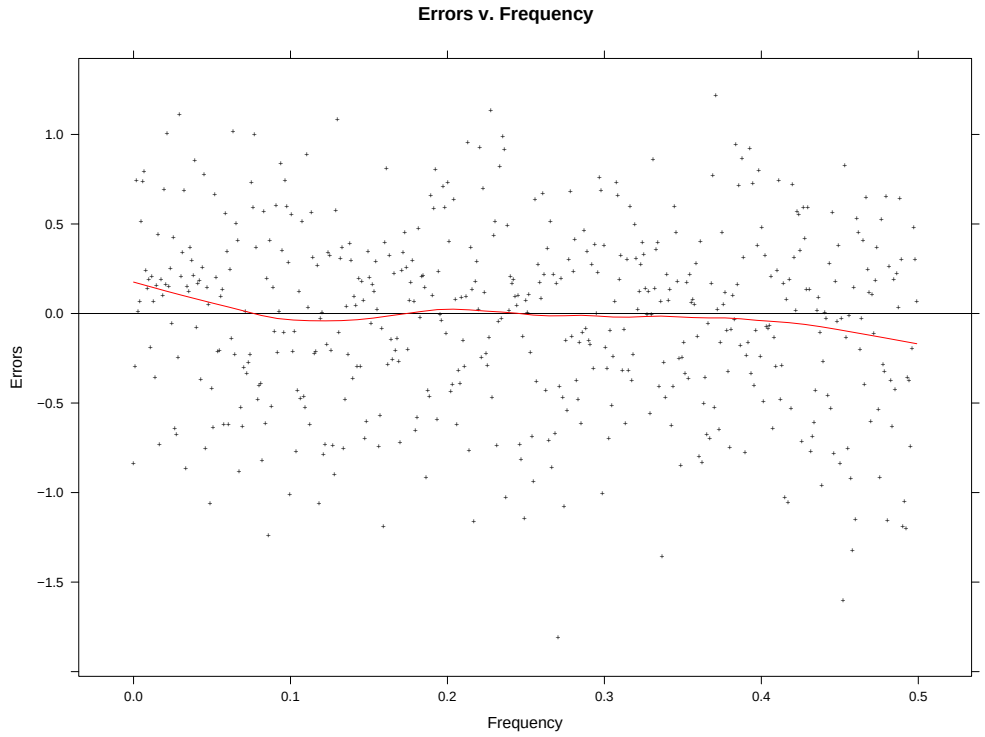

Fig. 5.2. AR(p = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

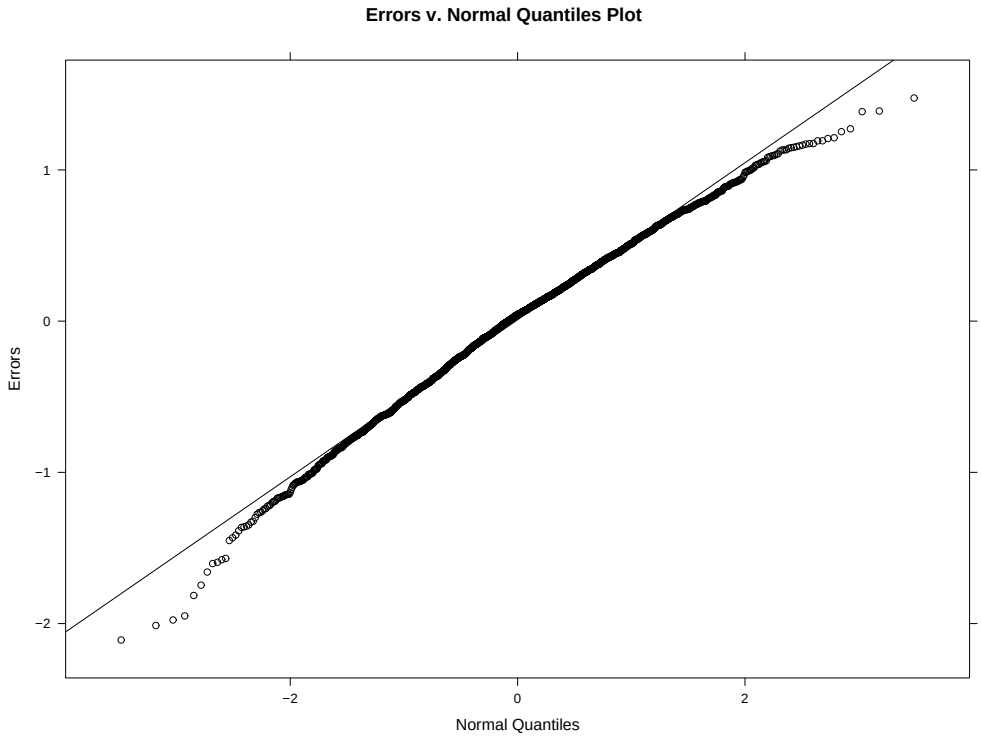

Fig. 5.3. AR(p = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.1.2 AR(p = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

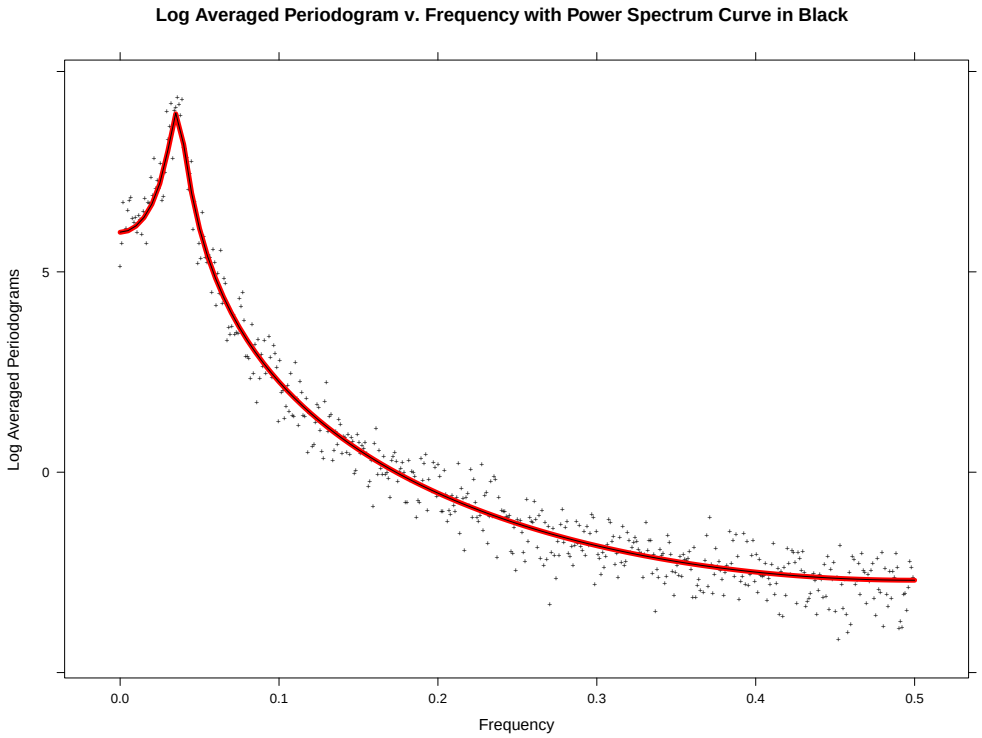

Fig. 5.4. AR(p = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

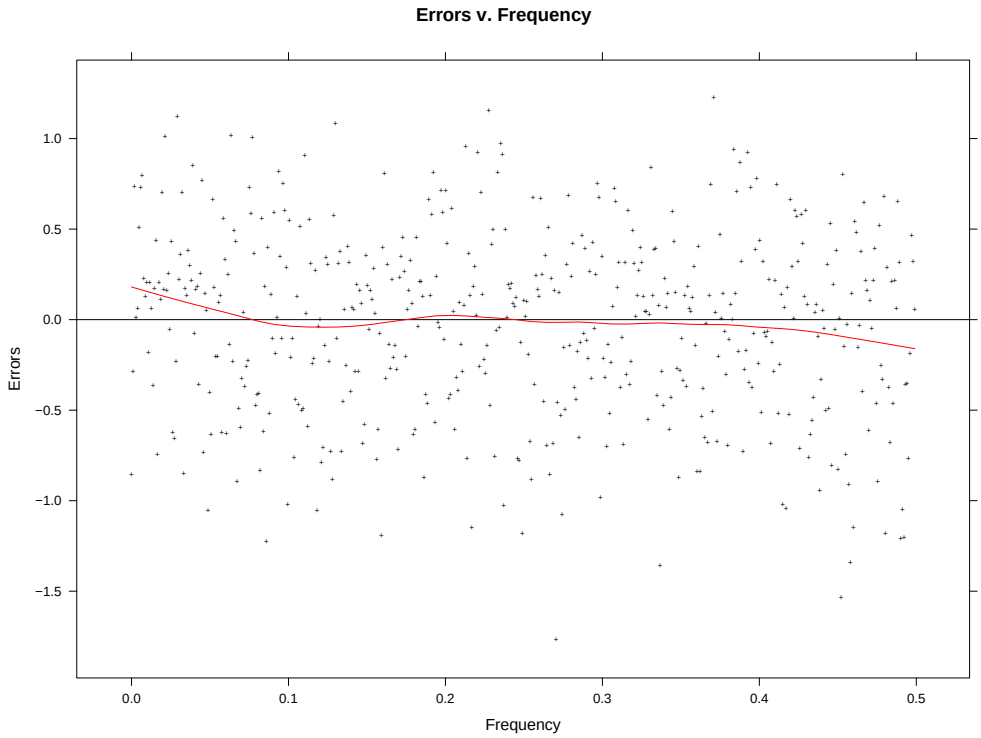

Fig. 5.5. AR(p = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

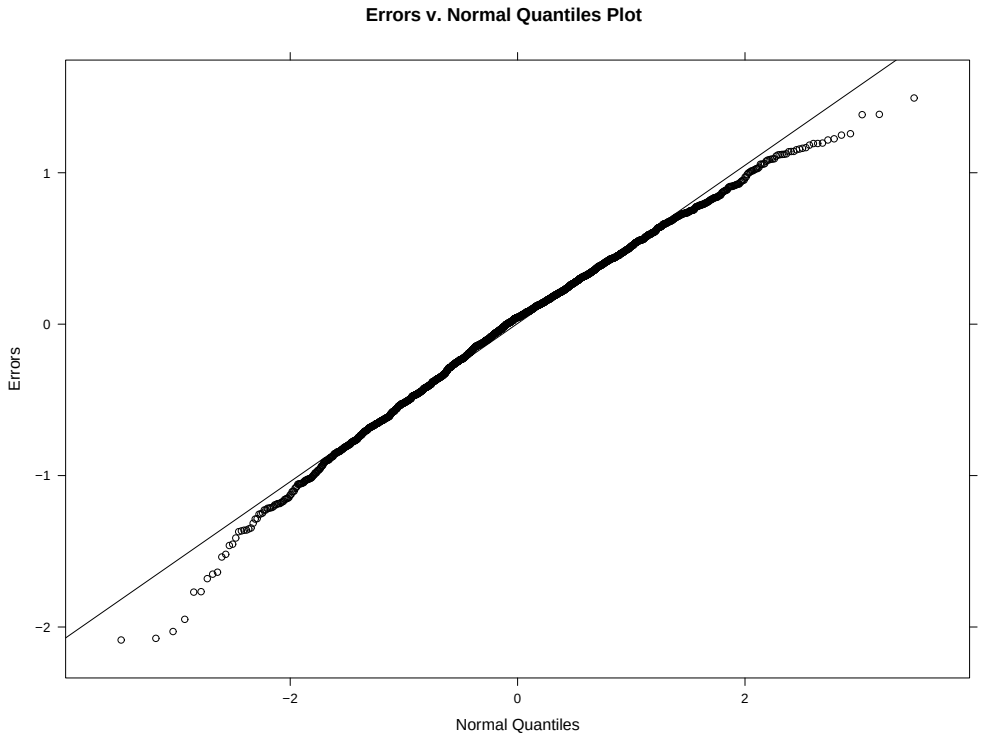

Fig. 5.6. AR(p = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

5.1.3 MA(q = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

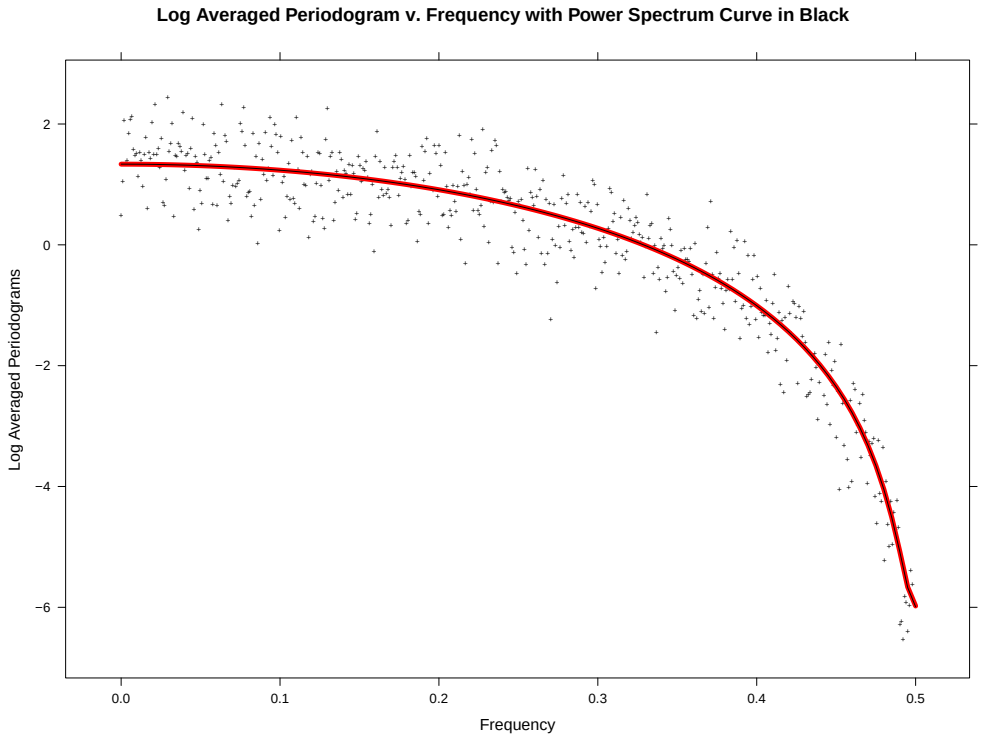

Fig. 5.7. MA(q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

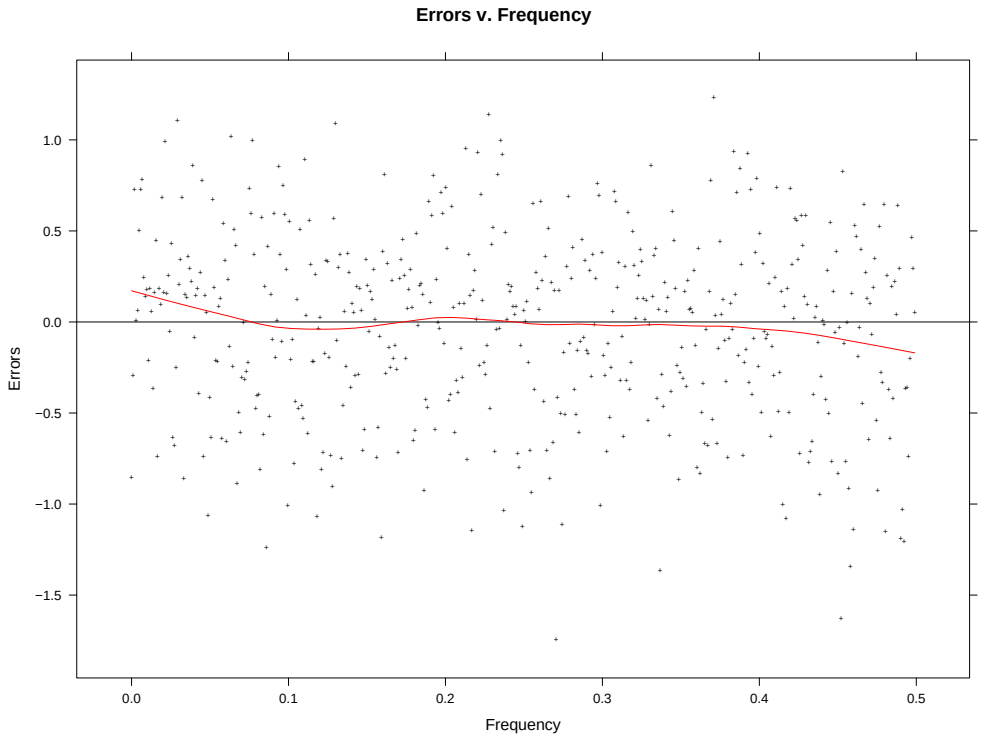

Fig. 5.8. MA(q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

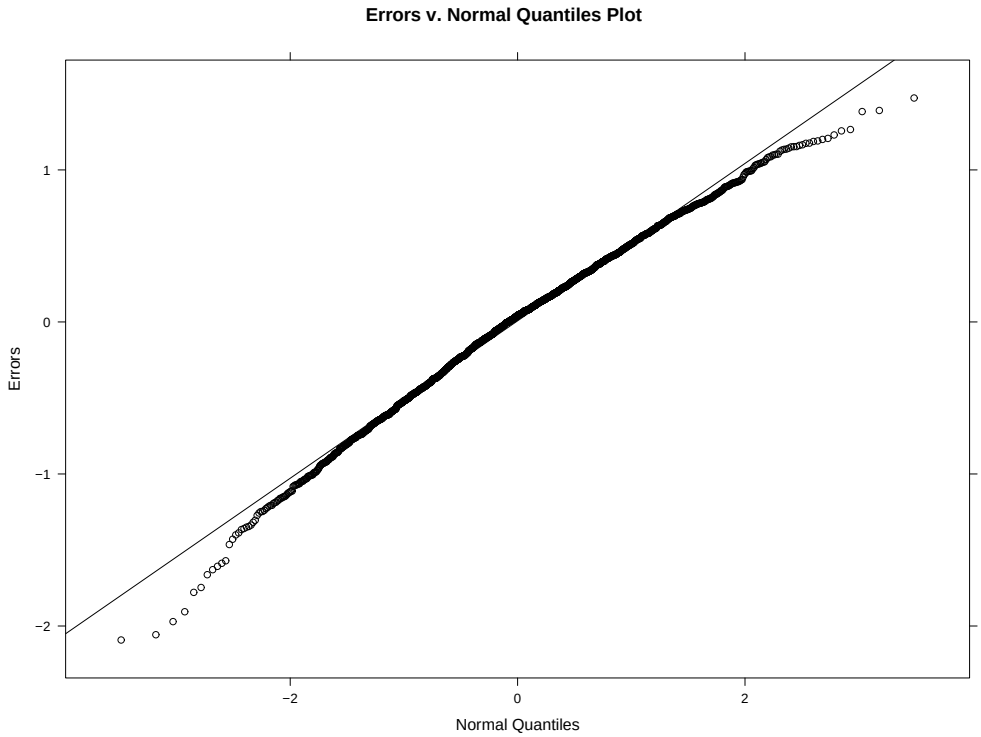

Fig. 5.9. MA(q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

5.1.4 MA(q = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

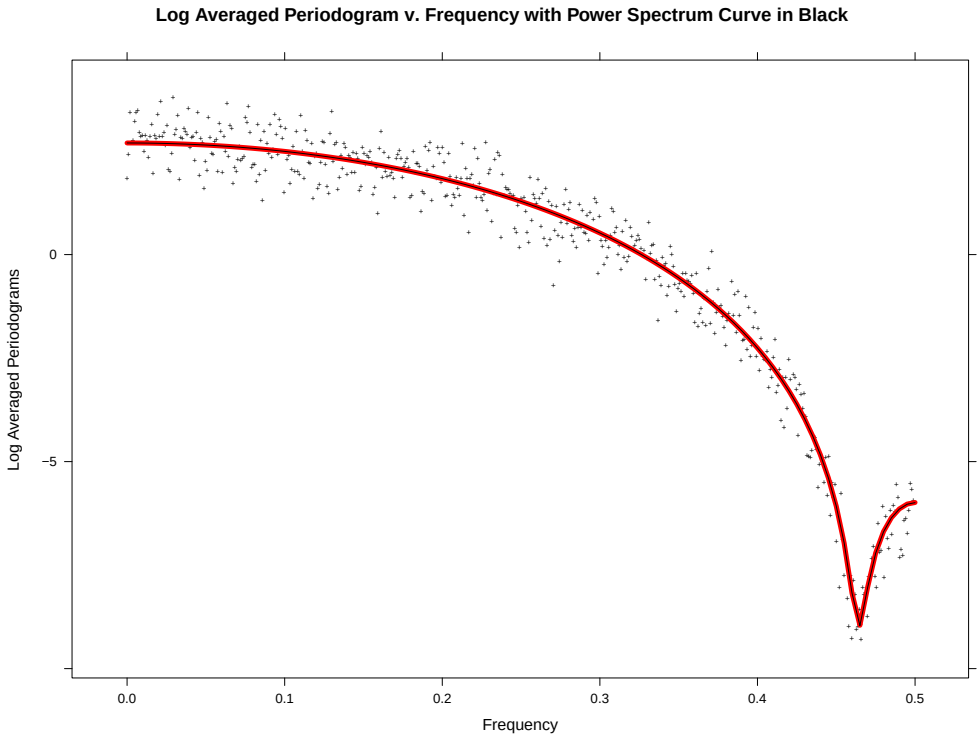

Fig. 5.10. MA(q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

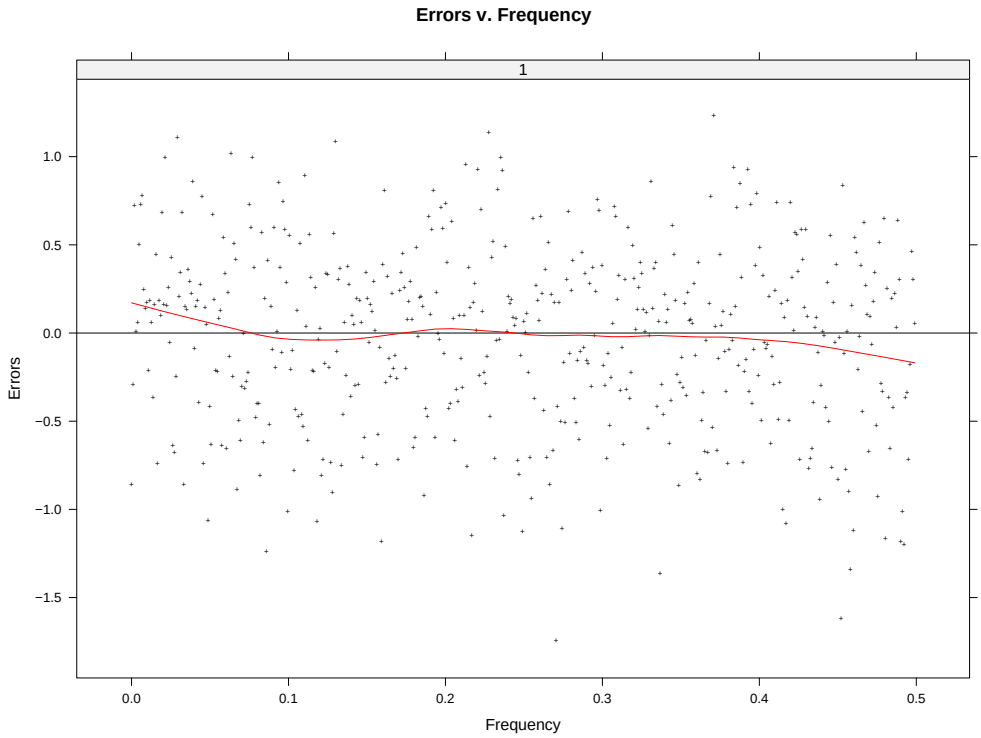

Fig. 5.11. MA(q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

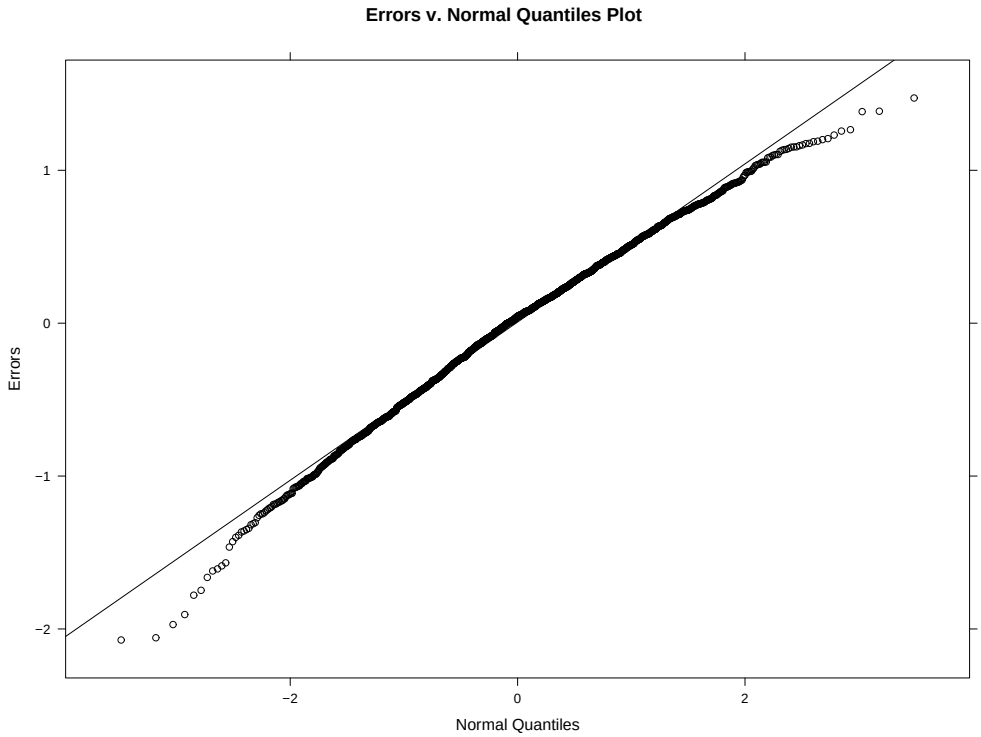

Fig. 5.12. MA(q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

5.1.5 ARMA(p = 1, q = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

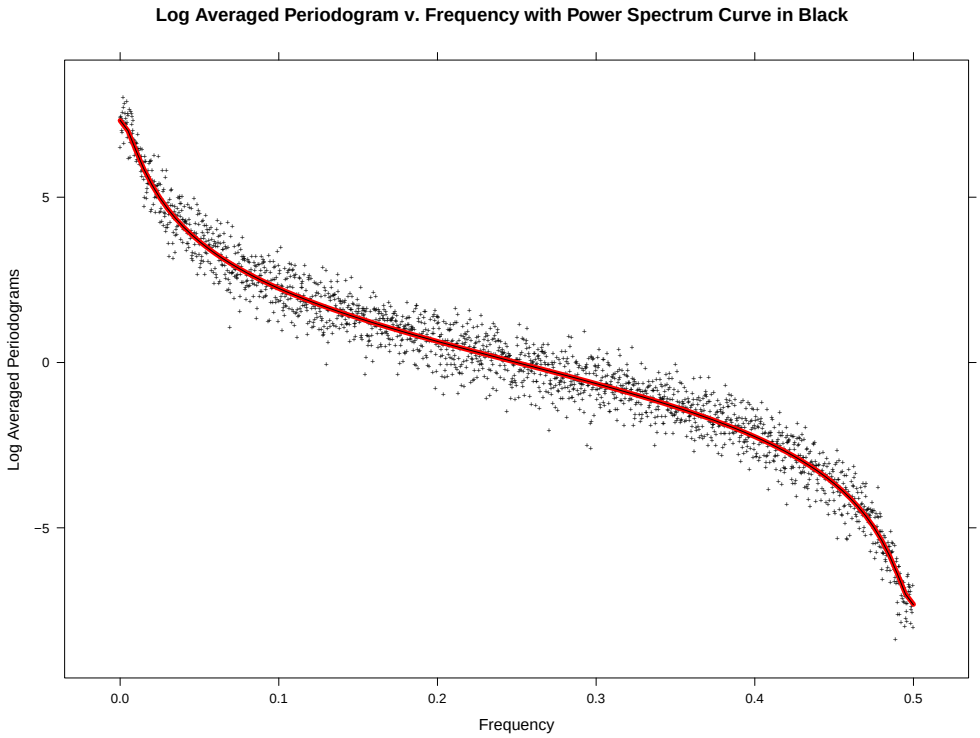

Fig. 5.13. ARMA(p = 1, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

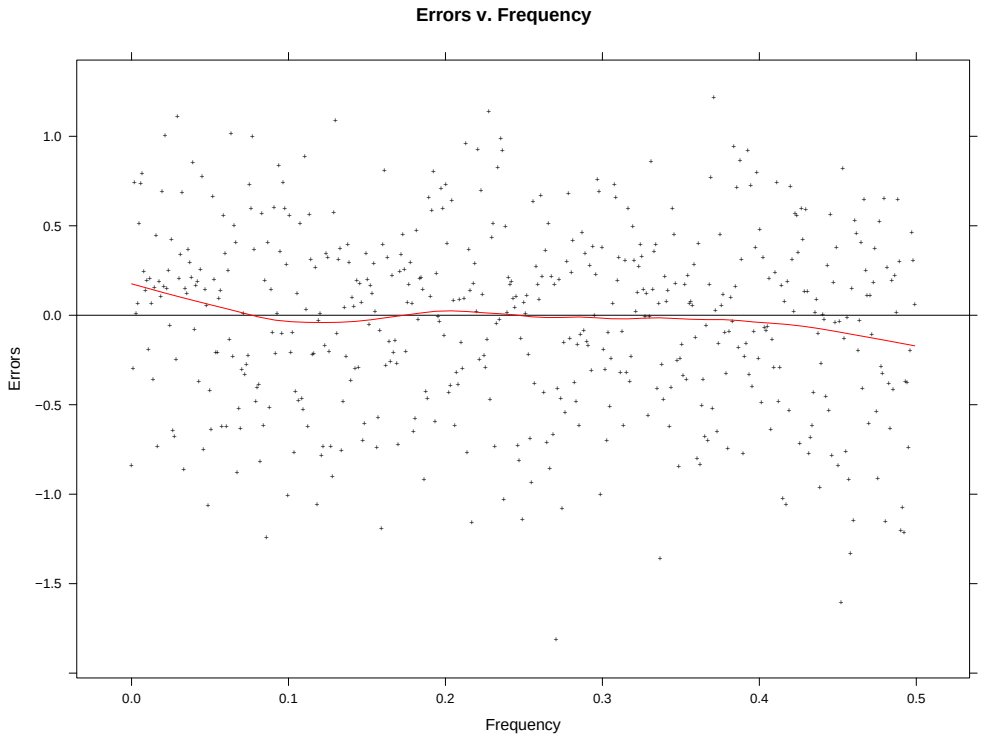

Fig. 5.14. ARMA(p = 1, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

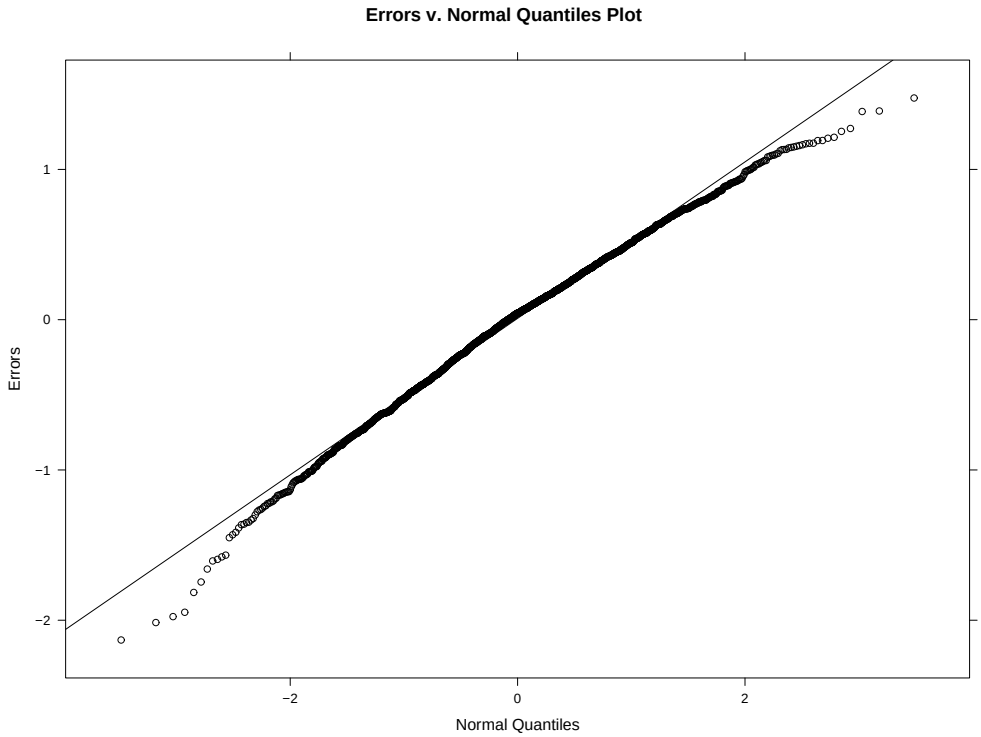

Fig. 5.15. ARMA(p = 1, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.1.6 ARMA(p = 2, q = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

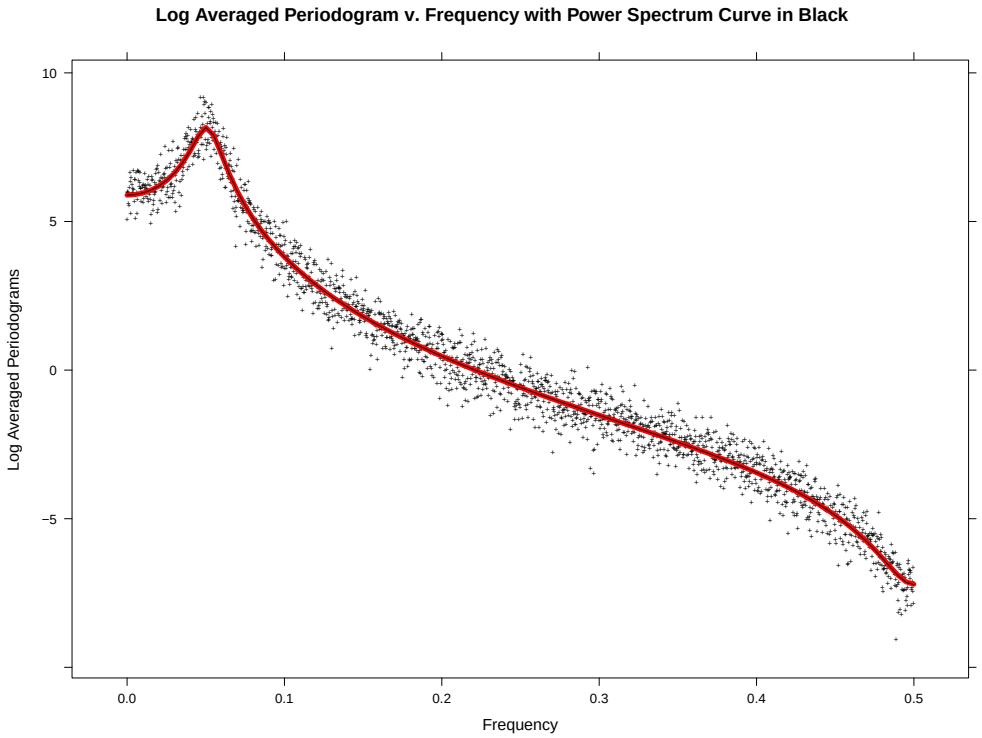

Fig. 5.16. ARMA(p = 2, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

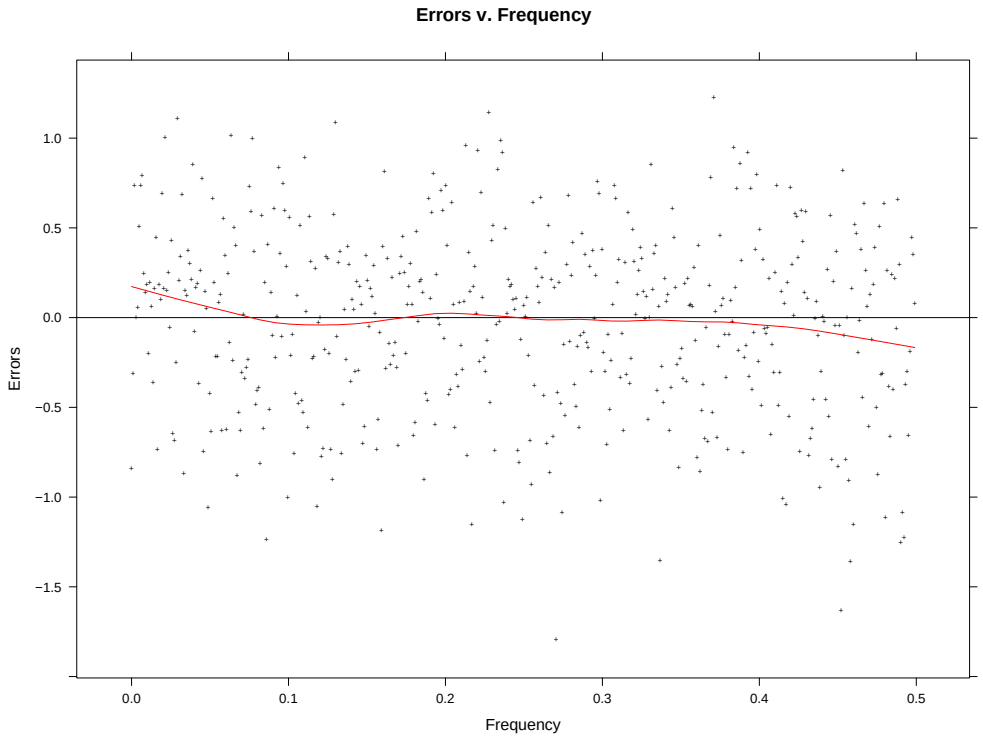

Fig. 5.17. ARMA(p = 2, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

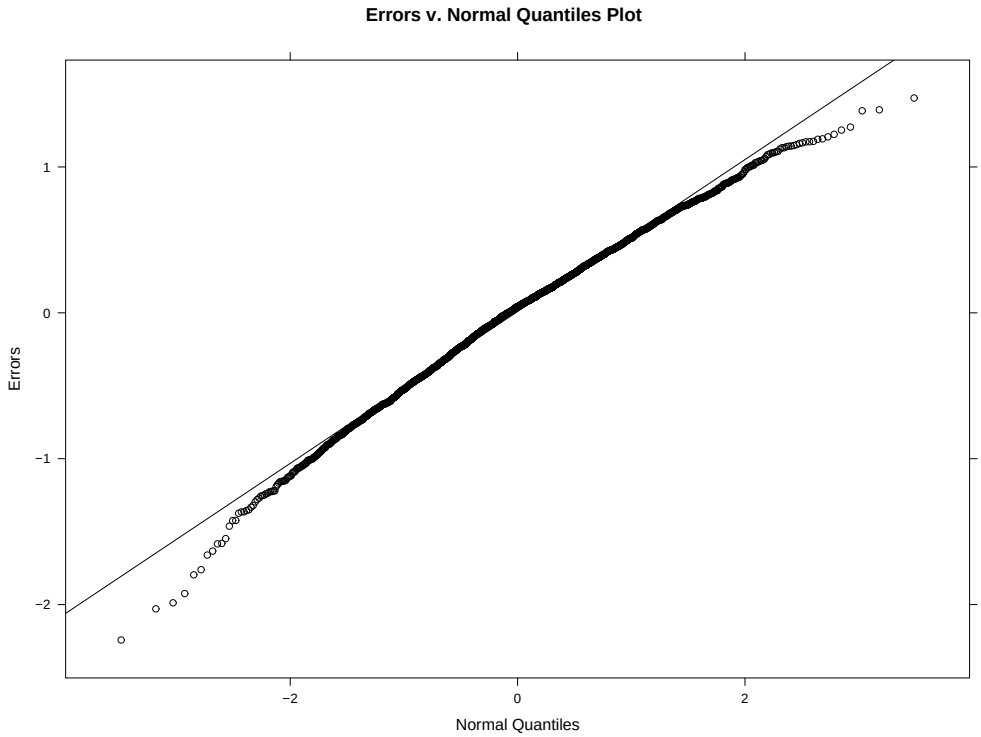

Fig. 5.18. ARMA(p = 2, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

5.1.7 ARMA(p = 1, q = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

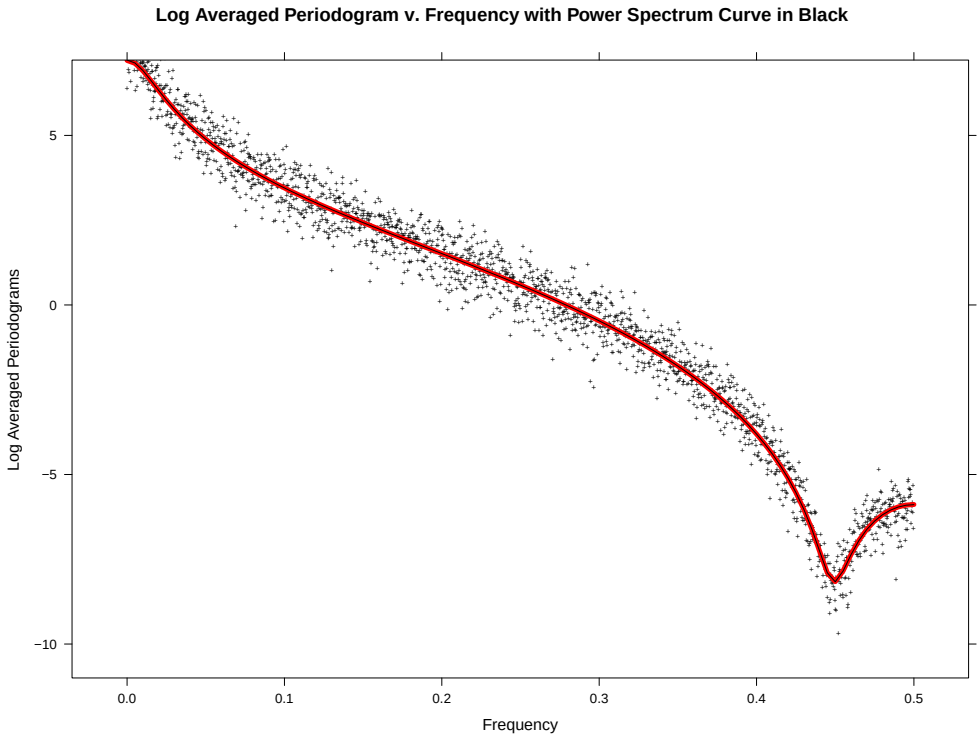

Fig. 5.19. ARMA(p = 1, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

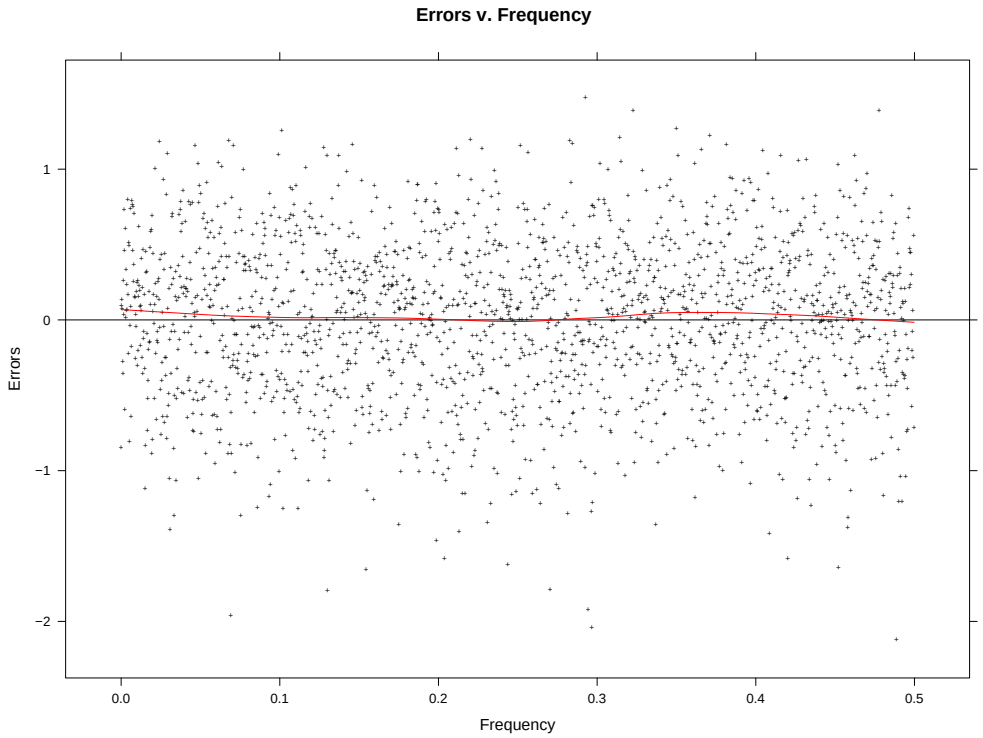

Fig. 5.20. ARMA(p = 1, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

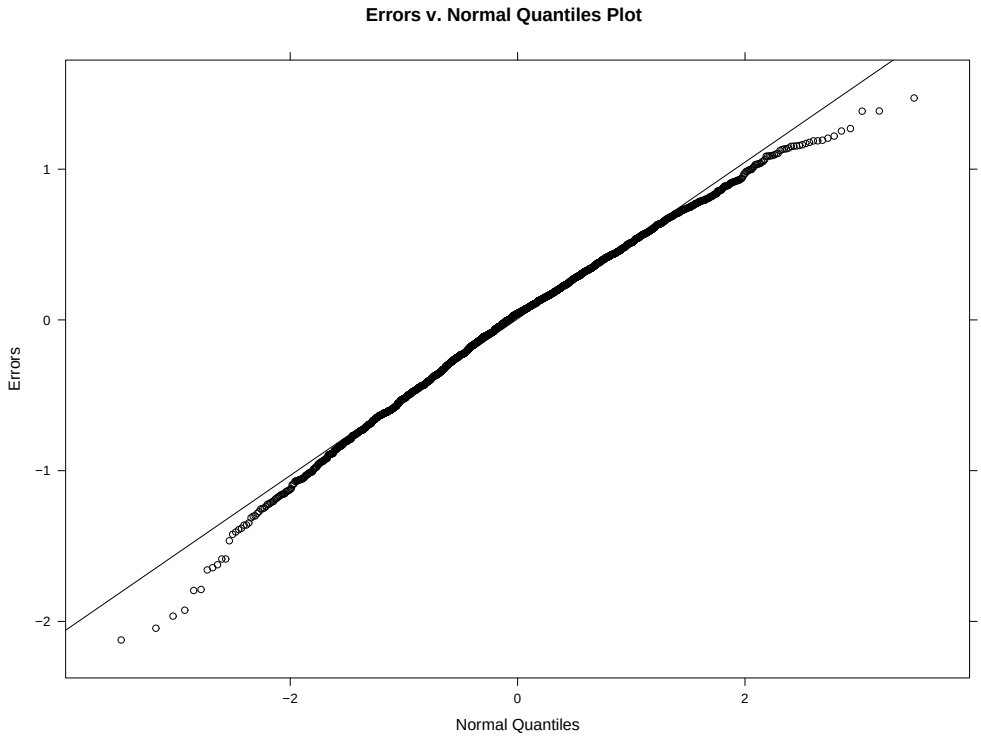

Fig. 5.21. ARMA(p = 1, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

5.1.8 ARMA(p = 2, q = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

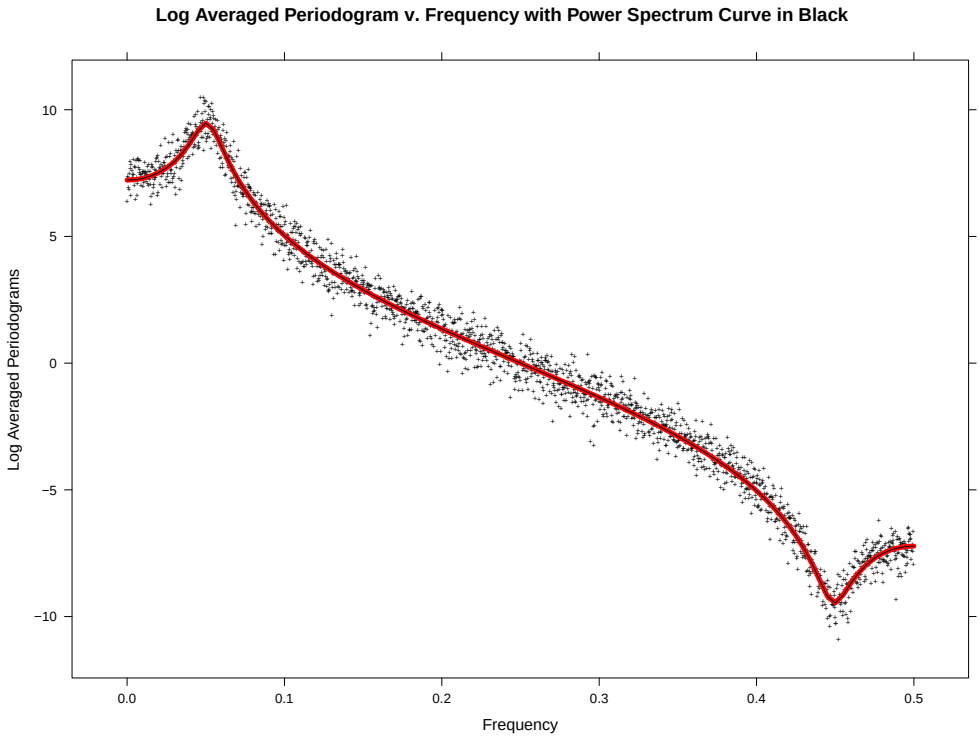

Fig. 5.22. ARMA(p = 2, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

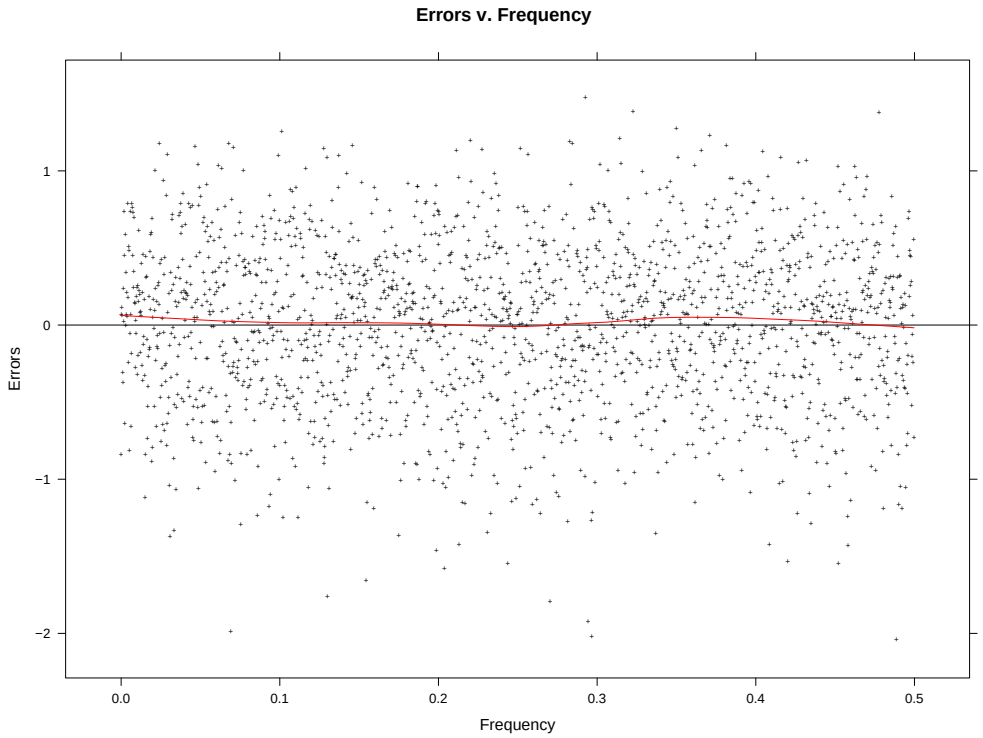

Fig. 5.23. ARMA(p = 2, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

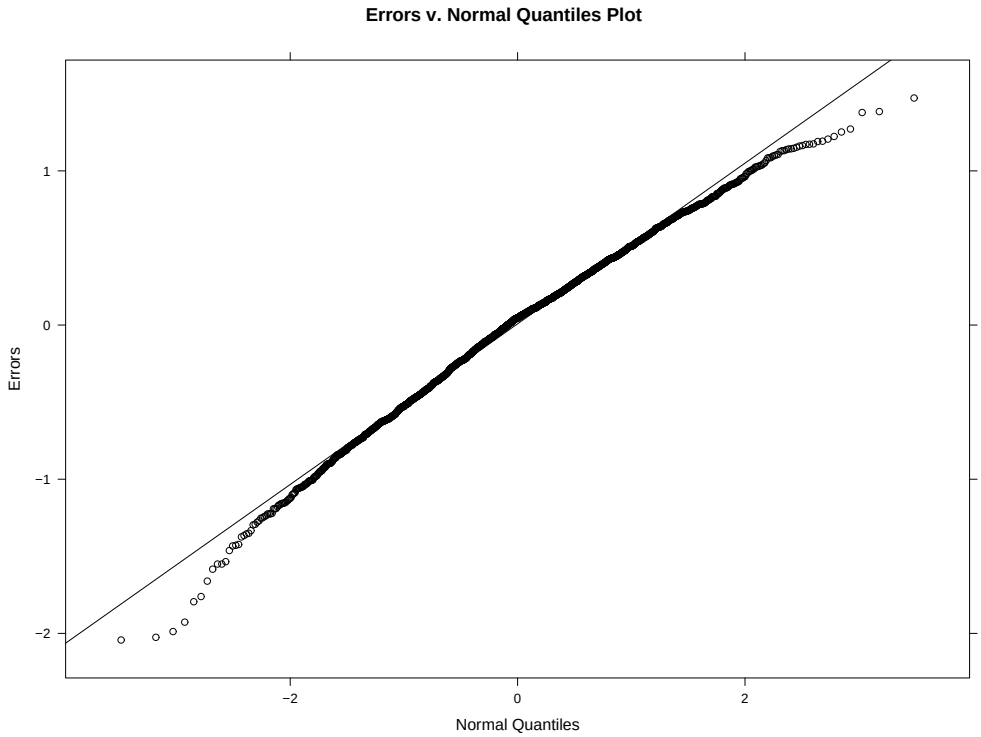

Fig. 5.24. ARMA(p = 2, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals

The residuals are normally distributed.

5.2 FARMA Models

5.2.1 ARFIMA(p = 0, d, q = 0)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

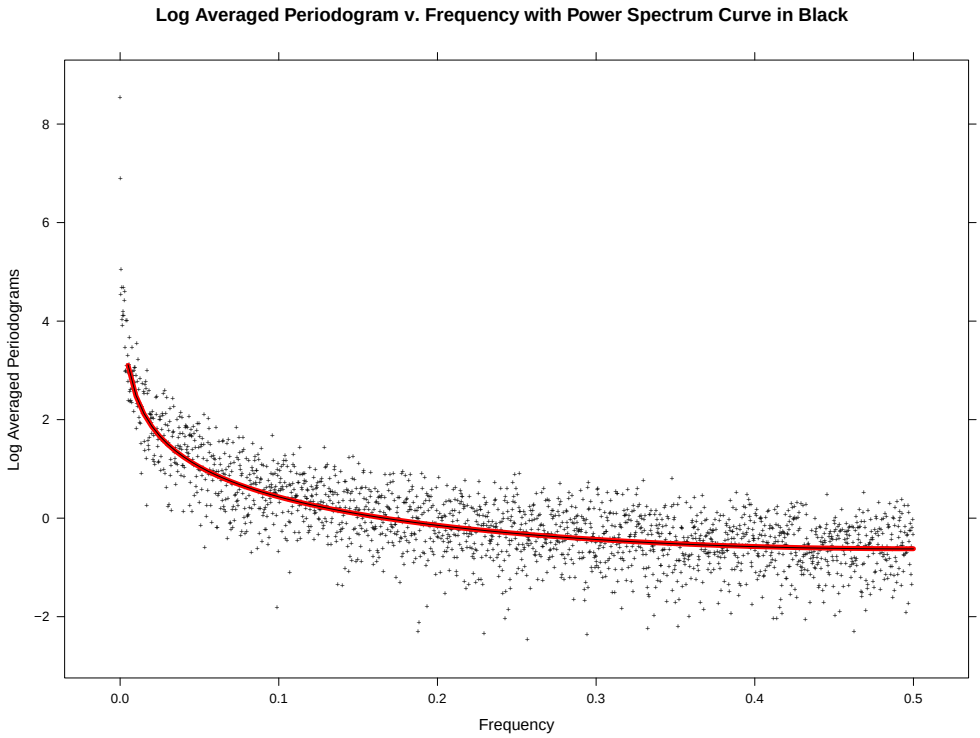

Fig. 5.25. FARMA(p = 0, d, q = 0) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

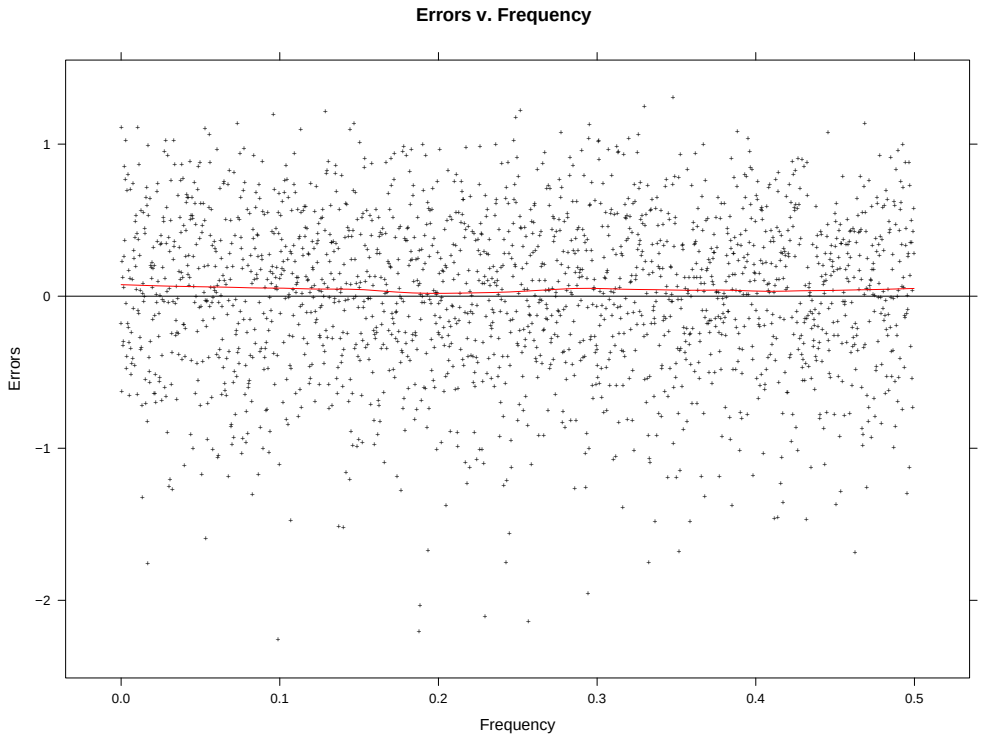

Fig. 5.26. FARMA(p = 0, d, q = 0) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

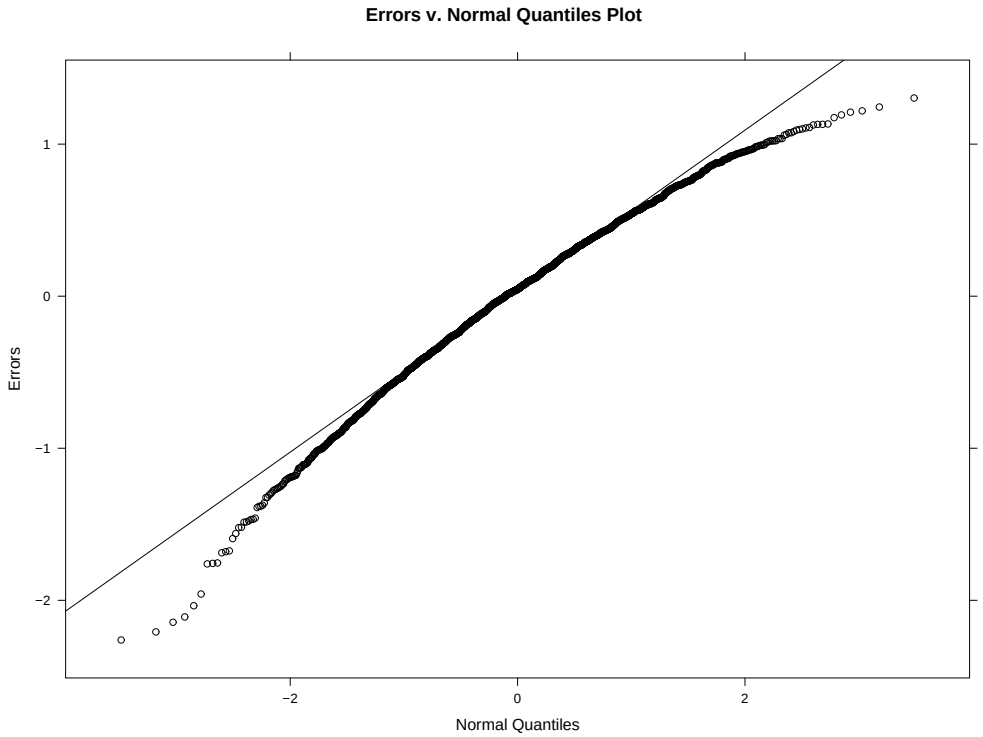

Fig. 5.27. FARMA(p = 0, d, q = 0) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.2 ARFIMA(p = 1, d, q = 0)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

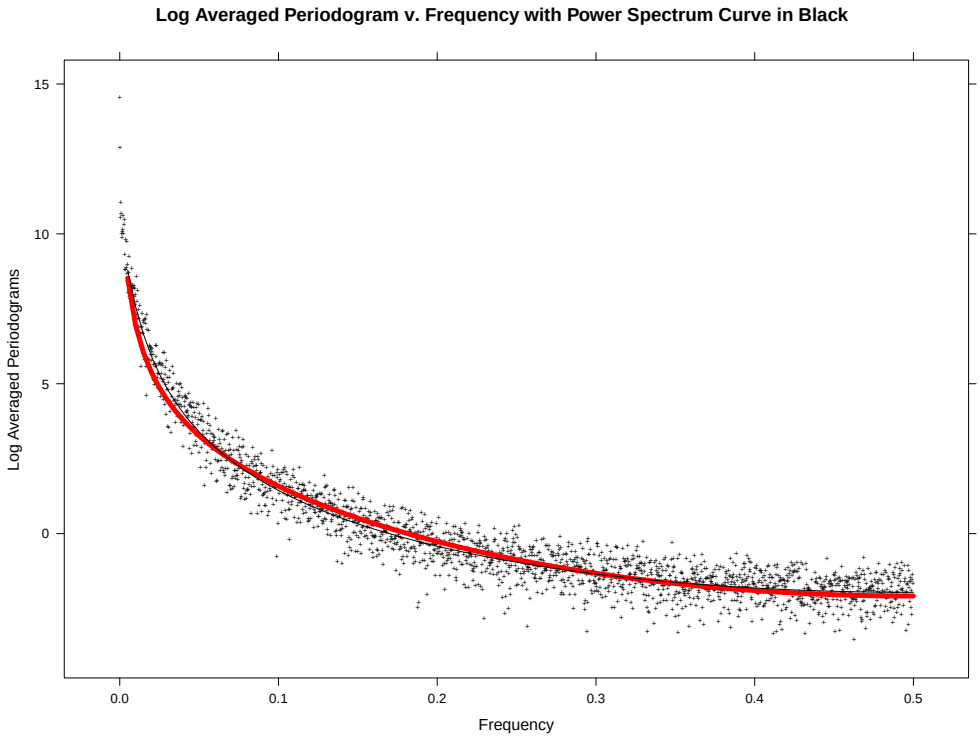

Fig. 5.28. FARMA(p = 1, d, q = 0) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

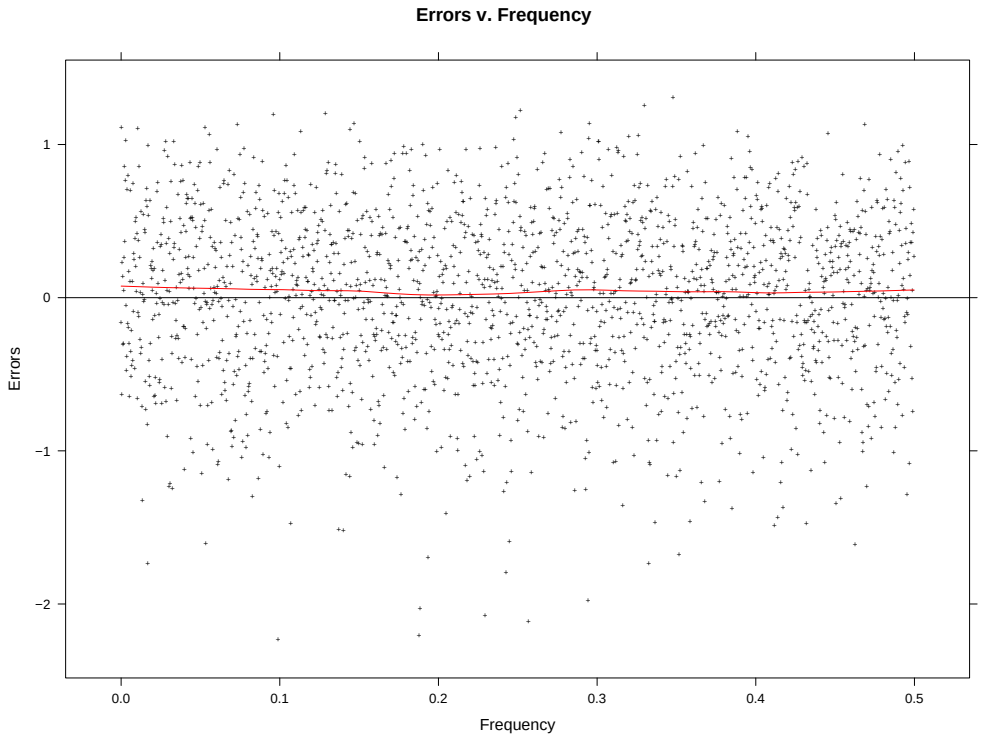

Fig. 5.29. FARMA(p = 1, d, q = 0) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

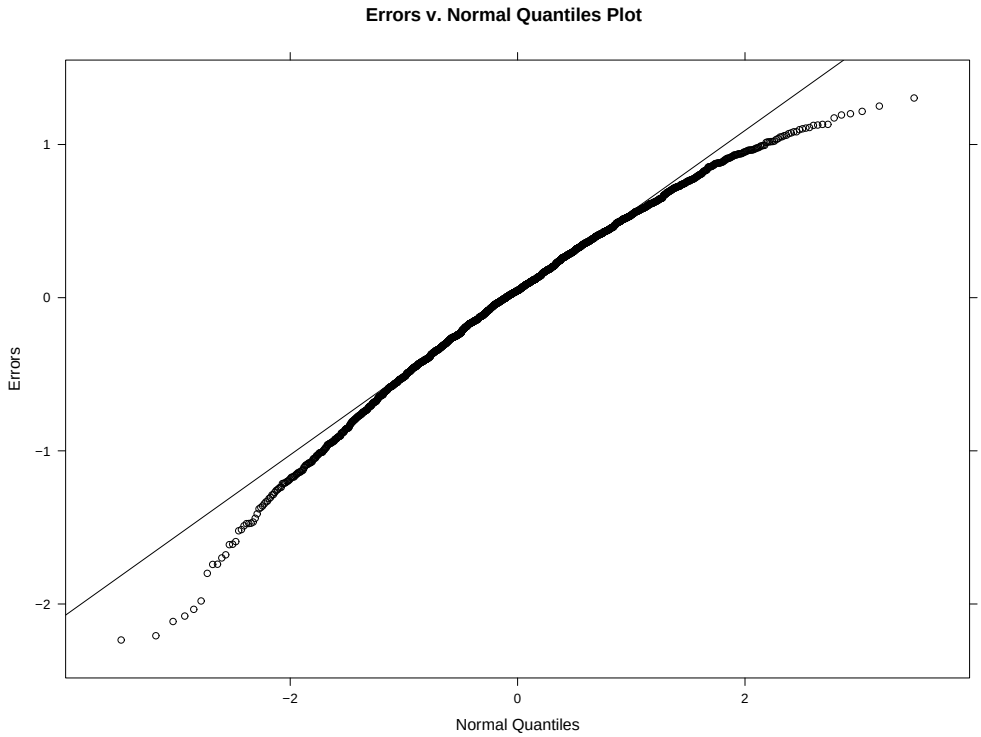

Fig. 5.30. FARMA(p = 1, d, q = 0) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.3 ARFIMA(p = 2, d, q = 0)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

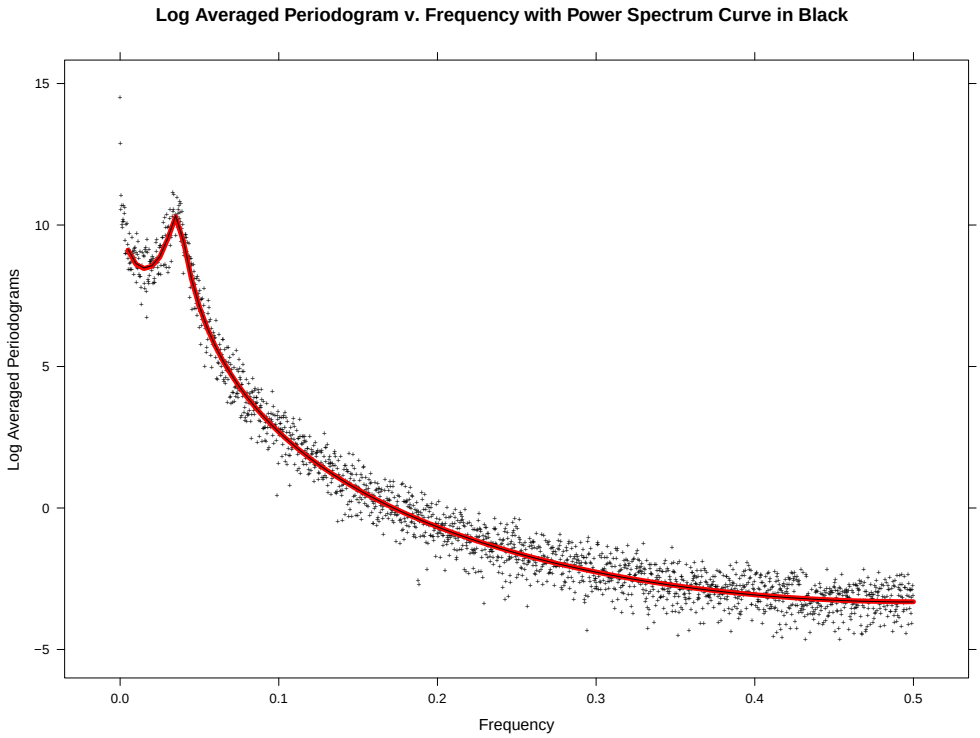

Fig. 5.31. FARMA(p = 2, d, q = 0) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

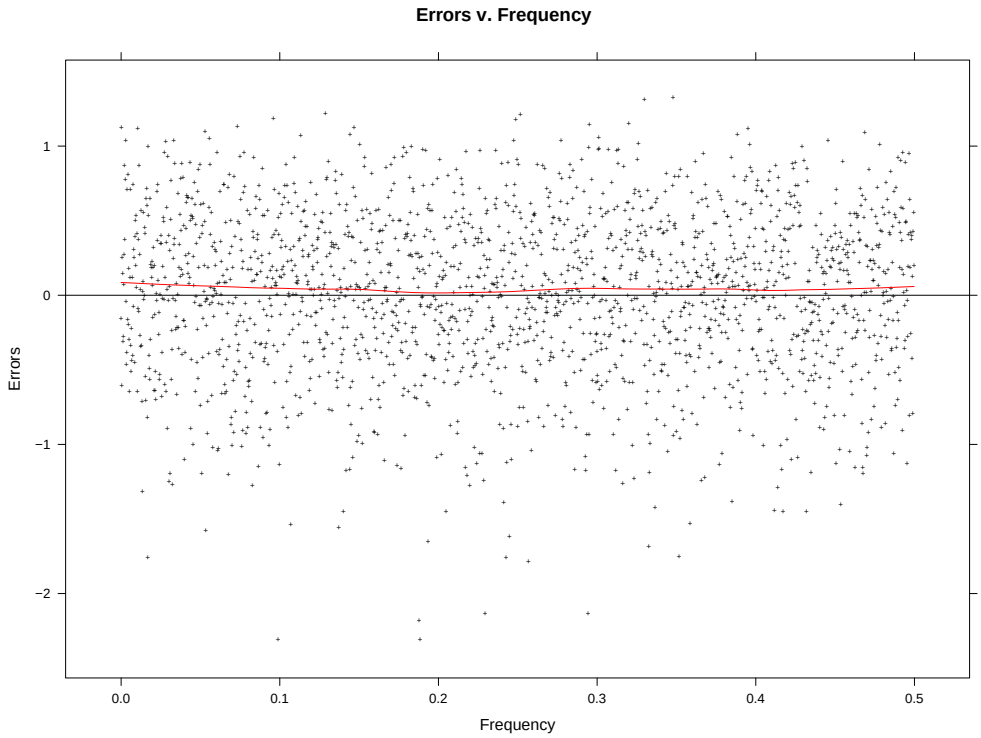

Fig. 5.32. FARMA(p = 2, d, q = 0) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

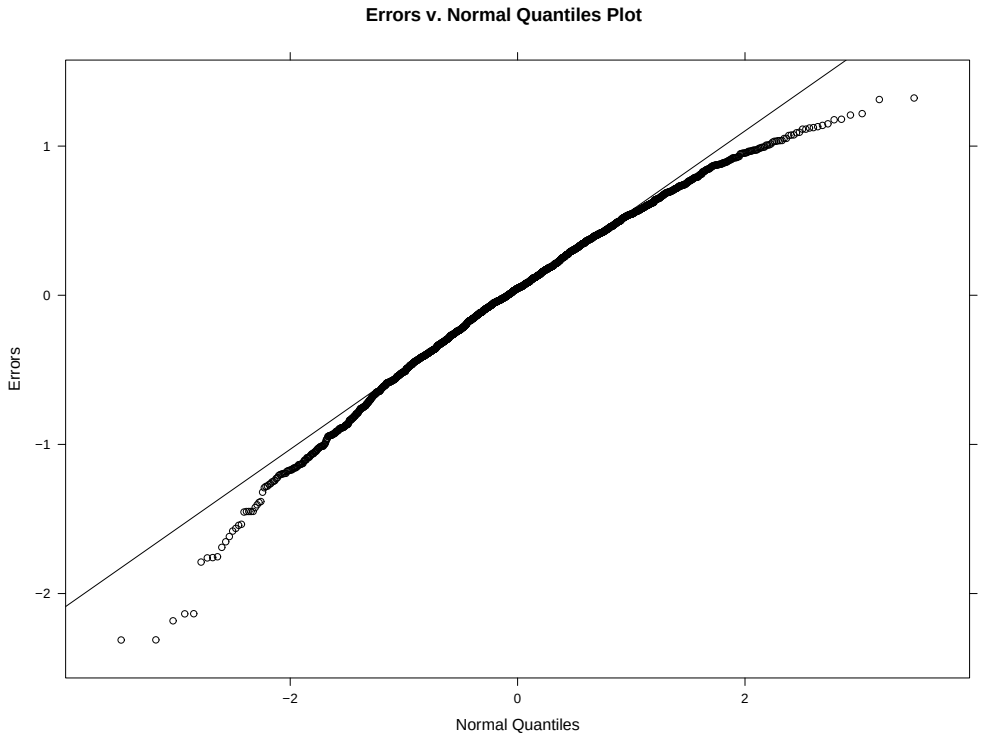

Fig. 5.33. ARFIMA(p = 2, d, q = 0) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.4 ARFIMA(p = 0, d, q = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

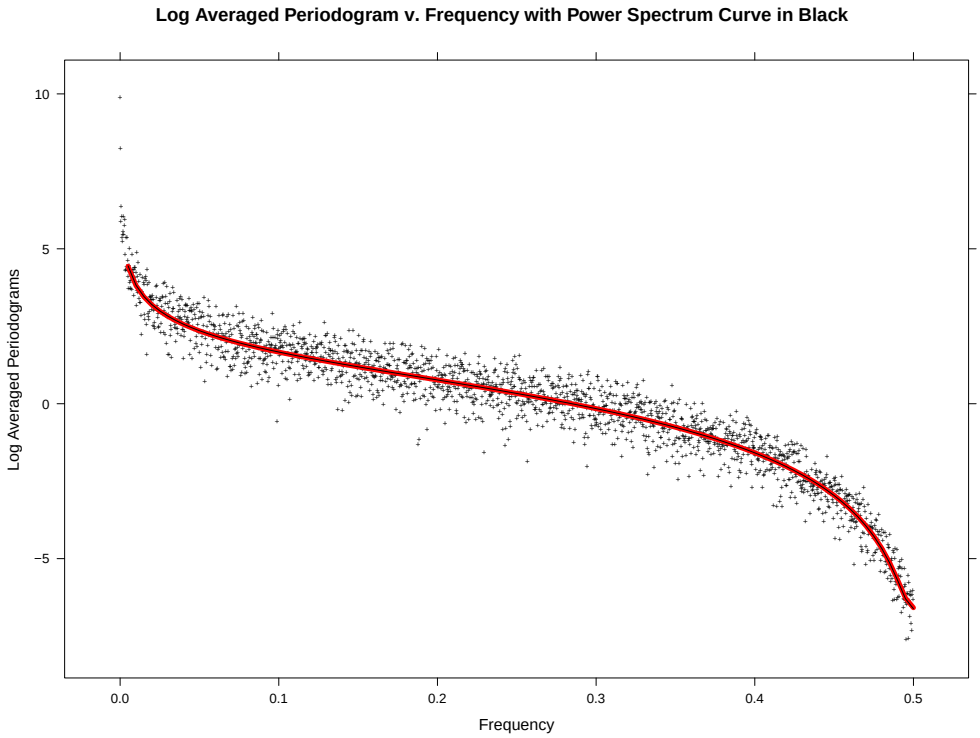

Fig. 5.34. FARMA(p = 0, d, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

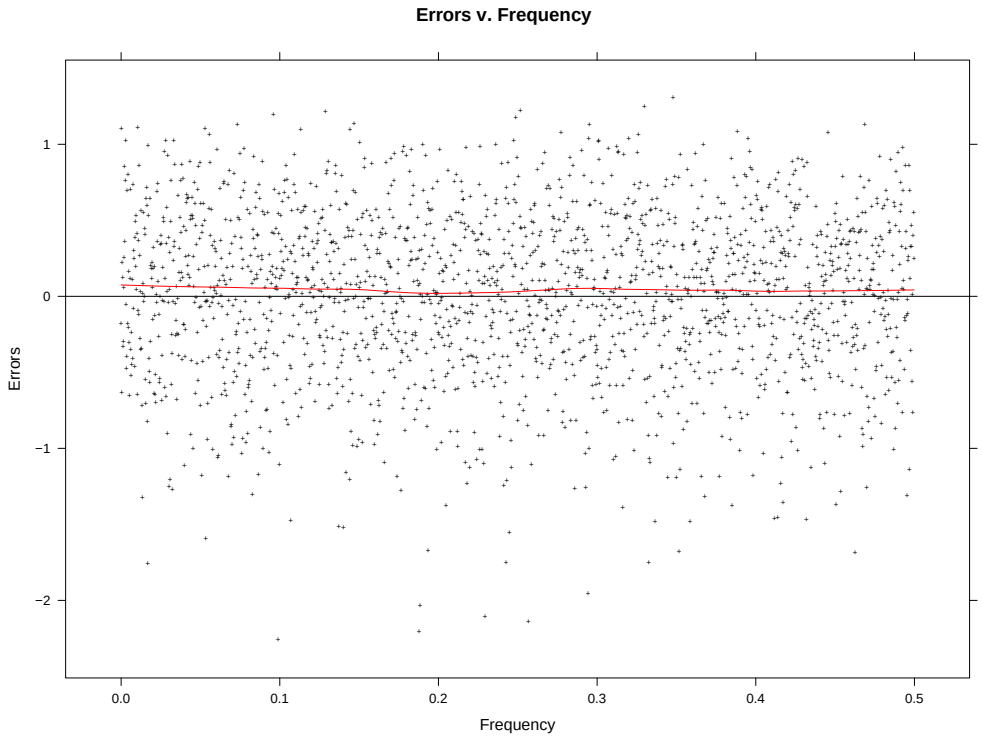

Fig. 5.35. FARMA(p = 0, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

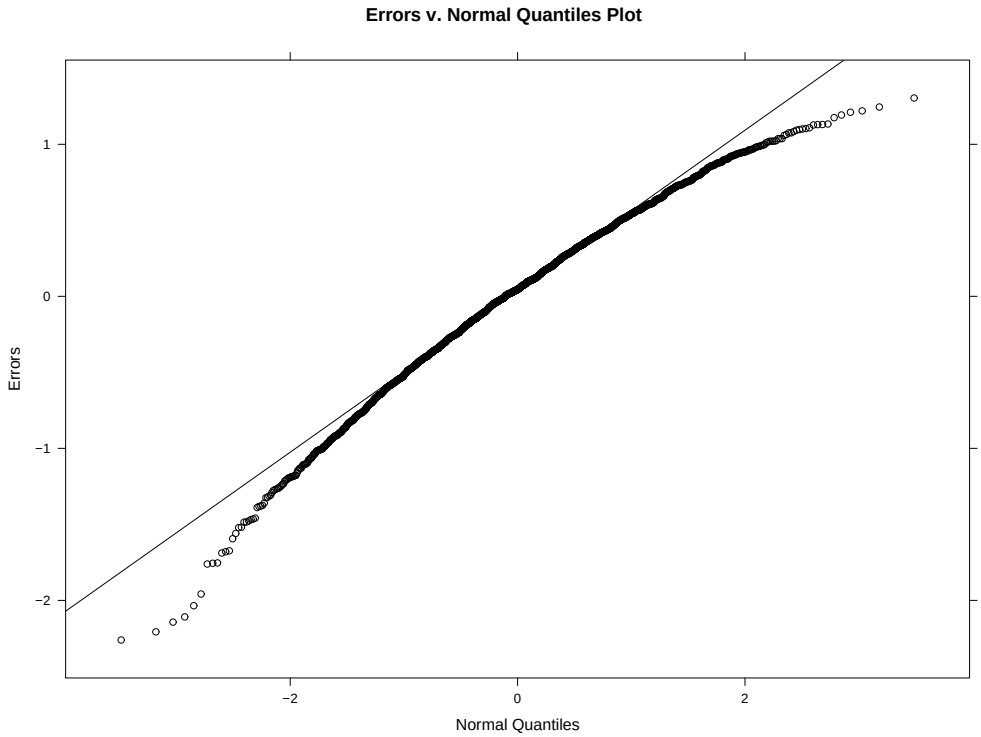

Fig. 5.36. FARMA(p = 0, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.5 ARFIMA(p = 0, d, q = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

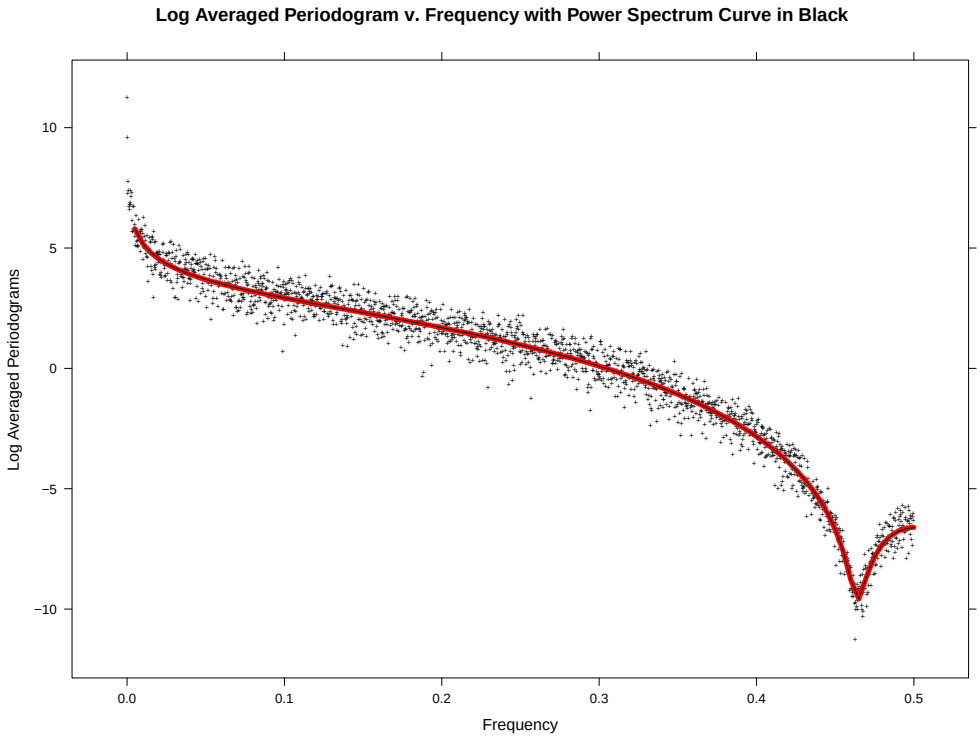

Fig. 5.37. FARMA(p = 0, d, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

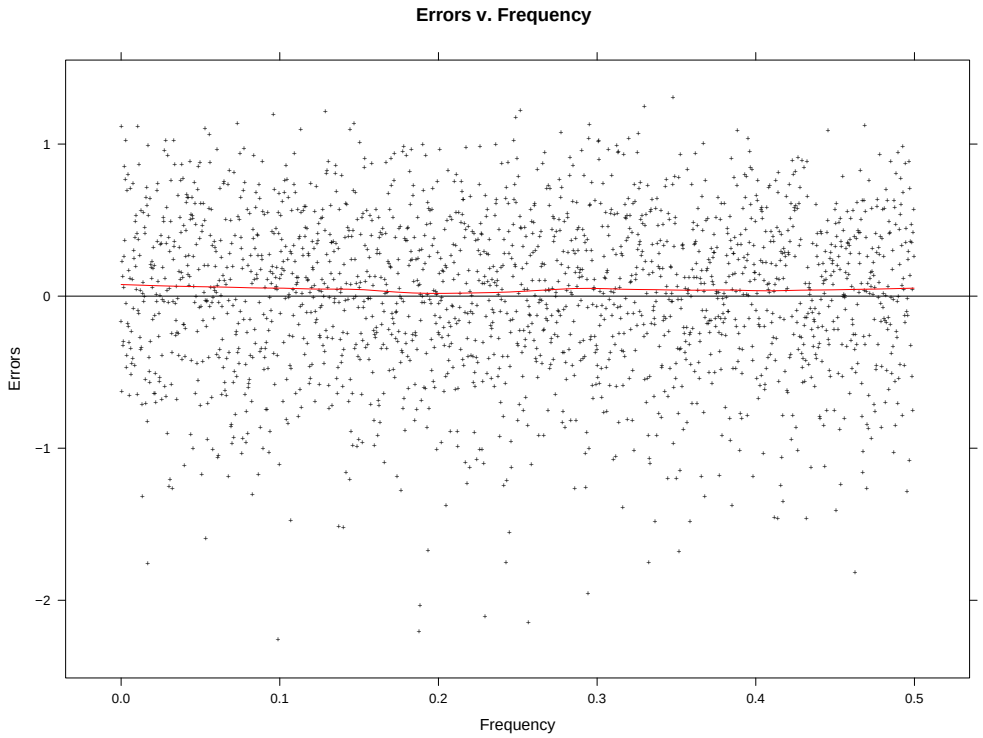

Fig. 5.38. FARMA(p = 0, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, Quantile-Quantile plots were constructed:

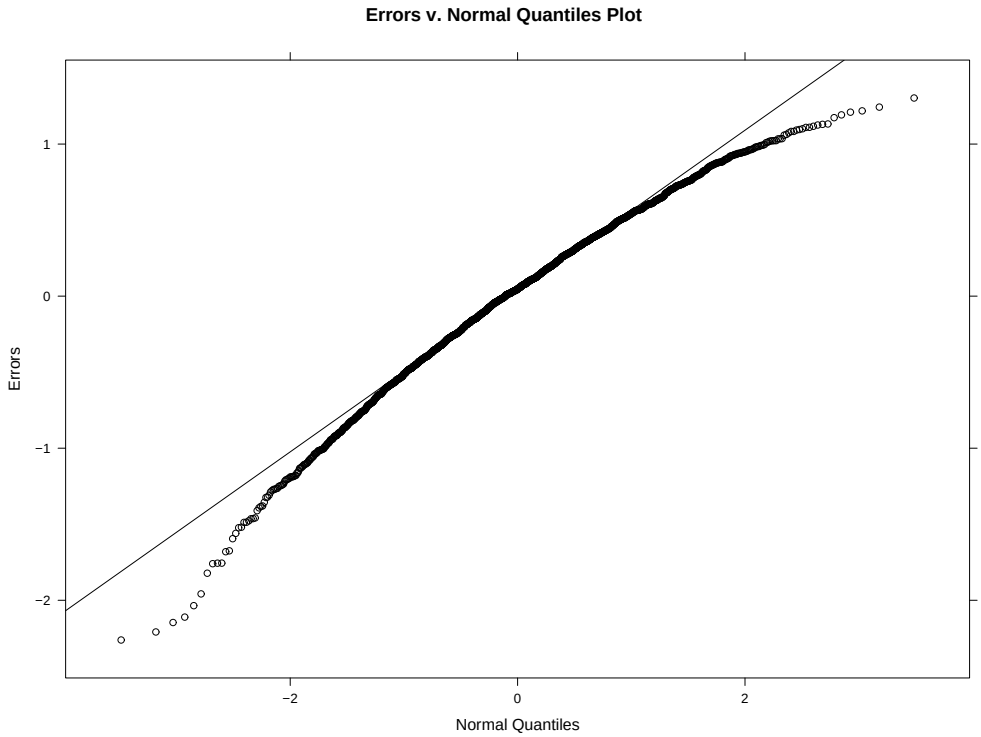

Fig. 5.39. FARMA(p = 0, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.6 ARFIMA(p = 1, d, q = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

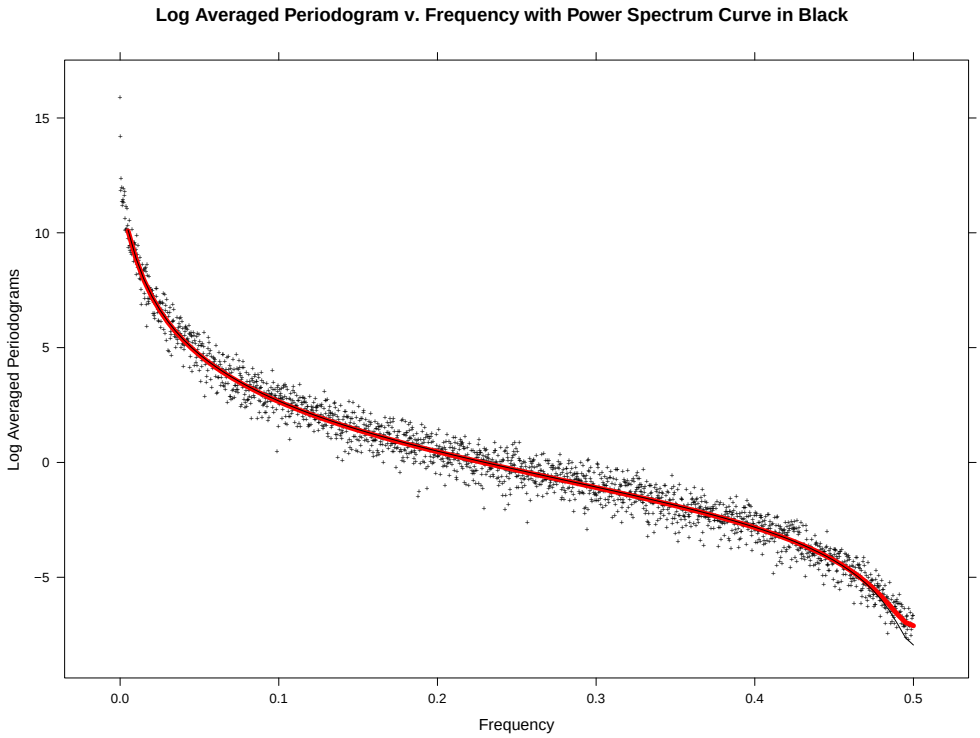

Fig. 5.40. FARMA(p = 1, d, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

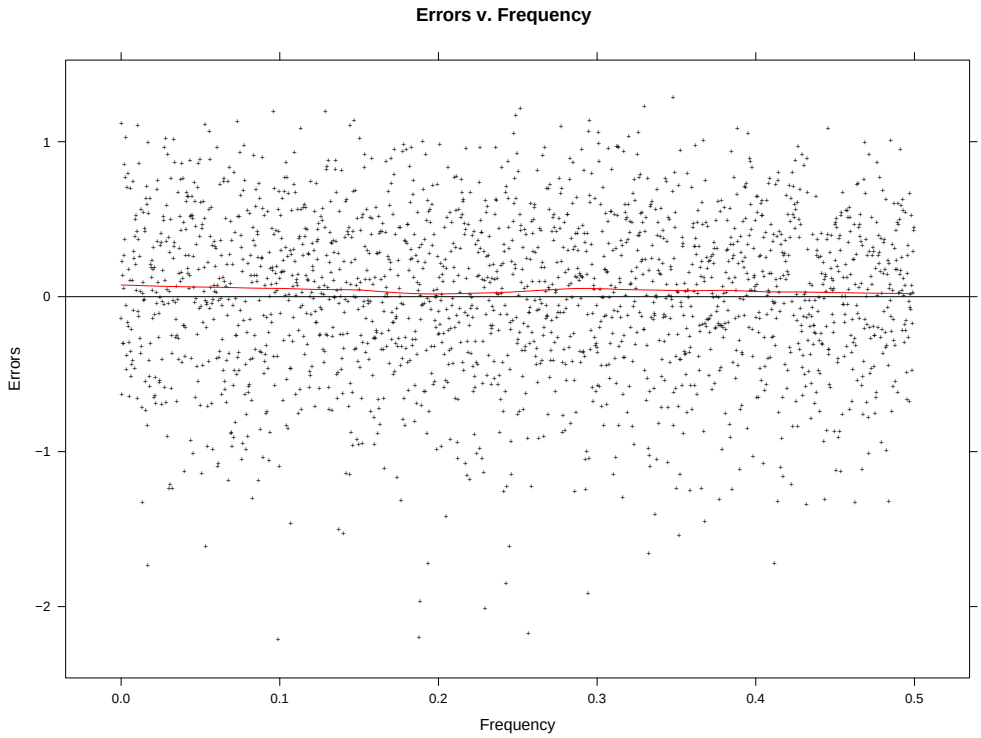

Fig. 5.41. FARMA(p = 1, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

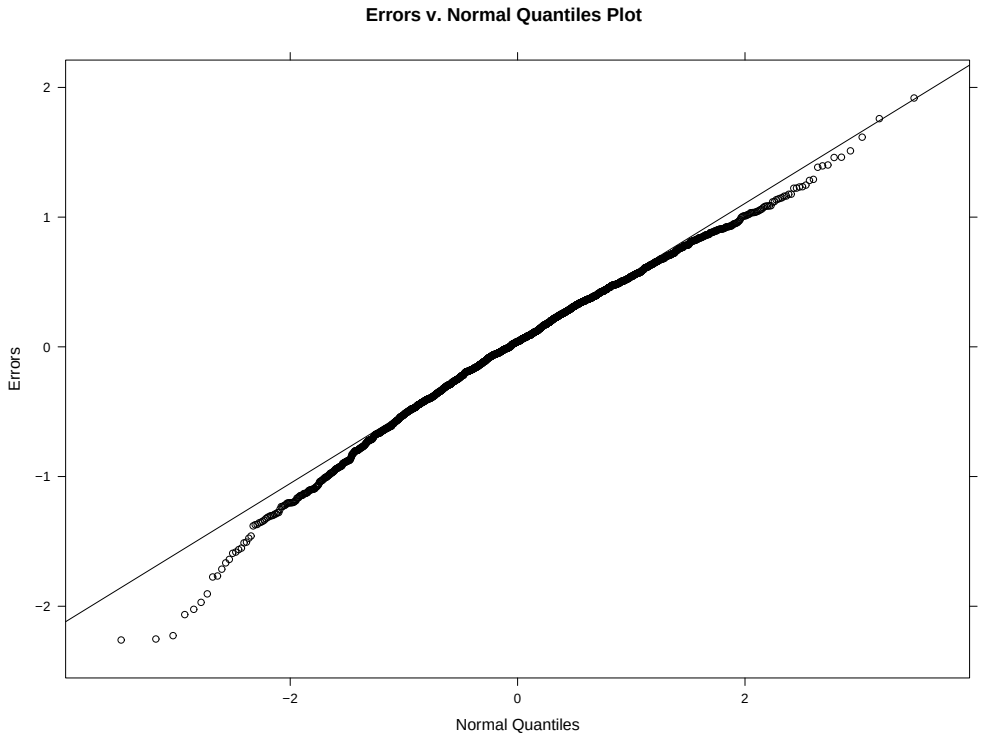

Fig. 5.42. FARMA(p = 1, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.7 ARFIMA(p = 2, d, q = 1)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

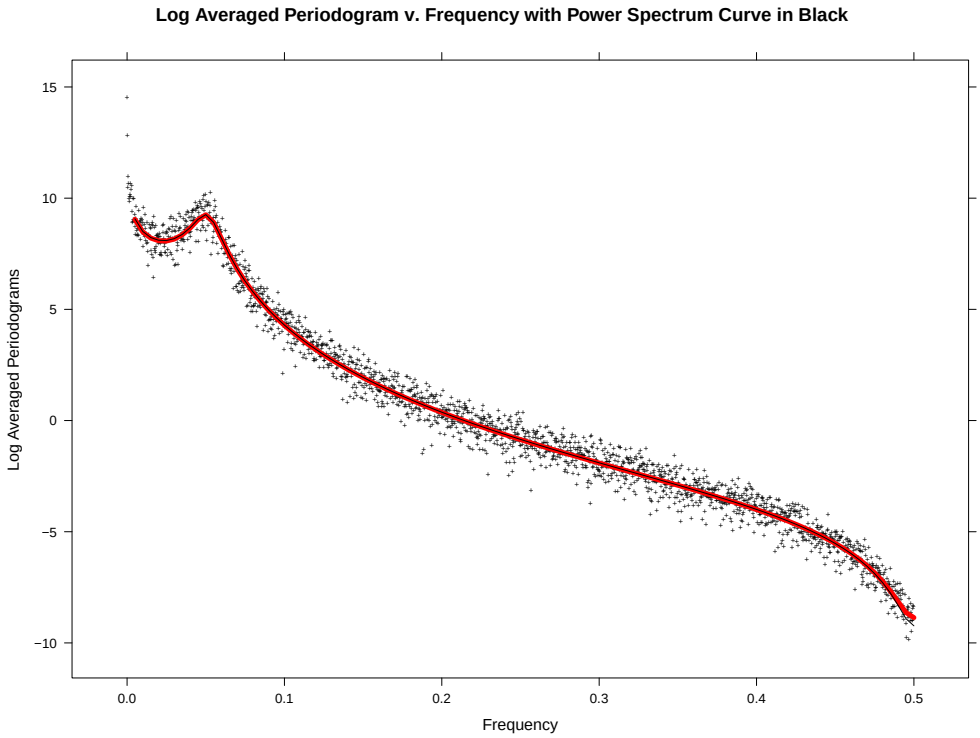

Fig. 5.43. FARMA(p = 2, d, q = 1) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

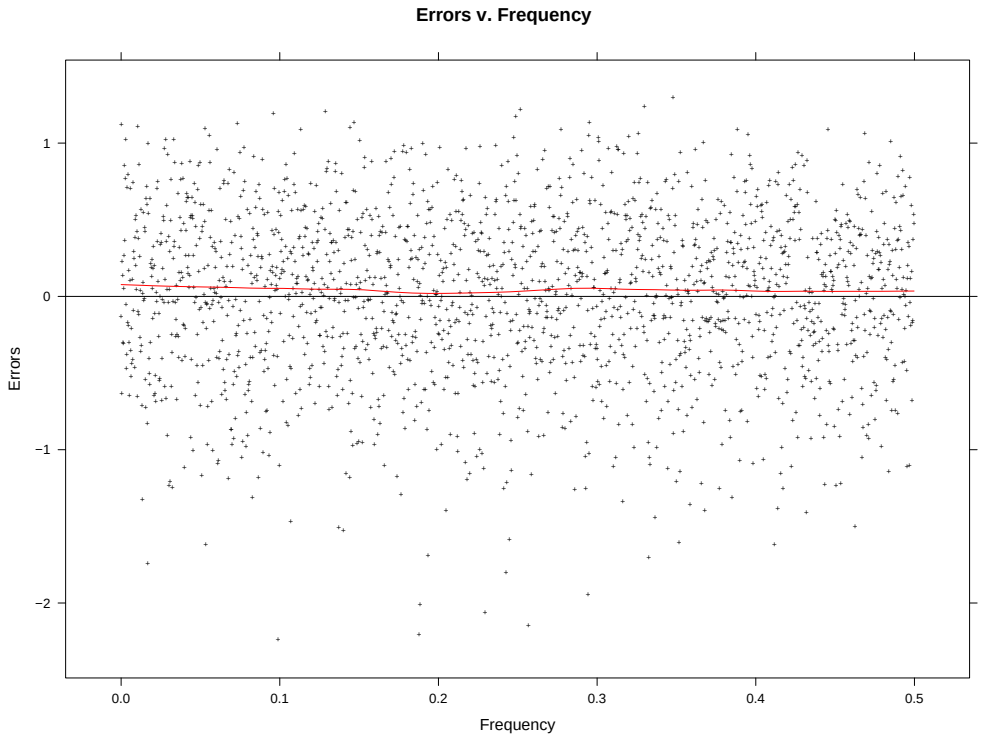

Fig. 5.44. FARMA(p = 2, d, q = 1) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

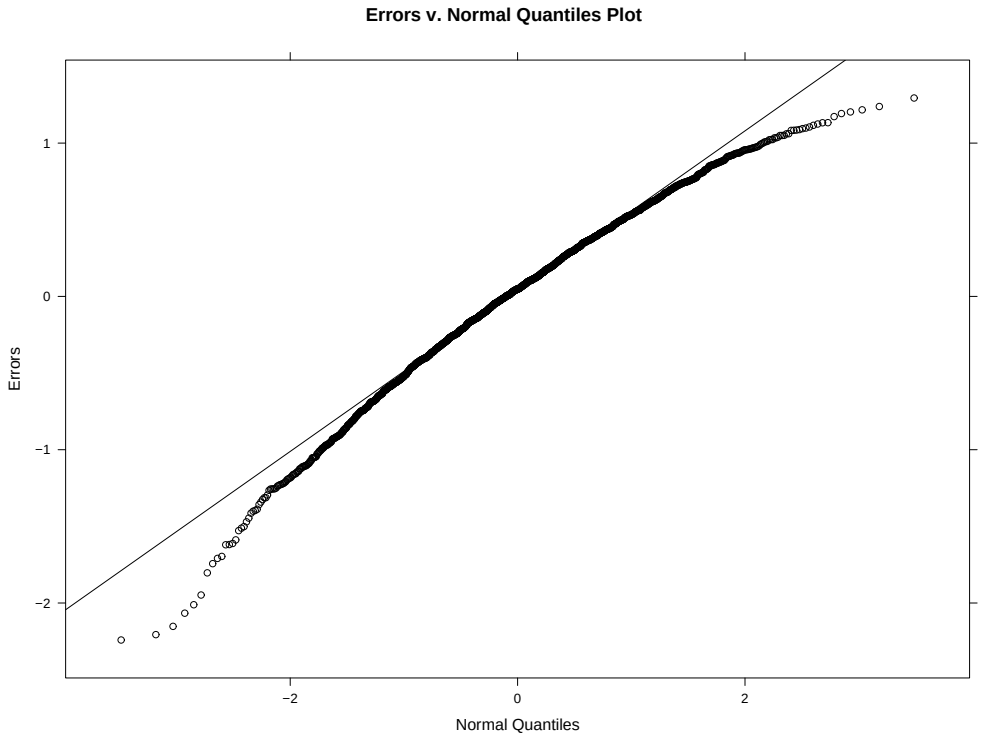

Fig. 5.45. FARMA(p = 2, d, q = 1) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.8 ARFIMA(p = 1, d, q = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

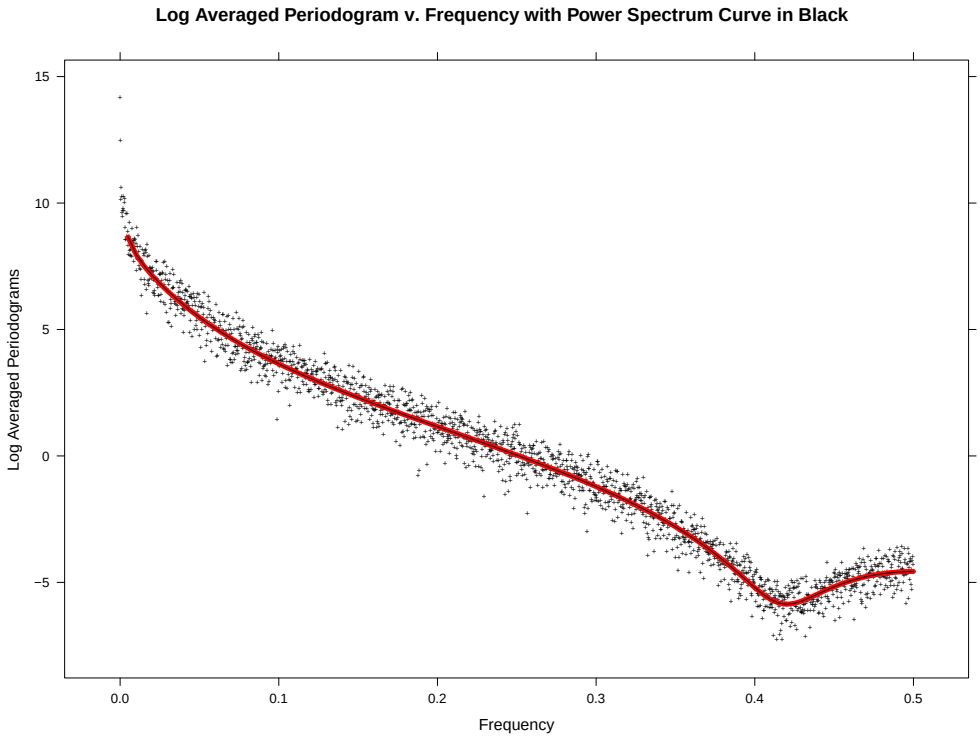

Fig. 5.46. FARMA(p = 1, d, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

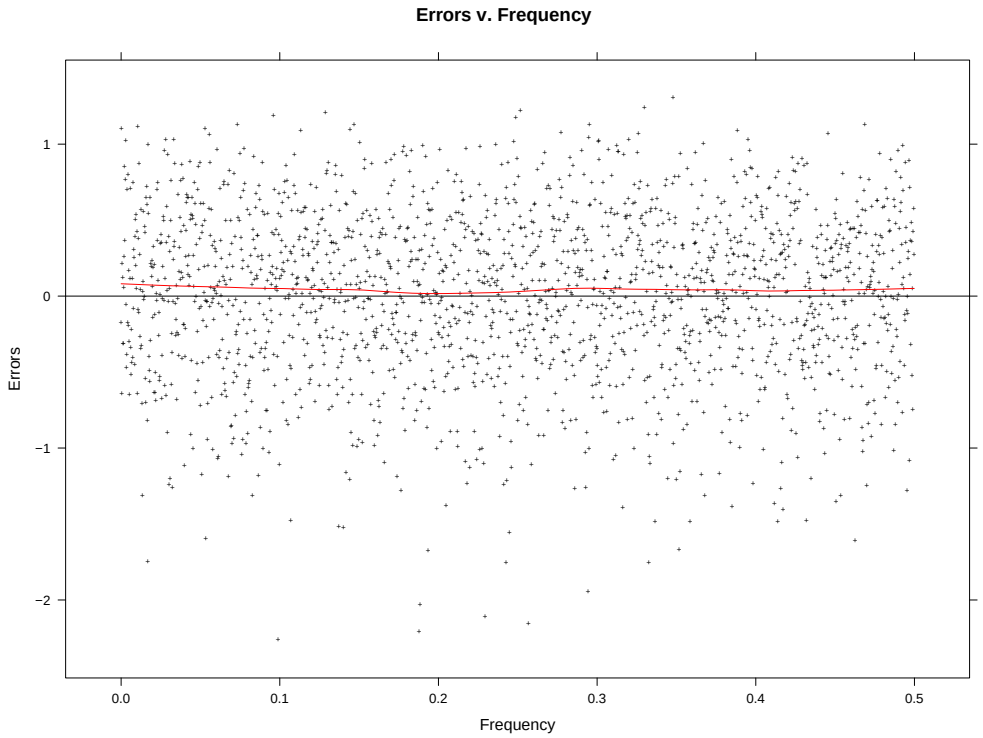

Fig. 5.47. FARMA(p = 1, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

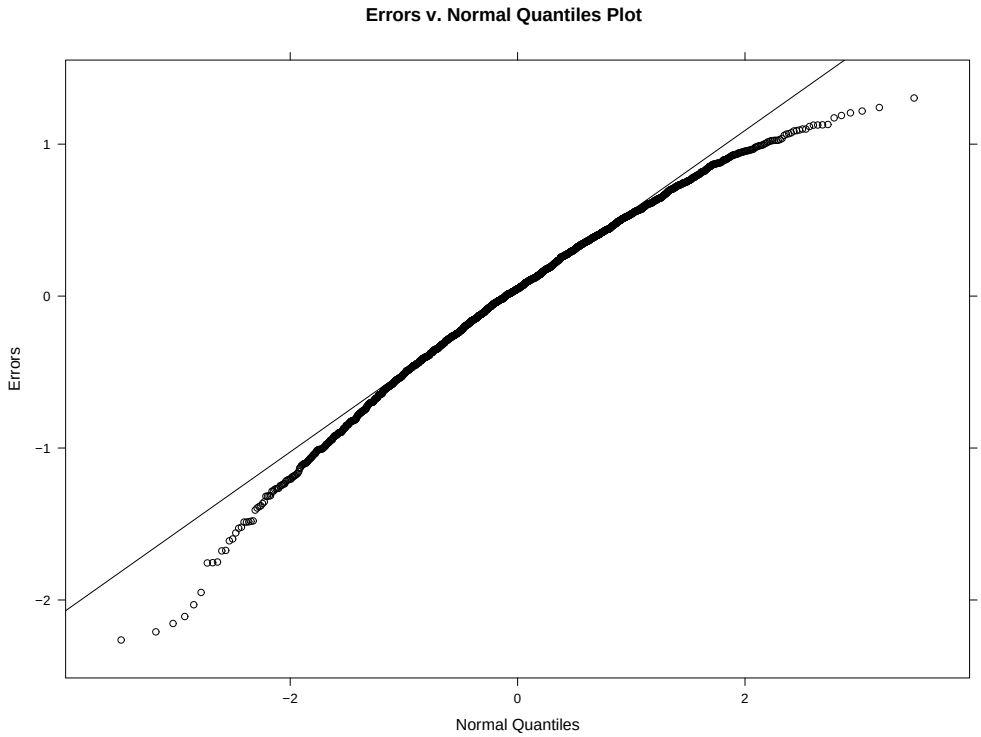

Fig. 5.48. FARMA(p = 1, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

5.2.9 ARFIMA(p = 2, d, q = 2)

The PPS-REG log power spectrum estimation yielded the following log averaged periodogram versus frequency results:

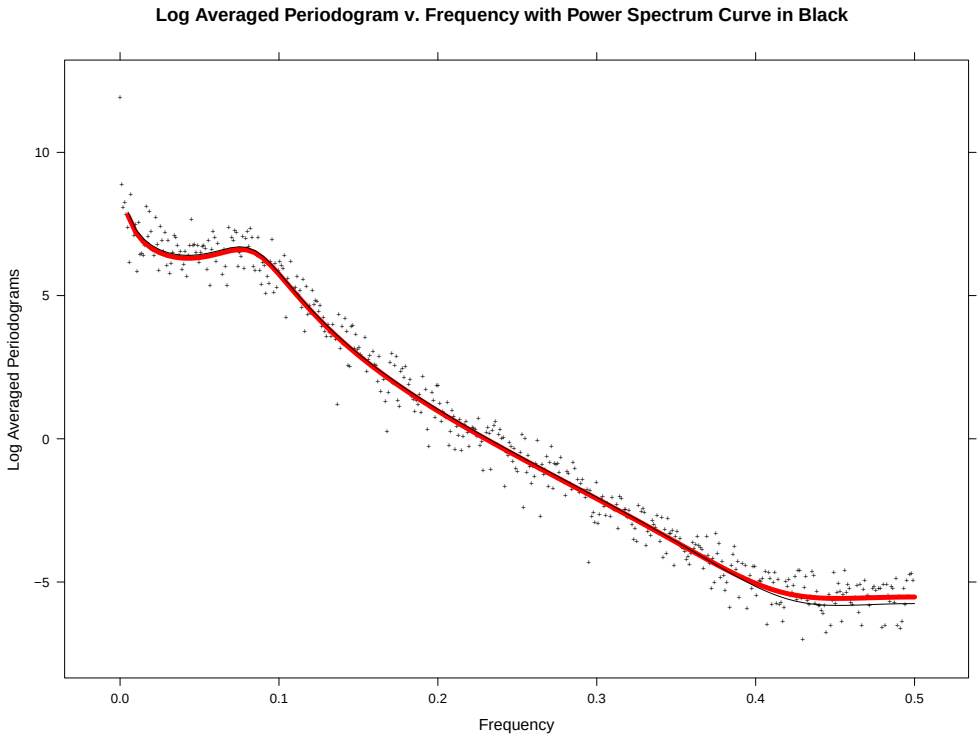

Fig. 5.49. FARMA(p = 2, d, q = 2) Log of the Averaged Periodogram v. Frequency, Estimate

The PPS-REG estimated log power spectrum is consistent with the true log power spectrum.

To assess the fit, the difference between the log averaged periodogram and the PPS-REG estimated log power spectrum versus frequency yielded the following residuals plot:

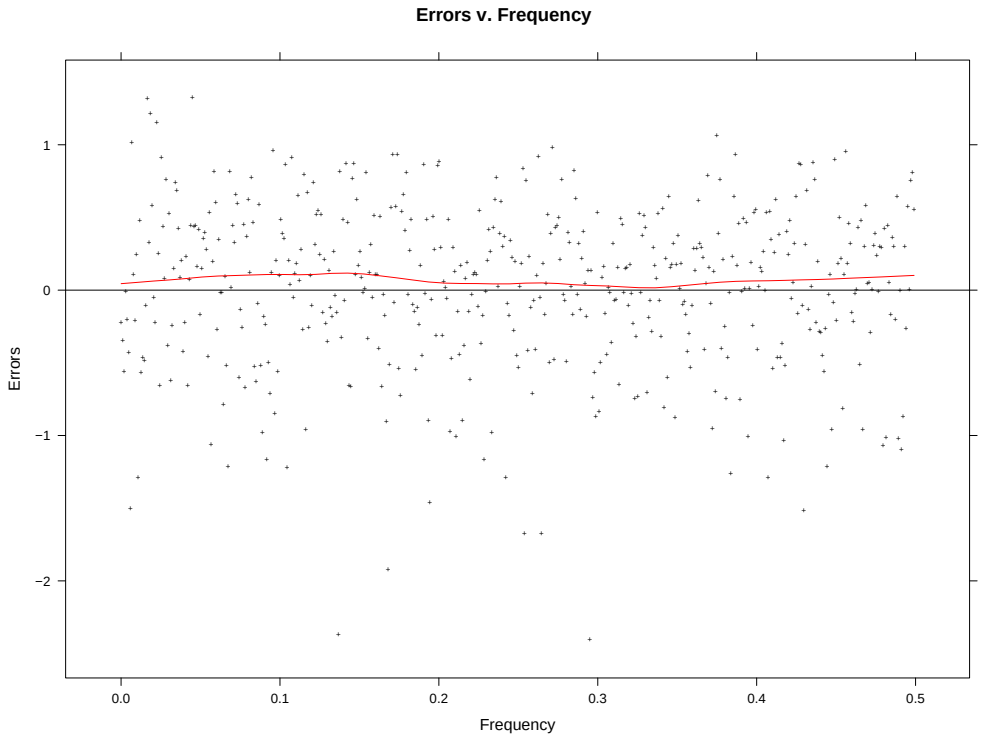

Fig. 5.50. FARMA(p = 2, d, q = 2) Log of the Averaged Periodogram Residuals v. Frequency, Estimate

The estimate residuals do not possess a lack of fit, which is supported diagnostically with a LOESS curve of degree one and span one third.

To assess the normality of the residuals from our estimated log power spectrum, QQ plots were constructed:

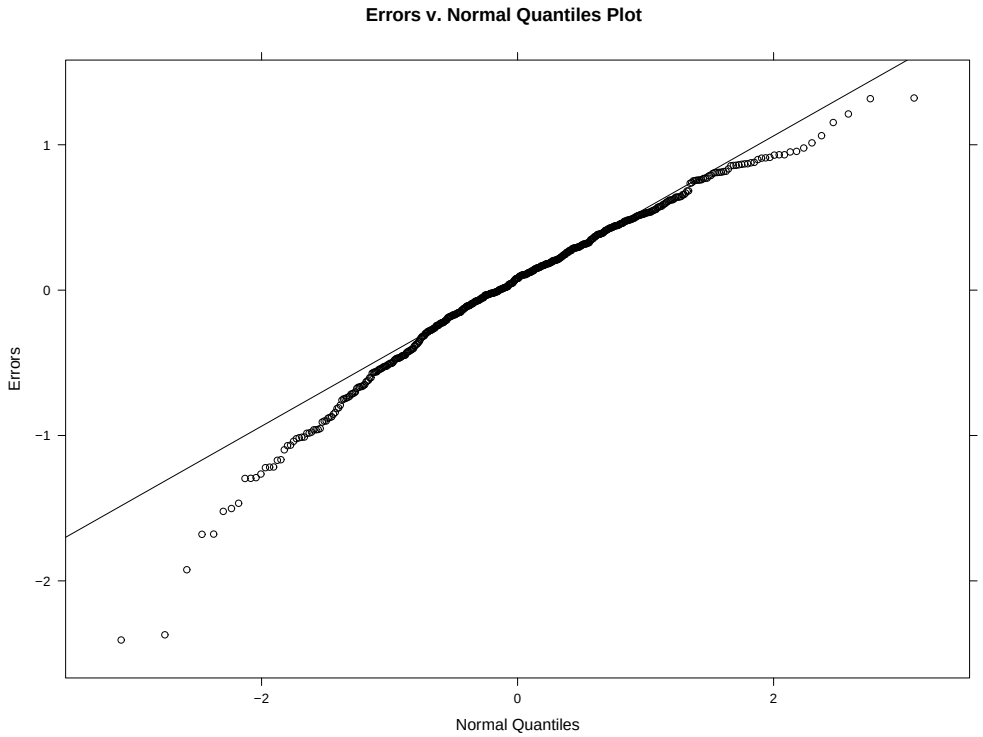

Fig. 5.51. FARMA(p = 2, d, q = 2) QQ-Plot of the Log of the Averaged Periodogram Residuals, Estimate

The residuals are normally distributed.

REFERENCES

REFERENCES

[1] C. L. M. William S. Cleveland and J. E. McRae, "ATS Methods: Nonparametric Regression for Non-Gaussian Data." American Statistical Association (ASA), 1993, pp. 821–835.

[2] D. R. Brillinger, "Time Series: Data Analysis and Theory." New York: Holt, Rinehart, and Winston, 1975.

[3] D. Sarkar, *Lattice: Multivariate Data Visualization with R*. New York: Springer, 2008, iSBN 978-0-387-75968-5. [Online]. Available: http://lmdvr.r-forge.r-project.org

[4] W. S. Cleveland and S. J. Devlin, "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting." Journal of the American Statistical Association (ASA), 1988, vol. 83, No. 403, pp. 596–610.

[5] M. A. Hauser, "Maximum liklihood estimators for ARMA and ARFIMA models: a Monte Carlo study." Journal of Statistical Planning and Inference, 1999, vol. 80, pp. 229–255.

[6] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: https://www.R-project.org/

[7] B. D. Ripley, *Time series in R 1.5.0.*, 2002, r package version 1.5.0. [Online]. Available: https://www.r-project.org/doc/Rnews/Rnews_2002-2.pdf

[8] S. original by Chris Fraley, U.Washington, S. R. port by Fritz Leisch at TU Wien; since 2003-12: Martin Maechler; fdGPH, fdSperio, etc by Valderio Reisen, and A. Lemonte., *fracdiff: Fractionally differenced ARIMA aka ARFIMA(p,d,q) models*, 2012, r package version 1.4-2. [Online]. Available: https://CRAN.R-project.org/package=fracdiff

[9] S. L. M. Jr., "Digital spectral analysis." Prentice Hall, 1987.

VITA

VITA

Jeremy Troisi was born in Lansing, MI. He earned a B.A. in Mathematics and Economics & Management from Albion College in May of 2008 before earning an M.S. in Statistics from Purdue University in May of 2010. Since September 2017 he has been working for Booz Allen Hamilton, a consulting firm, out of their Saint Louis, MO regional offices from where he remotely finished his PhD dissertation. His interests include soccer, yoga, running, and weight training.