Electronic Journal of Statistics
Vol. 0 (2017) 1–25
ISSN: 1935-7524

Parameter Estimation for Tempered Fractional Time Series

Farzad Sabzikar

Department of Statistics,
Iowa State University,
Ames, IA 50011
USA
e-mail: sabzikar@iastate.edu
url: sabzikar.public.iastate.edu/

A. Ian McLeod

Department of Statistical and Actuarial Sciences,
Western University,
London, Ontario N6A 5B7
Canada
e-mail: aimcleod@uwo.ca
url: www.stats.uwo.ca/faculty/aim/

Mark M. Meerschaert

Department of Statistics and Probability,
Michigan State University,
East Lansing MI 48823
USA
e-mail: mcubed@stt.msu.edu
url: www.stt.msu.edu/users/mcubed/

Abstract: This paper develops the Whittle and maximum likelihood estimators for tempered fractional time series. These time series exhibit semi-long range dependence and are useful in modeling geophysical turbulence.

MSC 2010 subject classifications: Primary 62M10, 60G12; secondary 62P12.
Keywords and phrases: power-law spectrum, geophysical time series, Whittle and maximum likelihood estimator, R software package.

Contents

1. Introduction

Fractionally integrated time series have proven useful for modeling long range dependence [6, 12]. Recently, a model extension has been proposed based on tempered fractional calculus [43]. The autoregressive fractionally integrated moving average (ARFIMA) time series $X_t = \Delta^d Y_t$ where $\Delta^d = (I - B)^d$ using the backward shift operator $BY_t = Y_{t-1}$, and Y_t is a classical autoregressive moving average (ARMA) time series. Similarly, the autoregressive tempered fractionally integrated moving average (ARTFIMA) time series is defined by $X_t = \Delta^{d,\lambda} Y_t$ where Y_t is an ARMA time series, and the tempered fractional difference operator $\Delta^{d,\lambda} = (I - e^{-\lambda} B)^d$. The ARTFIMA model is a short range dependent time series that nonetheless exhibits *semi-long range dependence* [43, Eq. (30)]: For small values of the tempering parameter $\lambda > 0$, the spectral density grows like a power law at low frequencies, but remains bounded as the frequency tends to zero. The same behavior is commonly observed in geophysical turbulence, and it was demonstrated in [33] that the ARTFIMA model provides a good fit to such data. In particular, the ARTFIMA model reproduces Kolmogorov scaling in the inertial range, but also fits the low frequency periodogram.

Fractional calculus was invented by Leibnitz, and recently has found numerous applications in science and engineering [30, 34, 35, 36, 37, 44]. Fractional derivatives are the limits of fractional difference quotients [30, Proposition 2.1], and similarly, tempered fractional derivatives are the limits of tempered fractional difference quotients [30, Proposition 7.8]. Fractional Brownian motion was invented by Kolmogorov [25] as a model for turbulence, using a spectral representation. Mandelbrot and Van Ness [29] gave the process its name, and wrote its moving average representation, a fractional integral in time of a continuous parameter Gaussian white noise. The simplest ARFIMA time series, a fractionally integrated Gaussian white noise, can be viewed as a fractional Brownian motion sampled at integer time points. Recently, a tempered fractional Brownian motion has been developed, as the tempered fractional integral of a Gaussian white noise [31, 32]. A tempered fractionally integrated Gaussian white noise can be obtained by sampling tempered fractional Brownian motion at integer time points.

This paper develops the basic theory and parameter estimation methods for ARTFIMA time series. Since these time series are mathematically short range dependent, the methods are classical. However, we develop explicit formulae needed for practical applications, including geophysical turbulence. This paper also documents an R package called `artfima` that is freely available on CRAN. An example application is included here to demonstrate the software, and to rigorously verify the results in [33].

2. The ARTFIMA model

In this section, define the ARTFIMA model, and develop some basic results including causality, invertibility, the covariance function and spectral density.

The tempered fractional difference operator is defined by:

$$\Delta^{d,\lambda} f(x) = (I - e^{-\lambda j} B)^d f(x) = \sum_{j=0}^{\infty} \omega_j^{d,\lambda} f(x-j) \tag{2.1}$$

where $d > 0$, $\lambda > 0$, and

$$\omega_j^{d,\lambda} := (-1)^j \binom{d}{j} e^{-\lambda j} \quad \text{where} \quad \binom{d}{j} = \frac{\Gamma(1+d)}{j!\Gamma(1+d-j)} \tag{2.2}$$

using the gamma function $\Gamma(d) = \int_0^\infty e^{-x} x^{d-1}\, dx$. Using the well-known property $\Gamma(d+1) = d\Gamma(d)$, we can extend (2.1) to non-integer values of $d < 0$. By a common abuse of notation, we call this a tempered fractional integral. If $\lambda = 0$, then equation (2.1) reduces to the usual fractional difference operator. See [32, 43] for more details.

Definition 2.1. The discrete time stochastic process $\{X_t\}_{t\in\mathbb{Z}}$ is called an *autoregressive tempered fractional integrated moving average* time series, denoted by ARTFIMA(p, λ, d, q), if $\{X_t\}$ is a stationary solution with zero mean of the tempered fractional difference equations

$$\Phi(B)\Delta_1^{d,\lambda} X_t = \Theta(B) Z_t, \tag{2.3}$$

where Z_t is a white noise sequence (i.i.d. with $\mathbb{E}[Z_t] = 0$ and $\mathbb{E}[Z_t^2] = \sigma^2$), $d \notin \mathbb{Z}$, $\lambda > 0$, and $\Phi(z) = 1+\phi_1 z+\phi_2 z^2+\ldots+\phi_p z^p$, and $\Theta(z) = 1+\theta_1 z+\theta_2 z^2+\ldots+\theta_q z^q$ are polynomials of degrees $p, q \geq 0$ with no common zeros.

Remark 2.2. It follows from Definition 2.1 that X_t is ARTFIMA(p, λ, d, q) if and only if $Y_t = \Delta_1^{d,\lambda} X_t$ is an ARMA(p, q) time series

$$X_t + \sum_{j=1}^{p} \phi_j X_{t-j} = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i}$$

where $\{Z_t\}$ is a white noise sequence.

Theorem 2.3. *Suppose that X_t is an ARTFIMA(p, λ, d, q) time series that satisfies Definition 2.1. If*

$$\big|\Phi(z)\big| > 0 \quad \textit{and} \quad \big|\Theta(z)\big| > 0 \quad \textit{for } |z| \leq 1, \tag{2.4}$$

then:

(a) X_t is causal, i.e.,

$$X_t = \sum_{j=0}^{\infty} a_j^{-d,\lambda} Z_{t-j},$$

where $\sum_{j=0}^{\infty} \big|a_j^{-d,\lambda}\big| < \infty$; and

(b) X_t is invertible, i.e.,

$$Z_t = \sum_{j=0}^{\infty} c_j^{d,\lambda} X_{t-j},$$

where $\sum_{j=0}^{\infty} \left|c_j^{d,\lambda}\right| < \infty$.

Proof. By inverting the operator $\Delta^{d,\lambda}$, we get

$$W_t = \Delta^{-d,\lambda} Z_t = \sum_{j=0}^{\infty} (-1)^j e^{-\lambda j} \binom{-d}{j} Z_{t-j} = \sum_{j=0}^{\infty} \omega_j^{-d,\lambda} Z_{t-j}. \tag{2.5}$$

Since $\omega_j^{d,\lambda}$ has the same sign for all large j (e.g., see [30, Eq. (2.4)]) and

$$\sum_{j=0}^{\infty} \omega_j^{d,\lambda} = \sum_{j=0}^{\infty} (-1)^j e^{-\lambda j} \binom{d}{j} = (1-e^{-\lambda})^d < \infty,$$

by the fractional binomial formula (e.g., see Hille [19, p. 147]) it follows that

$$\sum_{j=0}^{\infty} \left|\omega_j^{d,\lambda}\right| < \infty \tag{2.6}$$

for all $\lambda > 0$ and all $d \notin \mathbb{Z}$. Define $A_\lambda B := \Delta^{-d,\lambda}\ \Theta(B)\Phi(B)^{-1}$ and write $\Theta(z)\Phi(z)^{-1} = \sum_{j=0}^{\infty} b_j z^j$ for $|z| \le 1$. Then

$$A_\lambda(z) = (1-e^{-\lambda}z)^{-d}\ \Theta(z)\Phi(z)^{-1} = \Big(\sum_{i=0}^{\infty} \omega_i^{-d,\lambda} z^i\Big)\Big(\sum_{s=0}^{\infty} b_s z^s\Big) = \sum_{j=0}^{\infty} a_j^{-d,\lambda} z^j,$$

where

$$a_j^{-d,\lambda} = \sum_{s=0}^{j} \omega_s^{-d,\lambda} b_{j-s} \tag{2.7}$$

for $j \ge 0$. Since X_t satisfies (2.3), we can write

$$X_t = \Delta^{-d,\lambda} \frac{\Theta(B)}{\Phi(B)} Z_t = \left(\sum_{j=0}^{\infty} a_j^{-d,\lambda} B^j\right) Z_t = \sum_{j=0}^{\infty} a_j^{-d,\lambda} Z_{t-j}. \tag{2.8}$$

where $a_j^{-d,\lambda}$ is given by (2.7). Under assumption (2.4), $\left|\Theta(z)/\Phi(z)\right| < \infty$, for $|z| \le 1+\varepsilon$, and the convergence of the series $\Theta(z)/\Phi(z)$ implies that $|b_j| \le C(1+\varepsilon)^{-j}$ for $j \ge 0$ (e.g., see [12, Theorem 7.2.3] or [6, Theorem 3.1.1]). By

applying (2.7), we have

$$\begin{aligned}\sum_{j=0}^{\infty}\left|a_j^{-d,\lambda}\right| &= \sum_{j=0}^{\infty}\Big|\sum_{s=0}^{j}\omega_s^{-d,\lambda}b_{j-s}\Big| \le \sum_{j=0}^{\infty}\sum_{s=0}^{j}\left|\omega_s^{-d,\lambda}\right|\,\left|b_{j-s}\right| \\ &= \sum_{s=0}^{\infty}\sum_{j=s}^{\infty}\left|\omega_s^{-d,\lambda}\right|\,\left|b_{j-s}\right| = \sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\left|\omega_s^{-d,\lambda}\right|\,\left|b_t\right| \\ &\le C\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\left|\omega_s^{-d,\lambda}\right|(1+\varepsilon)^{-t} = C\sum_{s=0}^{\infty}\left|\omega_s^{-d,\lambda}\right|\sum_{t=0}^{\infty}(1+\varepsilon)^{-t} < \infty\end{aligned} \tag{2.9}$$

since $\sum_{s=0}^{\infty}\left|\omega_s^{-d,\lambda}\right|$ is finite by (2.6) and $\sum_{t=0}^{\infty}(1+\varepsilon)^{-t} < \infty$. Now it follows from Brockwell and Davis [6, Proposition 3.1.2] that $\{X_t\}$, given by (2.8), is stationary and converges absolutely with probability one and this proves (a).

The white noise sequence $\{Z_t\}$ has spectral representation $Z_t = \int_{-\pi}^{\pi} e^{it\nu}\, dW(\nu)$ where $W(\nu)$ is an orthogonal increment process on $(-\pi,\pi)$ with $\mathbb{E}[dW(\nu)] = 0$, $\mathbb{E}[|dW(\nu)|^2] = d\nu/2\pi$ and $\mathbb{E}[dW(\nu)dW(\eta)] = 0$ for $\nu \neq \eta$ (e.g., see [12, Chapter 2]). Because $\sum_{j=0}^{\infty} a_j^{-d,\lambda}e^{-ij\nu} = (1-e^{-(\lambda+i\nu)})^{-d}\Theta(e^{-i\nu})/\Phi(e^{-i\nu})$, [12, Theorem 2.2.1] implies that

$$X_t = \int_{-\pi}^{\pi} e^{it\nu}(1-e^{-(\lambda+i\nu)})^{-d}\,\frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})}dW(\nu). \tag{2.10}$$

Define $B_\lambda(B) := \Delta_1^{d,\lambda}\ \Phi(B)/\Theta(B)$. Write $\Phi(z)/\Theta(z) = \sum_{j=0}^{\infty} c_j z^j$ for $|z| \le 1$ so that

$$B_\lambda(z) = (1-e^{-\lambda}z)^{-d}\,\frac{\Phi(z)}{\Theta(z)} = \Big(\sum_{i=0}^{\infty}\omega_i^{d,\lambda}z^i\Big)\Big(\sum_{s=0}^{\infty}c_s z^s\Big) = \sum_{j=0}^{\infty}c_j^{d,\lambda}z^j,$$

where

$$c_j^{d,\lambda} = \sum_{s=0}^{j}\omega_s^{d,\lambda}c_{j-s} \tag{2.11}$$

for $j \ge 0$. It is easy to check that $\sum_{j=0}^{\infty} c_j^{d,\lambda}e^{-ij\nu} = (1-e^{-(\lambda+i\nu)})^d\,\Phi(e^{-i\nu})/\Theta(e^{-i\nu})$.

Now, apply (2.10) to get:

$$\begin{aligned}\sum_{j=0}^{\infty} c_j^{d,\lambda} B^j X_t &= \sum_{j=0}^{\infty} c_j^{d,\lambda} X_{t-j} \\ &= \sum_{j=0}^{\infty} c_j^{d,\lambda} \int_{-\pi}^{\pi} e^{i(t-j)\nu}(1-e^{-(\lambda+i\nu)})^{-d} \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})} dW(\nu) \\ &= \int_{-\pi}^{\pi} \Big(\sum_{j=0}^{\infty} c_j^{d,\lambda} e^{-ij\nu}\Big) e^{it\nu}(1-e^{-(\lambda+i\nu)})^{-d} \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})}\ dW(\nu) \\ &= \int_{-\pi}^{\pi} (1-e^{-(\lambda+i\nu)})^{d}(1-e^{-(\lambda+i\nu)})^{-d} \frac{\Phi(e^{-i\nu})}{\Theta(e^{-i\nu})} \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})} e^{it\nu}\ dW(\nu) \\ &= Z_t.\end{aligned} \tag{2.12}$$

Under assumption (2.4), $\big|\Phi(z)/\Theta(z)\big| < \infty$, for $|z| \le 1+\varepsilon$, and the convergence of the series $\Phi(z)/\Theta(z)$ implies that

$$|c_j| \le K(1+\varepsilon)^{-j}, \tag{2.13}$$

for $j \ge 0$, see [12, Theorem 7.2.3]. Now apply (2.11) and (2.13) to write

$$\begin{aligned}\sum_{j=0}^{\infty} \big|c_j^{d,\lambda}\big| &= \sum_{j=0}^{\infty} \big|\sum_{s=0}^{j} \omega_s^{d,\lambda} c_{j-s}\big| \le \sum_{j=0}^{\infty}\sum_{s=0}^{j} \big|\omega_s^{d,\lambda}\big|\ \big|c_{j-s}\big| = \sum_{s=0}^{\infty}\sum_{j=s}^{\infty} \big|\omega_s^{d,\lambda}\big|\ \big|c_{j-s}\big| \\ &= \sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \big|\omega_s^{d,\lambda}\big|\ \big|c_t\big| \le K \sum_{s=0}^{\infty}\sum_{t=0}^{\infty} \big|\omega_s^{d,\lambda}\big|(1+\varepsilon)^{-t} \\ &= K \sum_{s=0}^{\infty} \big|\omega_s^{d,\lambda}\big| \sum_{t=0}^{\infty} (1+\varepsilon)^{-t} < \infty\end{aligned}$$

since $\sum_{s=0}^{\infty} \big|\omega_s^{d,\lambda}\big|$ is finite by (2.6) and $\sum_{t=0}^{\infty}(1+\varepsilon)^{-t} < \infty$. Now [6, Proposition 3.1.2] implies that $\{Z_t\}$ in (2.12) is stationary and converges absolutely with probability one, so X_t is invertible, which proves (b). □

In order to compute the covariance function of the ARTFIMA(p,λ,d,q) process X_t, we first recall a useful computational form of the spectral density for the ARMA(p,q) process $\Phi(B)U_t = \Theta(B)Z_t$, where Z_t is a white noise with $\mathbb{E}[Z_t^2] = \sigma^2$. Under assumption (2.4), $\Phi(x)$ can be written

$$\Phi(x) = \prod_{j=1}^{p}(1-\rho_j x),$$

where $\rho_1, \ldots, \rho_p$ are the complex numbers such that $|\rho_n| < 1$ for $n = 1, 2, \ldots, p$

(see Sowell [46, Section 4]). Then the spectral density of U_t is

$$f_U(\nu) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(\omega)\right|^2}{\left|\Phi(\omega)\right|^2}$$
$$= \frac{\sigma^2}{2\pi} \left|\Theta(\omega)\right|^2 \prod_{j=1}^{p} (1-\rho_j\omega)^{-1}(1-\rho_j\omega^{-1})^{-1},$$

where $\omega = e^{-i\nu}$. According to Sowell [46, Section 4], the spectral density of U_t can also be written as

$$f_U(\nu) = \frac{\sigma^2}{2\pi} \sum_{l=-q}^{q} \psi(l)\omega^l \sum_{j=1}^{p} \omega^p \zeta_j \Big[\frac{{\rho_j}^{2p}}{(1-\rho_j\omega)} - \frac{1}{(1-\rho_j^{-1}\omega)}\Big], \tag{2.14}$$

where

$$\psi(l) = \sum_{s=\min[0,l]}^{\max[q,q+l]} \theta_s\theta_{s-l},$$

and

$$\zeta_j = \frac{\sigma^2}{2\pi}\Big[\rho_j \prod_{i=1}^{p}(1-\rho_i\rho_j) \prod_{m\neq j}(\rho_j-\rho_m)\Big]^{-1}.$$

Recall that the Gaussian hypergeometric function

$${}_2F_1(a;b;c;z) = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(c+j)\Gamma(j+1)} z^j$$
$$= 1 + \frac{a\cdot b}{c\cdot 1} z + \frac{a(a+1)b(b+1)}{c(c+1)\cdot 1\cdot 2} z^2 + \dots$$

is defined for all complex numbers a and b, all complex $|z| < 1$ and real c not a negative integer. As $c \to -n$, a negative integer, we have

$$\lim_{c\to -n} \frac{{}_2F_1(a;b;c;z)}{\Gamma(c)} = a(a+1)\dots(a+n)b(b+1)\dots(b+n)$$
$$\cdot\, z^{n+1}\ {}_2F_1(a+n+1;b+n+1;n+2;z).$$

Hence when $c = -n$ we define ${}_2F_1(a;b;c;z)$ using this limit. We refer the reader to [15, Chapter 9] for more details about the Gaussian hypergeometric function.

Theorem 2.4. *Suppose that X_t is an ARTFIMA(p,λ,d,q) time series that satisfies Definition 2.1, and that (2.4) holds. Then:*

(a) $\{X_t\}$ has the spectral density

$$f_X(\nu) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(e^{-i\nu})\right|}{\left|\Psi(e^{-i\nu})\right|} \left|1 - e^{-(\lambda+i\nu)}\right|^{-2d} \tag{2.15}$$

for $-\pi \le \nu \le \pi$.

(b) The covariance function of $\{X_t\}$ is

$$\gamma_X(k) = \frac{\sigma^2}{2\pi} \sum_{l=-q}^{q} \sum_{j=1}^{p} \psi(l)\zeta_j C(d, \lambda, s-l-p, \rho_j),$$

using the notation in (2.14), where

$$\begin{aligned} &C(d, \lambda, h, \rho) \\ &= \rho^{2p} \sum_{m=0}^{\infty} \frac{\rho^m e^{-\lambda(h-m)} \Gamma(h-m+d)\ {}_2F_1(d; h-m+d; h-m+1; e^{-2\lambda})}{\Gamma(d)\Gamma(h-m+1)} \\ &\quad + \sum_{n=1}^{\infty} \frac{\rho^n e^{-\lambda(h+n)} \Gamma(h+n+d)\ {}_2F_1(d; h+n+d; h+n+1; e^{-2\lambda})}{\Gamma(d)\Gamma(h+n+1)}. \end{aligned} \tag{2.16}$$

Proof. (a) Use (2.3) to write $X_t = \psi_X(B)Z_t$ where $\psi_X(z) = (1-e^{-\lambda}z)^{-d}\Theta(z)/\Psi(z)$. Then the general theory of linear filters implies that X_t has spectral density $f_X(\nu) = |\psi_X(e^{-i\nu})|^2 f_Z(\nu)$ using the complex absolute value (e.g., see [6]). Since the white noise process $\{Z_t\}$ has spectral density $f_Z(\nu) = (2\pi)^{-1}\sigma^2$, it follows that (2.15) holds.

(b) Using (2.14), we can rewrite the spectral density of X_t in the form

$$\begin{aligned} f_X(\nu) &= \left(1 - e^{-(\lambda+i\nu)}\right)^{-d} \left(1 - e^{-(\lambda-i\nu)}\right)^{-d} h_U(\nu) \\ &= \frac{\sigma^2}{2\pi} \frac{\left|\Theta(\omega)\right|^2}{\left|\Phi(\omega)\right|^2} \left(1 - e^{-(\lambda+i\nu)}\right)^{-d} \left(1 - e^{-(\lambda-i\nu)}\right)^{-d} \\ &= \frac{\sigma^2}{2\pi} \sum_{l=-q}^{q} \sum_{j=1}^{p} \psi(l)\zeta_j \Big[\frac{{\rho_j}^{2p}}{(1-\rho_j\omega)} - \frac{1}{(1-\rho_j^{-1}\omega)}\Big] \\ &\qquad \left(1 - e^{-(\lambda+i\nu)}\right)^{-d} \left(1 - e^{-(\lambda-i\nu)}\right)^{-d} \omega^{p+l}. \end{aligned} \tag{2.17}$$

Next, we compute the covariance function of X_t. Recall that the covariance function $\gamma(k)$ and spectral density $f(\nu)$ are connected via $\gamma(k) = \mathbb{E}[X_t X_{t+k}] = \int_{-\pi}^{\pi} f(\nu) e^{i\nu k}\, d\nu$. Therefore we have

$$\begin{aligned} \gamma_X(k) &= \int_{-\pi}^{\pi} f_X(\nu) e^{i\nu k}\, d\nu \\ &= \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \Bigg[\sum_{l=-q}^{q} \sum_{j=1}^{p} \psi(l)\zeta_j \Big[\frac{{\rho_j}^{2p}}{(1-\rho_j\omega)} - \frac{1}{(1-\rho_j^{-1}\omega)}\Big] \\ &\qquad \left(1 - e^{-(\lambda+i\nu)}\right)^{-d} \left(1 - e^{-(\lambda-i\nu)}\right)^{-d} \omega^{p+l} \Bigg] e^{i\nu k}\, d\nu \\ &= \frac{\sigma^2}{2\pi} \sum_{l=-q}^{q} \sum_{j=1}^{p} \psi(l)\zeta_j C(d, \lambda, s-l-p, \rho_j), \end{aligned}$$

where

$$C(d,\lambda,h,\rho)=\int_{-\pi}^{\pi}\Big[\frac{\rho^{2p}}{(1-\rho\omega)}-\frac{1}{(1-\rho^{-1}\omega)}\Big]\Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d}e^{i\nu h}\ d\nu. \tag{2.18}$$

Next we write another form of $C(d,\lambda,h,\rho)$ by using the geometric series expansion:

$$\frac{\rho^{2p}}{(1-\rho\omega)}=\frac{\rho^{2p}}{(1-\rho e^{-i\nu})}=\rho^{2p}\sum_{m=0}^{\infty}(\rho e^{-i\nu})^m$$

and

$$\frac{-1}{(1-\rho^{-1}\omega)}=\frac{-1}{(1-\rho^{-1}e^{-i\nu})}=-1+\sum_{n=0}^{\infty}(\rho e^{i\nu})^n=\sum_{n=1}^{\infty}(\rho e^{i\nu})^n.$$

Using the spectral density of the ARTFIMA$(0,d,\lambda,0)$ process $W_t=\Delta^{-d,\lambda}Z_t$ given by

$$f_W(\nu)=\frac{\sigma^2}{2\pi}\Big|1-e^{-(\lambda+i\nu)}\Big|^{-2d}=\frac{\sigma^2}{2\pi}\left(1-2e^{-\lambda}\cos\nu+e^{-2\lambda}\right)^{-d}. \tag{2.19}$$

we can then write

$$\begin{aligned}
&C(d,\lambda,h,\rho)\\
&=\int_{-\pi}^{\pi}\rho^{2p}\sum_{m=0}^{\infty}(\rho e^{-i\nu})^m\Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d}e^{i\nu h}\ d\nu\\
&+\int_{-\pi}^{\pi}\sum_{n=1}^{\infty}(\rho e^{i\nu})^n\Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d}e^{i\nu h}\ d\nu\\
&=\rho^{2p}\sum_{m=0}^{\infty}\rho^m\int_{-\pi}^{\pi}\frac{2\pi}{\sigma^2}f_W(\nu)e^{i\nu(h-m)}\ d\nu+\sum_{n=1}^{\infty}\rho^n\int_{-\pi}^{\pi}\frac{2\pi}{\sigma^2}f_W(\nu)e^{i\nu(h+n)}\ d\nu\\
&=\rho^{2p}\sum_{m=0}^{\infty}\rho^m\frac{2\pi}{\sigma^2}\gamma_W(h-m)\ d\nu+\sum_{n=1}^{\infty}\rho^n\frac{2\pi}{\sigma^2}\gamma_W(h+n).
\end{aligned} \tag{2.20}$$

Next note that the covariance function of $W_t=\Delta^{-d,\lambda}Z_t$ is given by

$$\begin{aligned}
\gamma_W(k)&=\int_{-\pi}^{\pi}\cos(k\nu)f_W(\nu)\ d\nu\\
&=\frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}\frac{\cos(k\nu)}{\left(1-2e^{-\lambda}\cos\nu+e^{-2\lambda}\right)^d}\ d\nu\\
&=\frac{\sigma^2}{2\pi}\int_{0}^{2\pi}\frac{(-1)^k\cos(k\nu')}{\left(1+2e^{-\lambda}\cos\nu'+e^{-2\lambda}\right)^d}\ d\nu'\ [\nu':=\nu+\pi]\\
&=\sigma^2\frac{e^{-\lambda k}\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)}\ {}_2F_1(d;k+d;k+1;e^{-2\lambda}),
\end{aligned} \tag{2.21}$$

where we applied the integral formula (see [15], Eq. 9.112):

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{\cos k\omega}{(1-2z\cos\omega+z^2)^d}\,d\omega=\frac{z^k\Gamma(d+k)}{\Gamma(d)\Gamma(k+1)}\;{}_2F_1(d;k+d;k+1;z^2).$$

Substituting (2.21) into (2.20) it follows that (2.16) holds.

To complete the proof, we need to justify the interchange of the sum and integral in (2.20), observe that for any $d\notin\mathbb{Z}$ and any $|\rho|<1$ we have

$$\begin{aligned}
&\int_{-\pi}^{\pi}\sum_{s=0}^{\infty}\left|\rho^s\left(1-e^{-(\lambda+i\nu)}\right)^{-d}\left(1-e^{-(\lambda-i\nu)}\right)^{-d}e^{i\nu h}\right|d\nu\\
&\le\int_{-\pi}^{\pi}\sum_{s=0}^{\infty}\left|\rho\right|^s\left|\left(1-e^{-(\lambda+i\nu)}\right)^{-d}\left(1-e^{-(\lambda-i\nu)}\right)^{-d}e^{i\nu h}\right|d\nu\\
&=\frac{1}{1-|\rho|}\int_{-\pi}^{\pi}\left|\left(1-e^{-(\lambda+i\nu)}\right)^{-d}\left(1-e^{-(\lambda-i\nu)}\right)^{-d}e^{i\nu h}\right|d\nu\\
&=\frac{1}{1-|\rho|}\int_{-\pi}^{\pi}(1-2e^{-\lambda}\cos\nu+e^{-2\lambda})^{-d}\;d\nu\\
&=2\pi\gamma_W(0)<\infty,
\end{aligned}$$

where we applied (2.21) for $k=0$. □

2.1. Tempered fractional noise

An important special case of the ARTFIMA(p,λ,d,q) process is when $p=q=0$. This tempered fractional noise $W_t=\Delta^{-d,\lambda}Z_t$ has spectral density $f_W(\nu)$ given by (2.19) and covariance function $\gamma_W(k)$ given by (2.21).

Theorem 2.5. *Let W_t be a tempered fractional noise, i.e., an ARTFIMA(p,d,λ,q) time series as in Definition 2.1 with $p=q=0$, $d\notin\mathbb{Z}$, and $\lambda>0$. Then*

$$\sum_{k=0}^{+\infty}\left|\gamma_W(k)\right|<\infty\qquad\text{and}\qquad\sum_{k=-\infty}^{+\infty}\gamma_W(k)=\sigma^2(1-e^{-\lambda})^{-2d}.\tag{2.22}$$

Proof. Recall that the covariance function of the zero mean moving average process $W_t=\sum_{j=0}^{\infty}a_j\;Z_{t-j}$ can be written as

$$\gamma_W(k)=\mathbb{E}[W_0W_k]=\sigma^2\sum_{j=0}^{\infty}a_ja_{j+k}$$

for $k\ge0$ (see [12], Chapter 2). Therefore, using (2.5) we can write

$$\gamma_W(k)=\mathbb{E}[W_0W_k]=\sigma^2\sum_{j=0}^{\infty}\omega_j^{-d,\lambda}\omega_{j+k}^{-d,\lambda}$$

for $k \geq 0$. Since $\sum_{j=0}^{\infty} \left|\omega_j^{-d,\lambda}\right| < \infty$ and $\sum_{j=0}^{\infty} \omega_j^{-d,\lambda} = (1-e^{-\lambda})^{-d}$, it follows that

$$\sum_{k=0}^{+\infty} \left|\gamma_W(k)\right| \leq \sigma^2 \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \left|\omega_j^{-d,\lambda}\omega_{j+k}^{-d,\lambda}\right| \leq \sigma^2 \Big(\sum_{j=0}^{\infty} \left|\omega_j^{-d,\lambda}\right|\Big)^2 < \infty$$

and

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} \gamma_W(k) &= \gamma(0) + 2\sum_{k=1}^{\infty} \gamma_W(k) \\
&= \sigma^2 \Big(\sum_{j=0}^{\infty} (\omega_j^{-d,\lambda})^2 + 2\sum_{k=1}^{\infty}\sum_{j=0}^{\infty} \omega_j^{-d,\lambda}\omega_{j+k}^{-d,\lambda}\Big) \\
&= \sigma^2 \Big(\sum_{j=0}^{\infty} (\omega_j^{-d,\lambda})^2 + 2\sum_{j=0}^{\infty}\sum_{l=j+1}^{\infty} \omega_j^{-d,\lambda}\omega_l^{-d,\lambda}\Big) \\
&= \sigma^2 \Big(\sum_{j=0}^{\infty} (\omega_j^{-d,\lambda})^2 + \sum_{j=0}^{\infty}\sum_{l=0, l\neq j}^{\infty} \omega_j^{-d,\lambda}\omega_l^{-d,\lambda}\Big) \\
&= \sigma^2 \Big(\sum_{j=0}^{\infty} \omega_j^{-d,\lambda}\Big)^2 = \sigma^2 (1-e^{-\lambda})^{-2d}
\end{aligned}$$

and this completes the proof. □

Remark 2.6. If W_t is a fractional noise, ARTFIMA$(0, d, 0, 0)$, we have:

$$\sum_{k=0}^{+\infty} \left|\gamma(k)\right| < \infty \qquad \text{and} \qquad \sum_{k=-\infty}^{+\infty} \gamma(k) = 0 \qquad \text{for } -\frac{1}{2} < d < 0;$$

and

$$\sum_{k=0}^{+\infty} \left|\gamma(k)\right| = \infty \qquad \text{for } 0 < d < \frac{1}{2},$$

see Proposition 3.2.1 in [12]. This is consistent with Theorem 2.5, as the last term in (2.22) tends to zero or infinity, depending on the sign of d, as $\lambda \to 0$.

Remark 2.7. Here we compare the covariance function (2.21) of a tempered fractional noise to that of a fractional noise. The Gaussian hypergeometric function is related to the gamma function by ([1], Eq. 15.1.20):

$$_2F_1(a; b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \tag{2.23}$$

for $c - a - b > 0$. For the sake of completeness, we give the proof of (2.23). The Gaussian hypergeometric function $_2F_1(a; b; c; z)$ has the integral representation

$$_2F_1(a; b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\, dt \tag{2.24}$$

for $c > b > 0$ (see [1], Eq. 15.3.1). By taking the limit of the above integral representation as $z \to 1$ we get

$$\begin{aligned}
{}_2F_1(a;b;c;1) &= \lim_{z\to 1} {}_2F_1(a;b;c;z) \\
&= \lim_{z\to 1} \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\, dt \\
&= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-a-b-1}\, dt \\
&= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\, B(b, c-a-b) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \frac{\Gamma(b)\Gamma(c-a-b)}{\Gamma(c-a)} \\
&= \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}
\end{aligned}$$

which gives the desired result. Here $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$ is beta function. Let $\lambda = 0$ in (2.21) and apply (2.23) to get

$$\begin{aligned}
\gamma(k) = \mathbb{E}(X_t X_{t+k}) &= \frac{\sigma^2 \Gamma(d+k)}{\Gamma(d)\Gamma(k+1)}\, {}_2F_1(d; k+d; k+1; 1) \\
&= \sigma^2 \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)\Gamma(k+1-d)}
\end{aligned} \tag{2.25}$$

which is the autocovariance function of ARFIMA(0, d, 0), see [47], for $|d| < \frac{1}{2}$.

Peiris [38] has proposed a generalized autoregressive GAR(p) time series model $(1-\beta B)^d V_t = Z_t$ for applications in finance, where $|\beta| < 1$ and $BZ_t = Z_{t-1}$ is the backward shift operator. Taking $\beta = e^{-\lambda}$ we obtain a tempered fractional noise ARTFIMA$(0, d, \lambda, 0)$.

Lemma 2.8. *The covariance function of the tempered fractional noise $W_t = \Delta^{-d,\lambda} Z_t$ satisfies*

$$\gamma_W(k) \sim \frac{\sigma^2 (1-e^{-2\lambda})^{-d}}{\Gamma(d)}\, e^{-\lambda k} k^{d-1} \quad \text{as } k \to \infty. \tag{2.26}$$

Proof. Hosking [20] derived the asymptotic covariance function of the GAR(p) time series model $(1-\beta B)^d V_t = Z_t$,

$$\gamma_V(k) = \beta^k \frac{(k+d-1)!}{k!\,(d-1)!}\, {}_2F_1(d; k+d; k+1; \beta^2) \sim \frac{(1-\beta^2)^{-d}}{(d-1)!} \beta^k k^{d-1} \tag{2.27}$$

as $k \to \infty$. Substitute $e^{-\lambda} = \beta$ in (2.21) and apply (2.27) to get the desired result. □

Remark 2.9. We say that a stationary process $\{X_t\}$ exhibits *long range dependence* if

$$\sum_{k=0}^{\infty} \Big|\gamma(k)\Big| = \infty, \tag{2.28}$$

where $\gamma(k) = \mathbb{E}[X_t X_{t+k}]$ is the covariance function. Fractional noise is long range dependent when $0 < d < \frac{1}{2}$. Theorem 2.5 shows that tempered fractional noise is not long range dependent, but it does exhibit *semi-long range dependence* when $0 < d < \frac{1}{2}$. That is, for $\lambda > 0$ sufficiently small, the last term in (2.19) is large, since it tends to infinity as $\lambda \to 0$. Furthermore, Lemma 2.8 shows that, for small $\lambda > 0$, the covariance (2.26) falls off like k^{d-1} for moderately large values of k, but then falls off exponentially for very large lags $k \to \infty$. Finally, for small $\lambda > 0$, the spectral density grows like $|\nu|^{-2d}$ for moderately small ν, but then remains bounded for very small frequencies $\nu \to 0$,

3. Parameter estimation

In this section, we prove the consistency and asymptotic normality of the maximum likelihood and Whittle estimators for the parameters of an ARTFIMA(p, λ, d, q) time series, and we explicitly compute the asymptotic covariance matrix.

Recall from (2.17) that the ARTFIMA(p, λ, d, q) spectral density can be written in the form

$$\begin{aligned} f_X(\nu;\boldsymbol{\theta}) &= \frac{\sigma^2}{2\pi}\frac{\left|\Theta(e^{-i\nu})\right|^2}{\left|\Phi(e^{-i\nu})\right|^2}(1-2e^{-\lambda}\cos\nu+e^{-2\lambda})^{-d} \\ &= \frac{\sigma^2}{2\pi}\frac{\left|\Theta(e^{-i\nu})\right|^2}{\left|\Phi(e^{-i\nu})\right|^2}\left|1-e^{-(\lambda+i\nu)}\right|^{-d} := \frac{\sigma^2}{2\pi}K(\nu,\boldsymbol{\theta}), \end{aligned} \tag{3.1}$$

for $\nu \in (-\pi, \pi)$, where $\sigma > 0$, and $\boldsymbol{\theta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d, \lambda)$.

Let $\mathbf{X} = (X_1, \ldots, X_N)$ be a realization of the ARTFIMA(p, λ, d, q) time series with sample size N and consider the periodogram

$$I_{\mathbf{X}}(\nu) := \frac{1}{2\pi N}\Big|\sum_{t=1}^{N} X_t e^{it\nu}\Big|. \tag{3.2}$$

Define

$$Q_{\mathbf{X}}(\boldsymbol{\theta}) := \int_{-\pi}^{\pi}\frac{I_{\mathbf{X}}(\nu)}{K(\nu,\boldsymbol{\theta})}d\nu \tag{3.3}$$

and

$$D_N(\mathbf{X},\sigma,\boldsymbol{\theta}) := \frac{1}{2\sigma^2}Q_{\mathbf{X}}(\boldsymbol{\theta}) + \log\sigma. \tag{3.4}$$

Let σ_0 and $\boldsymbol{\theta}_0$ denote the true parameter values of $\sigma \in (0, \infty)$ and $\boldsymbol{\theta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d, \lambda) \in \Xi$, respectively, where $\Xi = \mathbb{R}^{p+q+1} \times (0, \infty)$. Define $\Omega = (0, \infty) \times \Xi$.

Definition 3.1. The *Whittle estimators* of σ_0 and $\boldsymbol{\theta}_0$ based on $\mathbf{X} = (X_1, \ldots, X_N)$ are defined by

$$(\bar{\sigma}_N, \bar{\boldsymbol{\theta}}_N) := \arg\min\{D_N(\mathbf{X},\sigma,\boldsymbol{\theta}) : (\sigma,\boldsymbol{\theta}) \in \Omega\}, \tag{3.5}$$

so that $\bar{\boldsymbol{\theta}}_N = \arg\min\{Q_{\mathbf{X}}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Xi\}$ and $\bar{\sigma}_N^2 = Q_{\mathbf{X}}(\bar{\boldsymbol{\theta}}_N)$.

In order to obtain consistency, we make the standard assumption that the parameter vector $(\sigma, \boldsymbol{\theta})$ is restricted to a compact set Ω_0. Choose this set so that $d \in [-1/2, 1/2]$, and so that $(\sigma_0, \boldsymbol{\theta}_0)$ is an interior point. Other values of d can be obtained by differencing.

Theorem 3.2. *The Whittle estimators* (3.5) *over the compact parameter space* Ω_0 *are strongly consistent. That is,*

$$\lim_{N\to\infty} \bar{\boldsymbol{\theta}}_N = \boldsymbol{\theta}_0 \qquad a.s.$$

and

$$\lim_{N\to\infty} \bar{\sigma}_N^2(\bar{\theta}_N) = \sigma_0^2 \qquad a.s.$$

Proof. First we check that the ARTFIMA(p, λ, d, q) time series is ergodic. For this, it is sufficient (e.g., see Hamilton [16, p. 504]) that

$$\sum_{j=0}^{\infty} j|a_j^{-d,\lambda}| < \infty \tag{3.6}$$

where $a_j^{-d,\lambda}$ are the moving average weights defined by (2.7). As in the proof of Theorem 2.3 we have $|b_j| \le C_1(1+\varepsilon)^{-j}$ for $j \ge 0$. An application of Stirling's approximation (e.g., see [30, Eq. (2.5)]) yields that $|\omega_j^{-d,\lambda}| \le C_2 j^{\alpha-1} e^{-\lambda j}$ for all $j > 0$, and note that $\omega_0^{-d,\lambda} = 1$. Now write

$$\begin{aligned}\sum_{j=0}^{\infty} j|a_j^{-d_0,\lambda_0}| &\le \sum_{j=0}^{\infty}\sum_{s=0}^{j} j\left|\omega_s^{-d,\lambda} b_{j-s}\right| = \sum_{s=0}^{\infty}\sum_{j=s}^{\infty} j\left|\omega_s^{-d,\lambda} b_{j-s}\right| \\ &\le C_1 \sum_{s=0}^{\infty}\sum_{t=0}^{\infty}(s+t)\left|\omega_s^{-d,\lambda}\right|(1+\varepsilon)^{-t} := C_1 \sum_{s=0}^{\infty}\left|\omega_s^{-d,\lambda}\right|(A+Bs)\end{aligned}$$

where

$$\begin{aligned}A &= \sum_{t=1}^{\infty} t(1+\varepsilon)^{-t} = (1+\varepsilon)^{-1}(1-(1+\varepsilon)^{-1})^{-2} < \infty, \\ B &= \sum_{t=0}^{\infty}(1+\varepsilon)^{-t} = (1-(1+\varepsilon)^{-1})^{-1} < \infty.\end{aligned}$$

Since

$$\sum_{s=0}^{\infty} s\left|\omega_s^{-d,\lambda}\right| \le \sum_{s=1}^{\infty} C_2 j^{\alpha} e^{-\lambda j} < \infty$$

it follows using (2.6) that (3.6) holds, and hence the time series is ergodic. Then we can apply Theorem 8.2.1 in [12] (see also Theorem 1 in Hannan [17]). This requires us to verify the conditions:

(1) The parameters $(\sigma, \boldsymbol{\theta}) \in \Omega$ determine the spectral density function (3.1) uniquely.
(2) $1/(K(\nu, \boldsymbol{\theta}) + a)$ is continuous in $(\nu, \boldsymbol{\theta}) \in (-\pi, \pi) \times \Xi$, for all $a > 0$.
(3) $\sum_{j=0}^{\infty} (a_j^{-d,\lambda})^2 < \infty$ and $a_0^{-d,\lambda} = 1$ where $a_j^{-d,\lambda}$ is given by (2.7).

Since d is not an integer, it is easy to see that condition (1) holds. Since $K(\nu, \boldsymbol{\theta})$ is continuous and strictly positive for $\lambda > 0$, it is apparent that condition (2) holds. From (2.9) it follows that $|a_j^{-d,\lambda}| \leq C_3$ for all $j \geq 0$. Then it follows easily from (3.6) that

$$\sum_{j=0}^{\infty} j(a_j^{-d,\lambda})^2 \leq C_3 \sum_{j=0}^{\infty} j|a_j^{-d,\lambda}| < \infty. \tag{3.7}$$

It is also easy to see that $a_0^{-d,\lambda} = \omega_0^{-d,\lambda} b_0 = 1$. Then Condition (3) follows, and now the proof is complete. □

Given a function $f(\boldsymbol{\theta})$, we write the gradient vector $\partial f(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$, a column vector with components $\partial f(\boldsymbol{\theta}_j)/\partial \boldsymbol{\theta}_j$ for $j = 1, 2, \ldots, p + q + 2$, where $\boldsymbol{\theta}_j$ is the jth component of the vector $\boldsymbol{\theta}$. Then the outer product $\{\partial f(\boldsymbol{\theta})/\partial \boldsymbol{\theta}\}\{\partial f(\boldsymbol{\theta})/\partial \boldsymbol{\theta}\}'$ is a matrix. Here the prime denotes the transpose, a row vector.

Theorem 3.3. *The Whittle estimators (3.5) over the compact parameter space Ω_0 are asymptotically normal. That is, $N^{1/2}(\bar{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)$ converges in distribution to a Gaussian random vector with zero mean vector and covariance matrix $\mathbf{W}^{-1}$ where*

$$\mathbf{W} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \frac{\partial \log K(\nu, \boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \right\} \left\{ \frac{\partial \log K(\nu, \boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \right\}' d\nu. \tag{3.8}$$

Proof. The result follows from Theorem 2 in Hannan [17] or Theorem 8.3.1 in Giraitis et al. [12], once we verify the following conditions:

(1) $K(\nu, \boldsymbol{\theta}) \geq a > 0$ for $\nu \in (-\pi, \pi)$.
(2) $K(\nu, \boldsymbol{\theta})$ is a twice differentiable function of parameters $\lambda, d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$.
(3) Equation (3.7) holds.

Property (1) follows from the fact that $|1 - e^{-(\lambda + i\nu)}| > 1 - e^{-\lambda}$ for all ν. Property (2) is apparent from the first line of (3.1), since $\lambda > 0$. Condition (3) was verified in the proof of Theorem 3.2. □

Theorem 3.4. *The covariance matrix $\mathbf{W}$ in (3.8) has the form*

$$\mathbf{W} = \begin{pmatrix} \mathbf{I}_{(p+q)\times(p+q)} & \mathbf{J}_{(p+q)\times 2} \\ \mathbf{J}'_{2\times(p+q)} & \mathbf{V}_{2\times 2} \end{pmatrix}_{(p+q+2)\times(p+q+2)}$$

where:

(1) the upper left block can be written in the form

$$\mathbf{I} = \begin{pmatrix} w_{1,1} & \cdots & w_{1,p} & w_{1,p+1} & \cdots & w_{1,p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{p,1} & \cdots & w_{p,p} & w_{p,p+1} & \cdots & w_{p,p+q} \\ w_{p+1,1} & \cdots & w_{p+1,p} & w_{p+1,p+1} & \cdots & w_{p+1,p+q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ w_{p+q,1} & \cdots & w_{p+q,p} & w_{p+q,p+1} & \cdots & w_{p+q,p+q} \end{pmatrix}$$

where:

(1a) Taking $\Phi(B)M_t = Z_t$ where $\{Z_t\} \sim WN(0,\sigma^2)$, we can write $w_{j,k} = \mathbb{E}[M_{t-j+1}M_{t-k+1}] = \gamma_M(j-k)$ for $1 \le j \le p$ and $1 \le k \le p$;

(1b) Taking $\Theta(B)N_t = Z_t$ where $\{Z_t\} \sim WN(0,\sigma^2)$, we can write $w_{j,k} = \mathbb{E}[N_{t-j+1}N_{t-k+1}] = \gamma_N(j-k)$ for $p+1 \le j \le p+q$ and $p+1 \le k \le p+q$;

(1c) We can write $w_{j,p+m} = \mathbb{E}[M_{t-j+1}N_{t-m+1}]$ for $1 \le j \le p$ and $1 \le m \le q$;

(1d) By symmetry we have $w_{p+m,j} = w_{j,p+m}$ for $1 \le j \le p$ and $1 \le m \le q$;

(2) the upper right block can be written in the form

$$\mathbf{J} = \begin{pmatrix} J_{1,1} & J_{1,2} \\ \vdots & \vdots \\ J_{p,1} & J_{p,2} \\ J_{p+1,1} & J_{p+1,2} \\ \vdots & \vdots \\ J_{p+q,1} & J_{p+q,2} \end{pmatrix}$$

where:

(2a) For $1 \le j \le p$,

$$J_{j,1} = \frac{-1}{4\pi}\int_{-\pi}^{\pi}[e^{-i\nu j}\Phi^{-1}(e^{-i\nu}) + e^{i\nu j}\Phi^{-1}(e^{i\nu})]\log(1-2e^{-\lambda}\cos\nu + e^{-2\lambda})d\nu;$$

(2b) If $p+1 \le j \le p+q$,

$$J_{j,1} = \frac{-1}{4\pi}\int_{-\pi}^{\pi}[e^{-i\nu j}\Theta^{-1}(e^{-i\nu}) + e^{i\nu j}\Theta^{-1}(e^{i\nu})]\log(1-2e^{-\lambda}\cos\nu + e^{-2\lambda})d\nu;$$

(2c) For $1 \le j \le p$,

$$J_{j,2} = \frac{-1}{2\pi}\int_{-\pi}^{\Pi}[e^{-i\nu j}\Phi^{-1}(e^{-i\nu}) + e^{i\nu j}\Phi^{-1}(e^{i\nu})]de^{-\lambda}\Big(\frac{\cos\nu - e^{-\lambda}}{1-2e^{-\lambda}\cos\nu + e^{-2\lambda}}\Big)d\nu;$$

(2d) If $p+1 \le j \le p+q$,

$$J_{j,2} = \frac{-1}{4\pi}\int_{-\pi}^{\pi}[e^{-i\nu j}\Theta^{-1}(e^{-i\nu}) + e^{i\nu j}\Theta^{-1}(e^{i\nu})]de^{-\lambda}\Big(\frac{\cos\nu - e^{-\lambda}}{1-2e^{-\lambda}\cos\nu + e^{-2\lambda}}\Big)d\nu;$$

(3) the lower right block can be written in the form

$$\mathbf{V} = \begin{pmatrix} v_{1,1} & v_{1,2} \\ v_{2,1} & v_{2,2} \end{pmatrix}$$

where:

$$\begin{aligned} v_{1,1} &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\big(\log(1-2e^{-\lambda}\cos\nu + e^{-2\lambda})\big)^2\, d\nu \\ v_{2,2} &= \frac{d^2 e^{-2\lambda}}{1-e^{-2\lambda}} \\ v_{1,2} &= v_{2,1} = d\ln(1-e^{-2\lambda}) \end{aligned}$$

xxx

Proof. The matrix $\mathbf{I}$ is the same as the matrix $\mathbf{W}$ in Brockwell and Davis [6, Theorem 10.8.2, p. 386] for the asymptotic covariance of the ARMA(p,q) parameter estimates, see also Box and Jenkins [4, p.240] or Li and McLeod [26, Section 3]. This proves part (1).

To prove the remaining parts, write

$$\log K(\nu,\boldsymbol{\theta}) = \log(\big|\Theta(e^{-i\nu})\big|^2) - \log(\big|\Phi(e^{-i\nu})\big|^2) - d\log(1-2e^{-\lambda}\cos\nu + e^{-2\lambda}).$$

Observe that

$$\begin{aligned} \log(\big|\Phi(e^{-i\nu})\big|^2) &= \log(\Phi(e^{-i\nu})) + \log(\Phi(e^{i\nu})) \\ &= \log(1+\phi_1 e^{-i\nu} + \ldots + \phi_j e^{-i\nu j} + \ldots + \phi_p e^{-i\nu p}) \\ &+ \log(1+\phi_1 e^{i\nu} + \ldots + \phi_j e^{i\nu j} + \ldots + \phi_p e^{i\nu p}) \end{aligned}$$

and then

$$\begin{aligned} \frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial\phi_j} &= \frac{\partial\log(\big|\Phi(e^{-i\nu})\big|^2)}{\partial\phi_j} = \frac{\partial}{\partial\phi_j}\big\{\log(\Phi(e^{-i\nu})) + \log(\Phi(e^{i\nu}))\big\} \\ &= \frac{\partial}{\partial\phi_j}\log(1+\phi_1 e^{-i\nu} + \ldots + \phi_j e^{-i\nu j} + \ldots + \phi_p e^{-i\nu p}) \\ &+ \frac{\partial}{\partial\phi_j}\log(1+\phi_1 e^{i\nu} + \ldots + \phi_j e^{i\nu j} + \ldots + \phi_p e^{i\nu p}) \\ &= \frac{e^{-i\nu j}}{1+\phi_1 e^{-i\nu} + \ldots + \phi_j e^{-i\nu j} + \ldots + \phi_p e^{-i\nu p}} \\ &+ \frac{e^{i\nu j}}{1+\phi_1 e^{i\nu} + \ldots + \phi_j e^{i\nu j} + \ldots + \phi_p e^{i\nu p}} \\ &= e^{-i\nu j}\Phi^{-1}(e^{-i\nu}) + e^{i\nu j}\Phi^{-1}(e^{i\nu}). \end{aligned} \tag{3.9}$$

By a similar argument,

$$\begin{aligned}\log(\left|\Theta(e^{-i\nu})\right|^2) &= \log(\Theta(e^{-i\nu})) + \log(\Theta(e^{i\nu}))\\ &= \log(1+\theta_1 e^{-i\nu}+\ldots+\theta_j e^{-i\nu j}+\ldots+\theta_q e^{-i\nu q})\\ &+ \log(1+\theta_1 e^{i\nu}+\ldots+\theta_j e^{i\nu j}+\ldots+\theta_q e^{i\nu q})\end{aligned}$$

and consequently

$$\begin{aligned}\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial\theta_j} &= \frac{\partial \log(\left|\Theta(e^{-i\nu})\right|^2)}{\partial\theta_j} = \frac{\partial}{\partial\theta_j}\big\{\log(\Theta(e^{-i\nu}))+\log(\Theta(e^{i\nu}))\big\}\\ &= \frac{\partial}{\partial\theta_j}\log(1+\theta_1 e^{-i\nu}+\ldots+\theta_j e^{-i\nu j}+\ldots+\theta_q e^{-i\nu q})\\ &+ \frac{\partial}{\partial\theta_j}\log(1+\theta_1 e^{i\nu}+\ldots+\theta_j e^{i\nu j}+\ldots+\theta_q e^{i\nu q})\\ &= \frac{e^{-i\nu j}}{1+\theta_1 e^{-i\nu}+\ldots+\theta_j e^{-i\nu j}+\ldots+\theta_q e^{-i\nu q}}\\ &+ \frac{e^{i\nu j}}{1+\theta_1 e^{i\nu}+\ldots+\theta_j e^{i\nu j}+\ldots+\theta_q e^{i\nu q}}\\ &= e^{-i\nu j}\Theta^{-1}(e^{-i\nu}) + e^{i\nu j}\Theta^{-1}(e^{i\nu}).\end{aligned} \tag{3.10}$$

It is also easy to see that

$$\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial d} = -\log(1-2e^{-\lambda}\cos\nu+e^{-2\lambda}) \tag{3.11}$$

and

$$\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial \lambda} = -2de^{-\lambda}\Big(\frac{\cos\nu-e^{-\lambda}}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big) \tag{3.12}$$

Then part (2) follows easily.

As for (3), the formula for $v_{1,1}$ follows immediately from (3.11). A useful approximation will be provided in a remark following the proof. Next write

$$\begin{aligned}v_{2,2} &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\Big(-2de^{-\lambda}\Big(\frac{\cos\nu-e^{-\lambda}}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big)\Big)^2\,d\nu\\ &= \frac{d^2}{4\pi}\int_{-\pi}^{\pi}\Big(1+\frac{e^{-2\lambda}-1}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big)^2 d\nu\\ &= \frac{d^2}{4\pi}\int_{-\pi}^{\pi}\Big[1+\frac{(e^{-2\lambda}-1)^2}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)^2}+\frac{2(e^{-2\lambda}-1)}{1+e^{-2\lambda}-2e^{-\lambda}\cos\nu}\Big]d\nu.\end{aligned} \tag{3.13}$$

Observe that

$$\begin{aligned}\int_{-\pi}^{\pi}\frac{d\nu}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)^2} &= 2\int_0^{\pi}\frac{d\nu}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)^2}\\ &= \frac{2\pi(1+e^{-2\lambda})}{(1-e^{-2\lambda})^3}\end{aligned} \tag{3.14}$$

where we used the standard formula

$$\int_0^\pi \frac{d\nu}{(1+a^2-2a\cos\nu)^n} = \frac{\pi}{(1-a^2)^n}\sum_{k=0}^{n-1}\frac{(n+k-1)!}{(k!)^2(n-k-1)!}\Big(\frac{a^2}{1-a^2}\Big)^k \tag{3.15}$$

for $a^2<1$ (see [15, p. 608]). By applying (3.15) for $n=1$, we also have

$$\int_{-\pi}^\pi \frac{d\nu}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)} = \frac{2\pi}{1-e^{-2\lambda}}. \tag{3.16}$$

Therefore, from (3.13), (3.14) and (3.16):

$$v_{2,2} = \frac{d^2}{4\pi}\Big[2\pi + \frac{2\pi(1+e^{-2\lambda})}{(1-e^{-2\lambda})} - 4\pi\Big] = \frac{d^2e^{-2\lambda}}{1-e^{-2\lambda}} \tag{3.17}$$

as claimed. Finally, write

$$\begin{aligned} v_{1,2} = v_{2,1} &= \frac{-d}{4\pi}\int_{-\pi}^\pi \big(\log(1-2e^{-\lambda}\cos\nu+e^{-2\lambda})\big)\Big(1+\frac{e^{-2\lambda}-1}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big)\,d\nu \\ &= \frac{-d}{4\pi}(I_1+I_2) \end{aligned} \tag{3.18}$$

where

$$\begin{aligned} I_1 &:= \int_{-\pi}^\pi \ln(1-2e^{-\lambda}\cos\nu+e^{-2\lambda})\,d\nu \\ I_2 &:= \int_{-\pi}^\pi \frac{(e^{-2\lambda}-1)\ln(1-2e^{-\lambda}\cos\nu+e^{-2\lambda})}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\,d\nu. \end{aligned}$$

To calculate I_1, use integration by parts to write

$$\begin{aligned} I_1 &= \nu\ln(1-2e^{-\lambda}\cos\nu+e^{-2\lambda})\big]_{-\pi}^{\pi} - \int_{-\pi}^\pi \frac{2e^{-\lambda}\nu\sin\nu}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\,d\nu \\ &= 2\pi\ln(1+e^{-\lambda})^2 - 4e^{-\lambda}\int_0^\pi \frac{\nu\sin\nu}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\,d\nu \\ &= 4\pi\ln(1+e^{-\lambda}) - 4\pi\ln(1+e^{-\lambda}) = 0, \end{aligned} \tag{3.19}$$

since

$$\int_0^\pi \frac{\nu\sin\nu}{1-2a\cos\nu+a^2}\,d\nu = \frac{\pi}{a}\ln(1+a)$$

for $a^2<1$ and $a\neq 0$ (see [15, No 4, p. 696]).

In order to calculate I_2, we use the standard formula

$$\int_0^\pi \frac{\ln(1-2a\cos\nu+a^2)}{1-2b\cos\nu+b^2}\,d\nu = \frac{2\pi\ln(1-ab)}{1-b^2}$$

for $a^2 \leq 1$ and $b^2 < 1$ (see [15, No. 16, p. 925]). In our case $a = b = e^{-\lambda}$. Therefore

$$\begin{aligned} I_2 &= (e^{-2\lambda} - 1) \int_{-\pi}^{\pi} \frac{\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda})}{1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}}\, d\nu \\ &= 2(e^{-2\lambda} - 1) \int_{0}^{\pi} \frac{\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda})}{1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}}\, d\nu \\ &= \frac{4\pi \ln(1 - e^{-2\lambda})}{1 - e^{-2\lambda}}(e^{-2\lambda} - 1) = -4\pi \ln(1 - e^{-2\lambda}). \end{aligned} \tag{3.20}$$

Now, from (3.18), (3.19) and (3.20), we have

$$v_{1,2} = v_{2,1} = \frac{-d}{4\pi}\big[-4\pi \ln(1 - e^{-2\lambda})\big] = d \ln(1 - e^{-2\lambda}), \tag{3.21}$$

and this completes the proof of (3). □

Remark 3.5. A useful approximation of $v_{1,1}$ can be obtained by using the fact that

$$\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}) \sim -2e^{-\lambda}\cos\nu + e^{-2\lambda} \tag{3.22}$$

as $(\lambda, \nu) \to 0$. Then

$$v_{1,1} \sim \int_{-\pi}^{\pi} (-2e^{-\lambda}\cos\nu + e^{-2\lambda})^2\, d\nu = 4\pi(e^{-2\lambda} + \frac{e^{-4\lambda}}{2}) \tag{3.23}$$

at low frequencies ν, when the tempering parameter $\lambda > 0$ is small.

Given $\mathbf{X} = (X_1, \ldots, X_N)$, the maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}_N$ for parameters $\boldsymbol{\theta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d, \lambda)$ can be computed using the logarithm of the likelihood function

$$l(X_1, \ldots, X_N) = (2\pi\sigma^2)^{-\frac{N}{2}} |G_N|^{-\frac{1}{2}} \exp\left[\frac{-\mathbf{X}_N' \, G_N^{-1} \mathbf{X}_N}{2\sigma^2}\right]$$

where $G_N = \sigma^{-2}\mathbb{E}[\mathbf{X}_N \mathbf{X}_N']$ and $|G_N|$ is the determinant of G_N. The next results shows that the MLE is also asymptotically normal.

Theorem 3.6. *Under the same assumptions at Theorem 3.2, the maximum likelihood estimator is strongly consistent, i.e.,*

$$\lim_{N\to\infty} \hat{\boldsymbol{\theta}}_N = \boldsymbol{\theta}_0 \qquad a.s.$$

and

$$\lim_{N\to\infty} \hat{\sigma}_N^2(\hat{\boldsymbol{\theta}}_N) = \sigma_0^2 \qquad a.s.$$

Proof. The proof follows from Theorem 1 in [17]. □

Theorem 3.7. *Under the same assumptions at Theorem 3.3, $N^{1/2}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)$ converges in distribution to a Gaussian random vector with zero mean and covariance matrix $\mathbf{W}^{-1}$ where $\mathbf{W}$ is given by (3.8).*

Proof. This follows immediately from Theorem 3.3, Theorem 3.4, and Theorem 10.8.2 in Brockwell and Davis [6]. □

Remark 3.8. The following argument shows that, for large N, we can approximate the logarithm of the likelihood function using (3.4). That is, the MLE is approximately the same as the Whittle estimator. Since the Whittle estimator is much easier to compute, this is a useful approximation in practice. Write

$$\log l(X_1,\dots,X_N) = \frac{-N}{2}\log(2\pi) - N\log\sigma - \frac{1}{2}\log|G_N| - \Big(\frac{\mathbf{X}_N' \; G_N^{-1}\mathbf{X}_N}{2\sigma^2}\Big)$$

and hence

$$\begin{aligned}\frac{-\log l(X_1,\dots,X_N)}{N} &= \frac{1}{2}\log(2\pi) + \log\sigma + \frac{1}{2N}\log|G_N| + \frac{N^{-1}}{2\sigma^2}\Big(\mathbf{X}_N' \; G_N^{-1}\mathbf{X}_N\Big)\\ &= \frac{1}{2}\log(2\pi) + \log\sigma + \frac{1}{2N}\log|G_N|\\ &\quad + \Big[\frac{N^{-1}}{2\sigma^2}\Big(\mathbf{X}_N' \; G_N^{-1}\mathbf{X}_N\Big) - \frac{1}{2\sigma^2}\frac{2\pi}{N}\sum_{t=-N/2}^{[N/2]}\frac{I_{\mathbf{X}_N}(\nu_t)}{K(\nu_t,\boldsymbol{\theta})}\Big]\\ &\quad + \frac{1}{2\sigma^2}\frac{2\pi}{N}\sum_{t=-N/2}^{[N/2]}\frac{I_{\mathbf{X}_N}(\nu_t)}{K(\nu_t,\boldsymbol{\theta})},\end{aligned}$$

where $\nu_t = 2\pi t/N$. According to Lemma 3 in Hannan [17], $\log|G_N|/N \to 0$ as $N\to\infty$. Moreover, Lemma 4 in Hanna [17], implies that

$$\lim_{N\to\infty}\Big[N^{-1}\Big(\mathbf{X}_N' \; G_N^{-1}\mathbf{X}_N\Big) - \frac{2\pi}{N}\sum_{t=-N/2}^{[N/2]}\frac{I_{\mathbf{X}_N}(\nu_t)}{K(\nu_t,\boldsymbol{\theta})}\Big] = 0.$$

Therefore

$$\begin{aligned}\frac{-\log l(X_1,\dots,X_N)}{N} &\approx \frac{1}{2}\log(2\pi) + \log\sigma + \frac{1}{2\sigma^2}\frac{2\pi}{N}\sum_{t=-N/2}^{[N/2]}\frac{I_{\mathbf{X}_N}(\nu_t)}{K(\nu_t,\boldsymbol{\theta})}\\ &\approx \frac{1}{2}\log(2\pi) + \log\sigma + \frac{1}{2\sigma^2}\int_{-\pi}^{\pi}\frac{I_{\mathbf{X}_N}(\nu)}{K(\nu,\boldsymbol{\theta})}\,d\nu\\ &= \frac{1}{2}\log(2\pi) + \log\sigma + \frac{1}{2\sigma^2}Q_{X_{\mathbf{N}}} = \frac{1}{2}\log(2\pi) + D_N(\mathbf{X}_N,\sigma,\boldsymbol{\theta})\end{aligned}$$

for large N.

4. Applications

This section provides some data applications of the ARTFIMA(p,d,λ,q) time series model, using an R package called `artfima` that has been written to accompany this paper. The package is freely available on CRAN.

Geophysical turbulence in water velocity data (cm/s) was measured in Lake Michigan, Lake Huron, and the Red Cedar River in Michigan, see [33] for further details. Figure 1 shows the periodogram and fitted ARTFIMA(p, d, λ, q) spectral density function for a data set from Saginaw Bay. Using the `artfima` package in R and setting $p = q = 0$ for a tempered fractional noise, we also set $d = 5/6$ (from theory, Kolmogorov scaling, see [33] for further details), and this resulted in the parameter fit $\lambda = 0.045$ (0.00248) using the Whittle estimator, where the second number in parentheses is the standard error. The plot uses a log-log scale to highlight the power law relation between frequency ν and spectral density $f_X(\nu)$ for $\log(\nu) > -4$. The tempering causes a deviation from that line at the lowest frequencies, a feature also seen in the data. Without fixing d, the Whittle estimates are $\lambda = 0.027$ (.00229) and $d = 0.752$ (.00582).

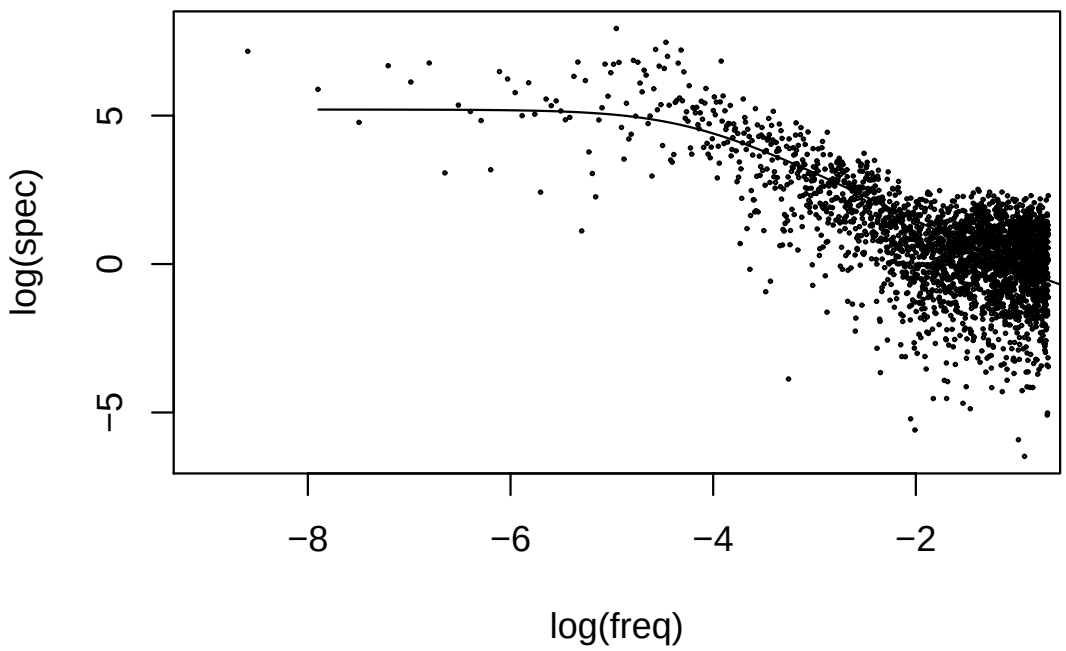

Fig 1. *Spectral density of water velocity data from Saginaw Bay (circles), along with fitted ARTFIMA spectrum (line).*

5. Acknowledgments

We would like to thank Mantha Phanikumar, Michigan State University, for access to the geophysical turbulence data sets.

References

[1] Abramowitz, M., Stegun, I. (1965). Handbook of mathematical functions, ninth edition, Dover, New York.
[2] Bai, S., Taqqu, M.,S. (2014). Generalized Hermite processes, discrete chaos and limit theorems. Stochastic processes and their applications **124**, 1710–1739.
[3] Billingsley, P. (1968): Convergence of Probability Measures. New York: Wiley.
[4] Box and Jenkins, 1976 Box, G., E., P., and Jenkins, G., M., Time series analysis forecasting and control, Holden-Day, San Francisco, CA (1976)
[5] Banerjee, S., Gelfand, A.E. (2003). On Smoothness Properties of Spatial Processes. Journal of Multivariate Analysis, **84**, 85–100.
[6] Brockwell P J and Davis R A, 1991 Time Series: Theory and Methods, 2nd Ed.(Springer-Verlag, New York)
[7] Cartea, Á., del-Castillo-Negrete, D. (2007). Fluid limit of the continuous-time random walk with general Lévy jump distribution functions. Physical Review E **76**, 041105.
[8] Cheridito, P. (2004). Gaussian moving averages, semimartingales and option pricing. Stochastic Processes and their Applications, **109(1)**, 47–68 .
[9] Dobrushin, R.L., Major, P. (1979). Non-central limit theorems for non-linear functionals of Gaussian fields. Z. Wahrscheinlichkeitstheorie verw. Gebiete 50, 27-52.
[10] Embrechts, P., Maejima, M. (2002). Selfsimilar Processes, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ.
[11] Goncalves, E,. Processus Fractionnaires (1988)
[12] L. Giraitis, H.L. Koul, and D. Surgailis. Large Sample Inference for Long Memory Processes. World Scientific Publishing Company Incorporated, (2012)
[13] Gneiting, T., Kleiber, W., Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. Journal of the American Statistical Association. **105**, 1167–1177.
[14] Friedlander. S., K. and Topper. L. (1962). Turbulence: Classical papers on statistical theory, (Interscience Publishers, New York)
[15] Gradshteyn, I.S., Ryzhik, I.M. (1980). Table of Integrals, seris and Products. Academic Press, New York.
[16] Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press, New Jersey.
[17] Hannan, E., J. (1973) The asymptotic theory of linear time series models. J. Appl. Probab. **10**, 130–145.

[18] Handcock, M.S., Stein, M.L. (1993) A Bayesian analysis of kriging. Technometrics **35**, 403–410.
[19] E. Hille, *Analytic Function Theory. Vol. 1*, Ginn, Boston, 1959.
[20] Hosking, J., R., M., Fractional differencing, Journal of Biometrica 68, pp. 165–176 (1981).
[21] Huang, S.T., Cambanis, S. (1978). Stochastic and multiple Wiener integrals for Gaussian processes. The Annals of Probability **6**, 585–614.
[22] Kallenberg, O. (2002). Foundations of Modern Probability. Second edition, Springer, New York.
[23] Kierat, W., Sztaba, U. (2003): Distributions, Integral Transforms and Applications. Taylor and Francis, CRC Press.
[24] Knight, F. (1992). Foundations of the prediction process. Oxford, Clarendon Press.
[25] Kolmogorov, A.N. (1940). Wiener spiral and some other interesting curves in Hilbert space. Dokl. Akad. Nauk SSSR **26**, 115–118.
[26] Li, W.K. ,and McLeod, A.I., Fractional time series modeling, Biometrika, 73 (1986), pp. 217-221
[27] Lim, S.C., Teo, L.P. (2007). Weyl and Riemann-Liouville multifractional Ornstein–Uhlenbeck processes. Journal of Physics A **40**, 6035–6060.
[28] Maejima, M. (1983a): On a class of self-similar processes. Z. Wahrscheinlichkeitstheor. Verw. Geb.62, 235-245.
[29] Mandelbrot, B., Van Ness, J. (1968). Fractional Brownian motion, fractional noises and applications. SIAM Review **10**, 422–437.
[30] Meerschaert, M.M., Sikorskii, A. (2012). Stochastic Models for Fractional Calculus. De Gruyter, Berlin/Boston.
[31] M.M. Meerschaert and F. Sabzikar, Tempered fractional Brownian motion. Statist. Prob. Lett. 83 (2013), 2269–2275.
[32] M.M. Meerschaert and F. Sabzikar, Stochastic integration for tempered fractional Brownian motion. Stoch. Proc. Appl. 124 (2014), 2363–2387.
[33] Meerschaert, M.M., Sabzikar, F. Phanikumar, M.S., Zeleke, A. (2014): Tempered fractional time series model for turbulence in geophysical flows preprint.
[34] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: A fractional dynamics approach. Phys. Rep. 339 (2000), 1–77.
[35] R. Metzler and J. Klafter The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. J. Physics A 37 (2004), R161-R208.
[36] K. Miller and B. Ross, An Introduction to the Fractional Calculus and Fractional Differential Equations. Wiley and Sons, New York, 1993.
[37] Oldham, K.B., Spanier, J. (1974). The Fractional Calculus. Academic Press.
[38] Peiris M S, 2003 Statist. Meth. 5 156
[39] Pipiras, V., Taqqu, M. (2000): Convergence of weighted sums of random variables with long range dependence. Stochastic Processes and their Applications 90:157–174.
[40] Pipiras, V., Taqqu, M. (2000): Integration questions related to fractional

Brownian motion, Probability Theory and Related Fields **118**, 251–291 .

[41] Rosenblatt, M. (1961). Independence and dependence. Proc. 4th Berkeley Symp. Math. Statist. Probab., 431–443.

[42] Rudin, W., Principles of mathematical analysis, third edition, McGraw-Hill Book Co., New York (1988).

[43] Sabzikar, F. Meerschaert, M. M. and Chen, J. (2014). Tempered Fractional Calculus, Journal of Computational Physics, to appear `www.stt.msu.edu/users/mcubed/TFC.pdf`

[44] Samko, S.G, Kilbas, A.A., Marichev, O.I. (1993). Fractional Integrals and Derivatives. Gordon and Breach.

[45] Shitan, M., and Peiris, S., Approximate asymptotic variancecovariance matrix for the whittle estimators of GAR(1) parameters Comm. Statist. Theory Methods, 42 (5) (2013), pp. 756-770

[46] Sowell, F., Maximum likelihood estimation of stationary univariate fractionally integrated time series models, Journal of Econometrics 53, pp. 165–183, North-Holland (1992)

[47] Sowell, F., The fractional unit root distrbution, Econometrica 58, pp. 495–505 (1990)

[48] Samorodnitsky, G., Taqqu, M. (1994). Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Chapman & Hall.

[49] Taqqu, M., S. (1979). Convergence of integrated processes of arbitrary Hermite rank Probab. Theory Related Fields, **50 (1)** ,53-83.

[50] Tudor, C. (2013). Analysis of Variations for Self-similar Processes: A Stochastic Calculus Approach. Springer (2013).