PARAMETER ESTIMATION FOR ARTFIMA TIME SERIES

FARZAD SABZIKAR, IAN MCLEOD, AND MARK M. MEERSCHAERT

Abstract. The ARTFIMA model applies a tempered fractional difference to the standard ARMA time series. This paper develops parameter estimation methods for the ARTFIMA model, and demonstrates a new R package. Several examples illustrate the utility of the method.

1. Introduction

The autoregressive tempered fractionally integrated moving average (ARTFIMA) time series model was introduced in Meerschaert et al. [20], see also Sabzikar et al. [25]. The ARTFIMA model extends the TFI model from Giraitis et al. [7], who note that tempered fractionally integrated times series exhibit *semi-long range dependence*: Their covariance function resembles long range dependence for a number of lags, depending on the tempering parameter, but eventually decays exponentially fast. In this paper, we develop the mathematical foundation for ARTFIMA parameter estimation, and present a new R package `artfima` to fit data. Several examples from finance, geophysics, turbulence, and climate illustrate the fitting procedure, and the utility of the ARTFIMA model.

There are three reasons to consider the ARTFIMA model. The first is mathematical. The ARFIMA model without tempering is challenging to analyze, since the covariance function is not summable. The ARTFIMA model, however, has a summable covariance function, and hence its properties can be derived using standard methods. Since the tempering parameter can be made as small as we like, the *mathematically more tractable* ARTFIMA model can fit data that is usually modeled using the ARFIMA model.

The second reason is statistical. The spectral density function of the ARFIMA model is asymptotically a power law that diverges as the frequency approaches zero. However, in many applications, data fit this power law model only up to a low frequency cutoff, after which the observed spectral density remains bounded. A classical example is turbulence: According to the Kolmogorov model, the spectral density is proportional to $k^{-5/3}$ in the inertial range, where k is the frequency. Since the observed power spectrum in actual turbulence data typically remains bounded as $k \to 0$, the ARTFIMA model provides a *better fit*. See [20] for additional discussion.

MMM was partially supported by ARO MURI grant W911NF-15-1-0562 and NSF grants DMS-1462156 and EAR-1344280.

The third reason is conceptual. The ARTFIMA model with any order of tempered fractional differencing is a stationary time series. It has now become common in many applications to model long range dependence using either an ARFIMA model or a fractional Brownian motion. A fractional Brownian motion with Hurst index $H = 1/3$ satisfies the Kolmogorov model, as does an ARFIMA with $d = 5/6$. Both of these models have stationary increments. However, typical applications involve data that most scientists believe form a stationary sequence, see Molz et al. [22] for additional discussion. The ARTFIMA model provides a more suitable *stationary* time series model for such data.

2. The ARTFIMA model

In this section, we 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:

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

where $d > 0$, $d \notin \mathbb{Z}$, $\lambda > 0$, $Bf(t) = f(t-1)$ is the shift operator, 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 [19, 25] for more details.

The exact form of the tempered fractional difference operator (2.1) comes from the emerging field of tempered fractional calculus [18, 25]. The tempered fractional derivative

$$\mathbb{D}_t^{d,\lambda} f(t) = \lim_{h \to 0} h^{-\alpha} \sum_{j=0}^{\infty} (-1)^j \binom{d}{j} e^{-\lambda j h} f(t - jh) \tag{2.3}$$

has been applied in Finance [4, 5], Hydrology [17, 30] and Geophysics [20] to model processes with *semi-heavy tails* [1]. If $\lambda = 0$ then (2.3) reduces to the Riemann-Liouville fractional derivative [18, 24, 26] and (2.1) reduces to the fractional difference.

Definition 2.1. The discrete time stochastic process $\{X_t\}_{t \in \mathbb{Z}}$ follows an *autoregressive tempered fractional integrated moving average* time series, denoted by ARTFIMA(p, d, λ, q), if

$$Y_t = \Delta^{d,\lambda} X_t = (I - e^{-\lambda} B)^d X_t \tag{2.4}$$

follows the ARMA(p,q) model

$$Y_t - \sum_{j=1}^{p} \phi_j Y_{t-j} = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i} \tag{2.5}$$

where $\{Z_t\}_{t\in\mathbb{Z}}$ 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.

Assumption A: In this paper, we always assume the polynomials $\Phi(\cdot)$ and $\Theta(\cdot)$ have no common zeros and

$$\left|\Phi(z)\right|>0 \quad \text{and} \quad \left|\Theta(z)\right|>0 \tag{2.6}$$

for $|z|\leq 1$.

Theorem 2.2. *Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ is an* ARTFIMA(p,d,λ,q) *time series that satisfies Definition 2.1 with polynomials $\Phi(\cdot)$ and $\Theta(\cdot)$ that satisfy assumption A. Then: (a) X_t is causal, i.e., $X_t=\sum_{j=0}^{\infty} a_j^{-d,\lambda} Z_{t-j}$, where $\sum_{j=0}^{\infty}\left|a_j^{-d,\lambda}\right|<\infty$; and (b) X_t is invertible, i.e., $Z_t=\sum_{j=0}^{\infty} c_j^{d,\lambda} X_{t-j}$, where $\sum_{j=0}^{\infty}\left|c_j^{d,\lambda}\right|<\infty$. In particular, $\{X_t\}$ is a stationary time series for all $d\notin\mathbb{Z}$ and all $\lambda>0$.*

The proofs of all the results in this paper can be found in the Appendix.

Remark 2.3. The well-known ARFIMA(p,d,q) model is a special case of the ARTFIMA(p,d,λ,q) model with $\lambda=0$. Although the ARFIMA model is only stationary for $-.5<d<.5$, Theorem 2.2 shows that the ARTFIMA model is stationary for any $d\notin\mathbb{Z}$. This is because tempering avoids the unit root in the ARFIMA model. For example, the ARFIMA characteristic polynomial $(1-z)$ with $d=1$ has a unit root, but the ARTFIMA characteristic polynomial $(1-e^{-\lambda}z)$ with $d=1$ does not.

Remark 2.4. The case $d\in\mathbb{Z}$ is excluded in Definition 2.1 because, in that case, the ARTFIMA(p,d,λ,q) model reduces to an ARMA model. For integer $d>0$, $(1-e^{-\lambda}z)^d$ is a polynomial of degree d that has no zeroes on the unit disk. Then Definition 2.1 reduces to the ARMA$(p+d,q)$ model $\Phi(B)(I-e^{-\lambda}B)^d X_t=\Theta(B)Z_t$. For integer $d<0$, the fractional binomial coefficients are undefined, since the recursion $\Gamma(d)=\Gamma(d+1)/d$ involves a divide by zero. In this case, Definition 2.1 is formally equivalent to the ARMA$(p,q-d)$ model $\Phi(B)X_t=\Theta(B)(I-e^{-\lambda}B)^{-d}Z_t$. These models are causal and invertible under Assumption A.

In order to compute the covariance function of the ARTFIMA(p,d,λ,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 (2.6), $\Phi(x)$ 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 [28, Section 4]). Then the spectral density of U_t is

$$\begin{aligned} f_U(\nu) &= \frac{\sigma^2}{2\pi} \frac{\Big|\Theta(\omega)\Big|^2}{\Big|\Phi(\omega)\Big|^2} \\ &= \frac{\sigma^2}{2\pi} \Big|\Theta(\omega)\Big|^2 \prod_{j=1}^{p} (1-\rho_j\omega)^{-1}(1-\rho_j\omega^{-1})^{-1}, \end{aligned}$$

where $\omega = e^{-i\nu}$. According to Sowell [28, Section 4], the spectral density of U_t can also be written as

$$f_U(\nu) = \frac{\sigma^2}{2\pi} \sum_{l=-q}^{q} \psi(l)\omega^l \sum_{j=1}^{p} \omega^p \zeta_j \Big[\frac{{\rho_j}^{2p}}{(1-\rho_j\omega)} - \frac{1}{(1-\rho_j^{-1}\omega)} \Big], \tag{2.7}$$

where

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

and

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

The next theorem gives the spectral density and explicit form of the covariance function of ARTFIMA(p, d, λ, q).

Theorem 2.5. *Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ is an* ARTFIMA(p, d, λ, q) *time series that satisfies Definition 2.1, and that* (2.6) *holds. Then:*

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

$$f_X(\nu) = \frac{\sigma^2}{2\pi} \frac{\big|\Theta(e^{-i\nu})\big|^2}{\big|\Phi(e^{-i\nu})\big|^2} \ \big|1 - e^{-(\lambda+i\nu)}\big|^{-2d} \tag{2.8}$$

for $-\pi \leq \nu \leq \pi$.

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

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

using the notation in (2.7)*, where*

$$
\begin{aligned}
&C(d,\lambda,h,\rho)\\
&\quad=\rho^{2p}\sum_{m=0}^{\infty}\frac{\rho^m e^{-\lambda(h-m)}\Gamma(h-m+d)\ {}_2F_1(d;h-m+d;h-m+1;e^{-2\lambda})}{\Gamma(d)\Gamma(h-m+1)}\\
&\qquad+\sum_{n=1}^{\infty}\frac{\rho^n e^{-\lambda(h+n)}\Gamma(h+n+d)\ {}_2F_1(d;h+n+d;h+n+1;e^{-2\lambda})}{\Gamma(d)\Gamma(h+n+1)},
\end{aligned}
\tag{2.9}
$$

where the Gaussian hypergeometric function

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

is defined for all complex numbers a and b, all complex $|z|<1$ and real c not a negative integer.

An important special case of the ARTFIMA(p,d,λ,q) process is when $p=q=0$. In this case, we have

$$
X_t=(1-\ e^{-\lambda}B)^{-d}Z_t=\sum_{k=0}^{\infty}\omega_k^{-d,\lambda}Z_{t-k},\qquad t\in\mathbb{Z}
\tag{2.10}
$$

where $\omega_k^{-d,\lambda}$ is given by (2.2). The TFI(d,λ) model in (2.10) appeared in Giraitis et al. [7], who noted that for small $\lambda>0$, X_t has a covariance function that resembles the covariance function of a long memory model for an arbitrarily large number of lags, but eventually decays exponentially fast. Following [7], we say that X_t exhibits *semi-long range dependence*. Meerschaert et al. [20] apply the TFI(d,λ) model (2.10) to turbulent water velocity data from the Great Lakes.

Figure 1 illustrates the autocorrelation function (ACF) of the ARTFIMA$(0,d,\lambda,0)$ model. The differencing parameter d acts the same as in the more familiar ARFIMA model. A higher value of d gives a stronger correlation, so the ACF falls off more slowly with lag. A higher value of the tempering parameter λ makes the ACF fall off more rapidly. Section 4 shows several more examples of the theoretical ACF, plotted against the periodogram.

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,d,λ,q) time series, and we explicitly compute the asymptotic covariance matrix.

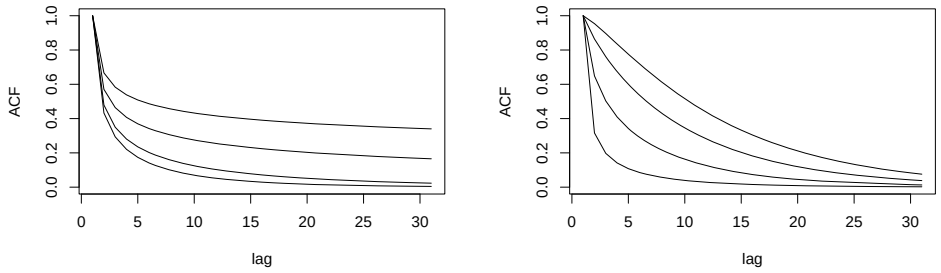

FIGURE 1. Theoretical autocorrelation function for the ARTFIMA$(0, d, \lambda, 0)$ model. The left panel shows the ACF for $d = .4$ and $\lambda = 0, .001, , 05, .1$ (descending). The right panel shows the ACF for $\lambda = .1$ and $d = .3, .6, .9, 1.2$ (ascending).

Recall from (2.8) that the ARTFIMA(p, d, λ, q) spectral density can be written in the form

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

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, d, λ, q) time series with sample size N and consider the periodogram

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

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 $\Omega_0 \subset \Omega$. Choose this set so that $(\sigma_0, \boldsymbol{\theta}_0)$ is an interior point.

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

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

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*

$$\mathbf{W} = \frac{1}{4\pi}\int_{-\pi}^{\pi} \left\{\frac{\partial \log K(\nu, \boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}}\right\}\left\{\frac{\partial \log K(\nu, \boldsymbol{\theta}_0)}{\partial\boldsymbol{\theta}}\right\}' d\nu. \tag{3.6}$$

Theorem 3.4. *The covariance matrix* $\mathbf{W}$ *in* (3.6) *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:

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

$$v_{2,2} = \frac{d^2e^{-2\lambda}}{1-e^{-2\lambda}}$$

$$v_{1,2} = v_{2,1} = d\ln(1-e^{-2\lambda}).$$

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

$$\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}) \sim -2e^{-\lambda}\cos\nu + e^{-2\lambda} \tag{3.7}$$

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

$$v_{1,1} \sim \int_{-\pi}^{\pi} (-2e^{-\lambda}\cos\nu + e^{-2\lambda})^2 \, d\nu = 4\pi(e^{-2\lambda} + \frac{e^{-4\lambda}}{2}) \tag{3.8}$$

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

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

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* (3.6).

Remark 3.8. It can be shown that, for large N, one 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.

4. Applications

This section provides several 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.

Example 4.1. Geophysical turbulence in water velocity data (cm/s) was measured in Lake Michigan, Lake Huron, and the Red Cedar River in Michigan, see [20] for further details. Figure 2 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 $d = 5/6$

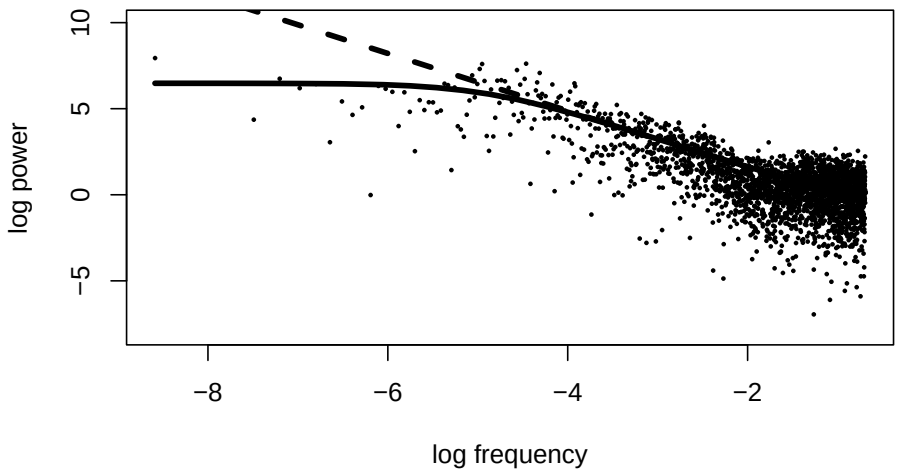

FIGURE 2. Spectral density of water velocity data (circles), fitted ARTFIMA spectrum (solid line), and ARFIMA spectrum (dotted line).

(from theory, Kolmogorov scaling, see [20] 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). The ARTFIMA model is stationary, and there is ample reason to consider the time series of velocity data as stationary. Figure 2 also plots the power law spectrum of the untempered ARFIMA with $d = 5/6$ (the classical Kolmogorov model). Both models fit the data for moderate frequencies, but the ARFIMA model shows a lack of fit for low frequencies. Hence we consider the ARFIMA model to be mis-specified, since its pure power law spectrum does not fit the data at low frequencies, and more importantly, the ARFIMA model with $d = 5/6$ is not stationary.

Example 4.2. The adjusted closing price C_t for AMZN stock from 1/3/2000 to 12/19/2017 ($n = 4520$) was used to compute log returns $R_t = \ln(C_t/C_{t-1})$, which appear uncorrelated. However, the squared log returns $X_t = R_t^2$ exhibit strong dependence, with an autocorrelation function that remains positive and statistically significant for more than 35 lags. An ARTFIMA model with $p = q = 0$ was fit using the `artfima` package. The fitted parameters are $d = 0.3$ and $\lambda = 0.025$. Figure 3 shows that the resulting model spectral density provides a reasonable fit to the periodogram, which follows a power law at moderate frequencies, but levels off at low frequencies. The untempered ARFIMA model fails to fit the periodogram at low

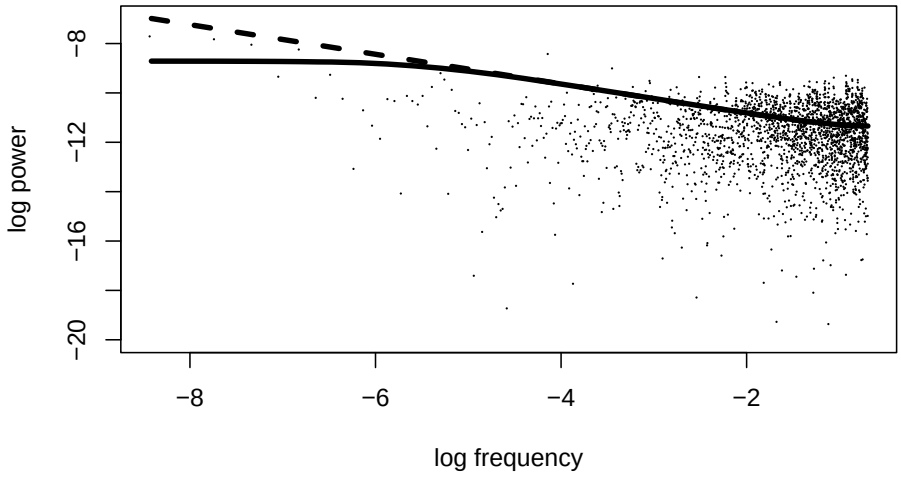

FIGURE 3. Spectral density of squared log returns for AMZN stock (circles), fitted ARTFIMA spectrum (solid line), and ARFIMA spectrum (dotted line).

frequencies, hence we consider the ARTFIMA model fit to be superior, evidence of semi-long range dependence in the squared log-returns.

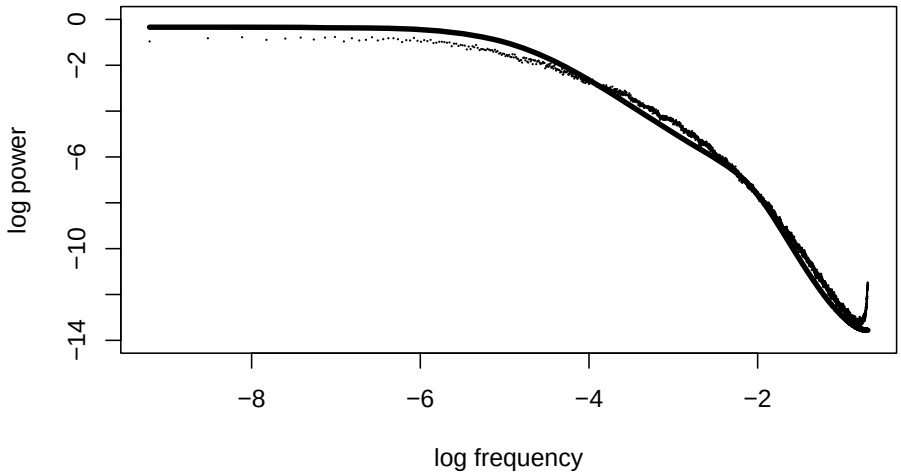

FIGURE 4. Spectral density of wall turbulence data (circles), and fitted ARTFIMA spectrum (solid line).

Example 4.3. The vorticity in a turbulent velocity flow near a wall was measured by Morrill-Winter et al. [23] at the High Reynolds Number Boundary Layer Wind Tunnel in Melbourne. This resulted in a time series with $n = 18,000,000$ observations.

We fit an ARTFIMA model with $p = q = 0$ to the first 10,000 observations, but the fit was inadequate. We obtained an adequate fit by including an autoregressive component with $p = 2$ and $q = 0$. The fitted parameters were $d = 1.34$ (.0486), $\lambda = .0531$ (.00646), $\phi_1 = 1.025$ (.0164), and $\phi_2 = -.0450$ (.01582). The model is causal and invertible. We plot the fitted spectral density against the smoothed periodogram in Figure 4. Welch's method with a segment length of $L = 2000$ was used to smooth the periodogram. The smoothed periodogram for the entire data set (not shown) is similar. The velocity field in this experiment is designed to be stationary. With $d = 1.34$, the corresponding ARFIMA time series model would have stationary increments. Hence the ARTFIMA model seems more appropriate. In fact, the ARFIMA model is mis-specified, since it is not stationary.

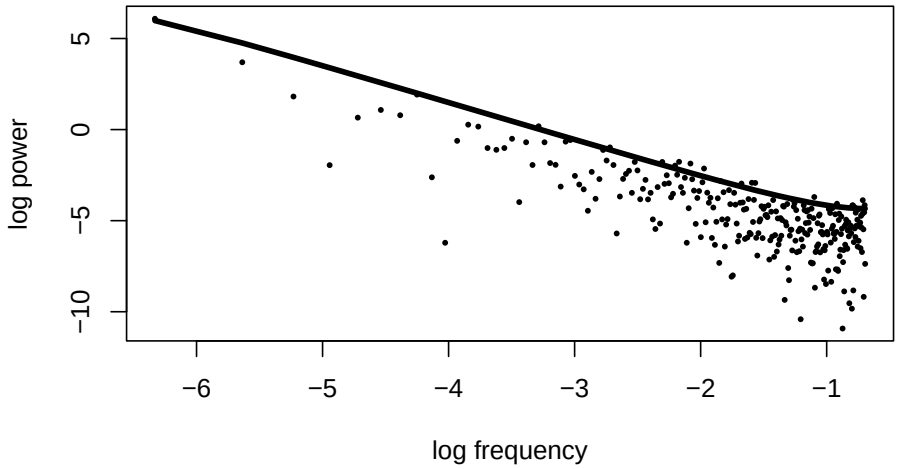

FIGURE 5. Spectral density of hydraulic conductivity data (circles), along with fitted ARTFIMA spectrum (line).

Example 4.4. In Hydrology, the hydraulic conductivity is measured in the field, and used to parameterize contaminant transport models. High resolution (1.5 cm) hydraulic conductivity data was measured at the MAcroDispersion Experimental (MADE) site in Mississippi by Liu et al. [15]. We fit the data from one vertical borehole (A121108) using the `artfima` package. The fitted parameters are $d = 1.01$ (.0245) and $\lambda = 0.00593$ (.00332). The resulting model spectral density and periodogram are plotted in Figure 5. Note that the fitted spectral density is almost a straight line. The same data was fit by Meerschaert et al. [21] to an ARFIMA model with $p = q = 0$ and d=0.9. The ARTFIMA model is very close to a simple first difference, with stationary increments, and the ARFIMA model with $d = 0.9$ also has stationary increments. However, the K field should be stationary based on the geology, and hence the ARTFIMA model with $\lambda > 0$ is preferred, as it provides a stationary time series model.

Indeed, the ARFIMA model is mis-specified, since the data comes from a stationary process, and the ARFIMA with $d > 0.5$ is nonstationary.

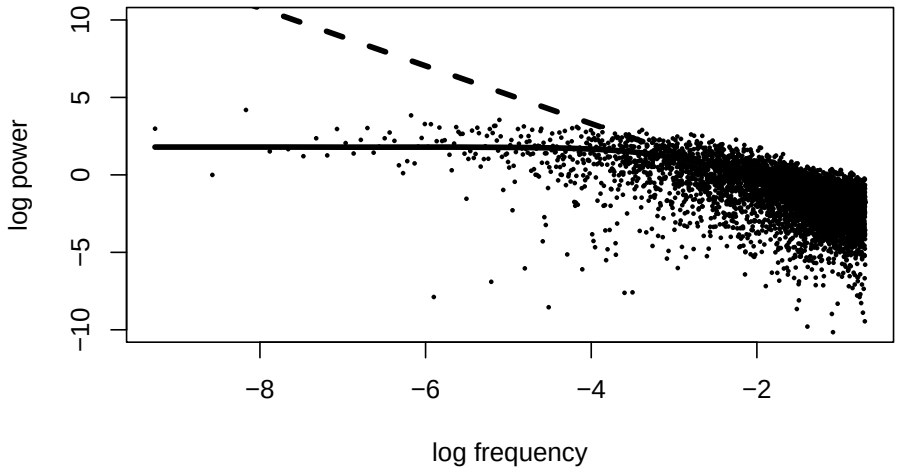

FIGURE 6. Spectral density of NARCCAP data climate data (circles), fitted ARTFIMA spectrum (solid line) and untempered ARFIMA spectrum (dotted line).

Example 4.5. A North American Regional Climate Change Assessment Program (NARCCAP) climate model was used to generate 29 years of daily maximum temperature data at 16,100 spatial locations in North America. The data is freely available at `www.narccap.ucar.edu`. We examine this data at one location, with sample size $n = 10,585$. The time series shows no evidence of a trend at this location, but significant mean variation and heteroscedasticity with respect to the season. Hence we subtract the seasonal mean, and divide by the seasonal standard deviation. We fit an ARTFIMA$(0, d, \lambda, 0)$ model to the standardized time series with $d = 0.933$ and $\lambda = 0.300$ using the `artfima` package. The fitted spectral density and the periodogram are shown in Figure 6. The spectral density levels off at low frequencies, consistent with the periodogram. The spectrum of the untempered ARFIMA model in Figure 6 shows a lack of fit at low frequencies. As there is no evidence of nonstationarity in the standardized time series, the stationary ARTFIMA model seems highly preferable to the stationary increment ARFIMA model, which is mis-specified due to its lack of stationarity.

Example 4.6. Tree-ring time series provide information about past climate [6, 29]. Hipel and McLeod [12, p. 364] fit an ARMA$(2, 1)$ model to a tree-ring time series from a Douglas Fir tree in Colorado. The time series contains annual indices for $n = 858$ consecutive values corresponding to the years 1107–1964. Table 1 shows

that an ARTFIMA$(0, d, \lambda, 0)$ = TFD model provides a better fit in terms of AIC and BIC. Both ARTFIMA and ARFIMA models were considered, as well as traditional ARMA models. Both the AIC and BIC selected TFD as the best model. The fitted parameters of the TFD model are $\lambda = 0.079$ (0.038) and $d = 0.529$ (0.0455). The diagnostic plot of the fitted ARTFIMA spectral density function and the observed periodogram (not shown) does not suggest lack-of-fit. The residual diagnostic plots and portmanteau diagnostic checks also confirm that the model is satisfactory.

The rows pAIC and pBIC give the plausibility scores. The plausibility is defined in a way similar to the relative likelihood in Fisher likelihood inference. If AIC_1, AIC_2, AIC_3 and AIC_4 are four AIC scores arranged in ascending order, the plausibility for the i-th score is $\exp(-(\text{AIC}_1 - \text{AIC}_i)/2), i = 1, 2, 3, 4$, and similarly for the BIC.

TABLE 1. Best models for Tree-ring Time Series

	best	2nd best	3rd best	4th best
Ω	ARTFIMA(0,0,0)	ARFIMA(2,0,1)	ARFIMA(1,0,2)	ARTFIMA(2,0,0)
AIC	8405.03	8405.83	8405.98	8406.39
pAIC	1.000	0.668	0.621	0.507
Ω	ARTFIMA(0,0,0)	ARIMA(1,0,1)	ARFIMA(0,0,0)	ARTFIMA(1,0,0)
BIC	8419.29	8424.30	8424.62	8425.94
pBIC	1.000	0.082	0.070	0.036

For forecasting purposes, the BIC is often recommended [16, and references therein]. From Table 1 we see that the ARMA(1,1) provides the second best BIC fit, and the FD = ARFIMA$(0, d, 0)$ model ranks third. Although there is little difference in the BIC values between the FD and ARMA(1,1), Figure 7 shows that the forecast performance is very different. The FD forecast very slowly damps out to the mean level, which is shown by the dotted line. The TFD forecast damps out much faster to the mean. The ARMA$(1, 1)$ forecasts are similar to the TFD forecasts. The last ten years of observed data, shown in Figure 7, suggest there is little long-term predictability, so the TFD or ARMA$(1, 1)$ forecasts may be preferable.

5. Appendix

Proof of Theorem 2.2: By inverting the operator $\Delta^{d,\lambda}$, we get

$$W_t = \Delta^{-d,\lambda} Z_t = \sum_{j=0}^{\infty} (-1)^j e^{-\lambda j} \binom{-d}{j} Z_{t-j} = \sum_{j=0}^{\infty} \omega_j^{-d,\lambda} Z_{t-j}. \tag{5.1}$$

Since $\omega_j^{d,\lambda}$ has the same sign for all large j (e.g., see [18, Eq. (2.4)]) and

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

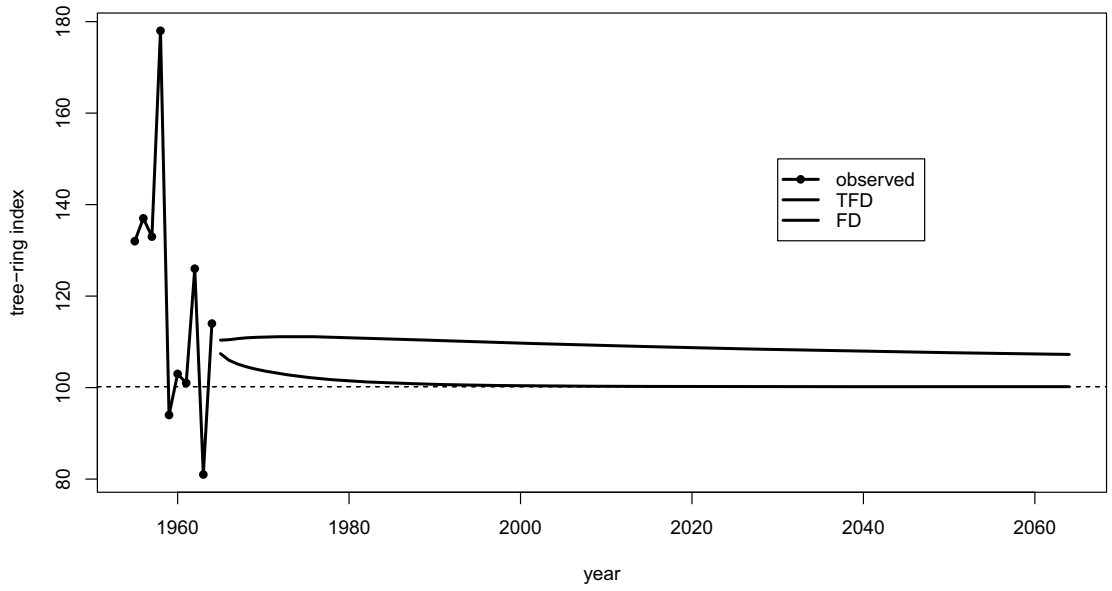

FIGURE 7. Last ten years of observed tree-ring data, along with mean (dotted line), TFD forecast (bottom solid line), and FD (top solid line) forecast for 100 years.

for any fixed $d \notin \mathbb{Z}$ by the fractional binomial formula (e.g., see Hille [13, p. 147]) it follows that

$$\sum_{j=0}^{\infty} \left|\omega_j^{d,\lambda}\right| < \infty \tag{5.2}$$

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

$$a_j^{-d,\lambda} = \sum_{s=0}^{j} \omega_s^{-d,\lambda} b_{j-s} \tag{5.3}$$

for $j \geq 0$. Since X_t satisfies (2.4) and (2.5), we can write

$$X_t = \Delta^{-d,\lambda} \frac{\Theta(B)}{\Phi(B)} Z_t = \left(\sum_{j=0}^{\infty} a_j^{-d,\lambda} B^j \right) Z_t = \sum_{j=0}^{\infty} a_j^{-d,\lambda} Z_{t-j}. \tag{5.4}$$

where $a_j^{-d,\lambda}$ is given by (5.3). Under assumption (2.6), $\left|\Theta(z)/\Phi(z)\right| < \infty$, for $|z| \leq 1+\varepsilon$, and the convergence of the series $\Theta(z)/\Phi(z)$ implies that $|b_j| \leq C(1+\varepsilon)^{-j}$ for $j \geq 0$ (e.g., see [8, Theorem 7.2.3] or [3, Theorem 3.1.1]). By applying (5.3), we have

$$\begin{aligned} \sum_{j=0}^{\infty} \left|a_j^{-d,\lambda}\right| &= \sum_{j=0}^{\infty} \Big| \sum_{s=0}^{j} \omega_s^{-d,\lambda} b_{j-s} \Big| \leq \sum_{j=0}^{\infty} \sum_{s=0}^{j} \left|\omega_s^{-d,\lambda}\right| \, \left|b_{j-s}\right| \\ &= \sum_{s=0}^{\infty} \sum_{j=s}^{\infty} \left|\omega_s^{-d,\lambda}\right| \, \left|b_{j-s}\right| = \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} \left|\omega_s^{-d,\lambda}\right| \, \left|b_t\right| \\ &\leq C \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} \left|\omega_s^{-d,\lambda}\right| (1+\varepsilon)^{-t} = C \sum_{s=0}^{\infty} \left|\omega_s^{-d,\lambda}\right| \sum_{t=0}^{\infty} (1+\varepsilon)^{-t} < \infty \end{aligned} \tag{5.5}$$

since $\sum_{s=0}^{\infty} \left|\omega_s^{-d,\lambda}\right|$ is finite by (5.2) and $\sum_{t=0}^{\infty} (1+\varepsilon)^{-t} < \infty$. Now it follows from Brockwell and Davis [3, Proposition 3.1.2] that $\{X_t\}$, given by (5.4), 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 [8, Chapter 2]). Because $\sum_{j=0}^{\infty} a_j^{-d,\lambda} e^{-ij\nu} = (1-e^{-(\lambda+i\nu)})^{-d} \Theta(e^{-i\nu})/\Phi(e^{-i\nu})$, [8, Theorem 2.2.1] implies that

$$X_t = \int_{-\pi}^{\pi} e^{it\nu} (1 - e^{-(\lambda+i\nu)})^{-d} \, \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})} dW(\nu). \tag{5.6}$$

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

$$c_j^{d,\lambda} = \sum_{s=0}^{j} \omega_s^{d,\lambda} c_{j-s} \tag{5.7}$$

for $j \geq 0$. 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 (5.6) to get:

$$
\begin{aligned}
\sum_{j=0}^{\infty} c_j^{d,\lambda} B^j X_t &= \sum_{j=0}^{\infty} c_j^{d,\lambda} X_{t-j} \\
&= \sum_{j=0}^{\infty} c_j^{d,\lambda} \int_{-\pi}^{\pi} e^{i(t-j)\nu} (1 - e^{-(\lambda+i\nu)})^{-d} \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})} dW(\nu) \\
&= \int_{-\pi}^{\pi} \Big(\sum_{j=0}^{\infty} c_j^{d,\lambda} e^{-ij\nu} \Big) e^{it\nu} (1 - e^{-(\lambda+i\nu)})^{-d} \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})} \ dW(\nu) \\
&= \int_{-\pi}^{\pi} (1 - e^{-(\lambda+i\nu)})^d (1 - e^{-(\lambda+i\nu)})^{-d} \frac{\Phi(e^{-i\nu})}{\Theta(e^{-i\nu})} \frac{\Theta(e^{-i\nu})}{\Phi(e^{-i\nu})} e^{it\nu} \ dW(\nu) \\
&= Z_t.
\end{aligned}
\tag{5.8}
$$

Under assumption (2.6), $\big|\Phi(z)/\Theta(z)\big| < \infty$, for $|z| \leq 1 + \varepsilon$, and the convergence of the series $\Phi(z)/\Theta(z)$ implies that

$$
|c_j| \leq K(1 + \varepsilon)^{-j}, \tag{5.9}
$$

for $j \geq 0$, see [8, Theorem 7.2.3]. Now apply (5.7) and (5.9) to write

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

since $\sum_{s=0}^{\infty} \big|\omega_s^{d,\lambda}\big|$ is finite by (5.2) and $\sum_{t=0}^{\infty} (1+\varepsilon)^{-t} < \infty$. Now [3, Proposition 3.1.2] implies that $\{Z_t\}$ in (5.8) is stationary and converges absolutely with probability one, so X_t is invertible, which proves (b). □

Proof of Theorem 2.5: (a) Use (2.4) and (2.5) to write $X_t = \psi_X(B) Z_t$ where $\psi_X(z) = (1 - e^{-\lambda} z)^{-d} \Theta(z)/\Phi(z)$. 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 [3]). Since the white noise process $\{Z_t\}$ has spectral density $f_Z(\nu) = (2\pi)^{-1}\sigma^2$, it follows that (2.8) holds.

(b) Using (2.7), we can rewrite the spectral density of X_t in the form

$$
\begin{aligned}
f_X(\nu) &= \Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d} h_U(\nu) \\
&= \frac{\sigma^2}{2\pi}\frac{\Big|\Theta(\omega)\Big|^2}{\Big|\Phi(\omega)\Big|^2}\Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d} \\
&= \frac{\sigma^2}{2\pi}\sum_{l=-q}^{q}\sum_{j=1}^{p}\psi(l)\zeta_j\Big[\frac{{\rho_j}^{2p}}{(1-\rho_j\omega)}-\frac{1}{(1-\rho_j^{-1}\omega)}\Big] \\
&\qquad \Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d}\omega^{p+l}.
\end{aligned}
\tag{5.10}
$$

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

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

where

$$
\begin{aligned}
C(d,\lambda,h,\rho) = \int_{-\pi}^{\pi}&\Big[\frac{\rho^{2p}}{(1-\rho\omega)}-\frac{1}{(1-\rho^{-1}\omega)}\Big] \\
&\Big(1-e^{-(\lambda+i\nu)}\Big)^{-d}\Big(1-e^{-(\lambda-i\nu)}\Big)^{-d}e^{i\nu h}\,d\nu.
\end{aligned}
\tag{5.11}
$$

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

$$f_W(\nu) = \frac{\sigma^2}{2\pi}\left|1 - e^{-(\lambda+i\nu)}\right|^{-2d} = \frac{\sigma^2}{2\pi}\left(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}\right)^{-d}. \tag{5.12}$$

we can then write

$$\begin{aligned}
&C(d,\lambda,h,\rho)\\
&= \int_{-\pi}^{\pi} \rho^{2p} \sum_{m=0}^{\infty} (\rho e^{-i\nu})^m \Big(1 - e^{-(\lambda+i\nu)}\Big)^{-d} \Big(1 - e^{-(\lambda-i\nu)}\Big)^{-d} e^{i\nu h}\, d\nu\\
&+ \int_{-\pi}^{\pi} \sum_{n=1}^{\infty} (\rho e^{i\nu})^n \Big(1 - e^{-(\lambda+i\nu)}\Big)^{-d} \Big(1 - e^{-(\lambda-i\nu)}\Big)^{-d} e^{i\nu h}\, d\nu\\
&= \rho^{2p} \sum_{m=0}^{\infty} \rho^m \int_{-\pi}^{\pi} \frac{2\pi}{\sigma^2} f_W(\nu) e^{i\nu(h-m)}\, d\nu + \sum_{n=1}^{\infty} \rho^n \int_{-\pi}^{\pi} \frac{2\pi}{\sigma^2} f_W(\nu) e^{i\nu(h+n)}\, d\nu\\
&= \rho^{2p} \sum_{m=0}^{\infty} \rho^m \frac{2\pi}{\sigma^2} \gamma_W(h-m)\, d\nu + \sum_{n=1}^{\infty} \rho^n \frac{2\pi}{\sigma^2} \gamma_W(h+n).
\end{aligned} \tag{5.13}$$

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

$$\begin{aligned}
\gamma_W(k) &= \int_{-\pi}^{\pi} \cos(k\nu) f_W(\nu)\, d\nu\\
&= \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \frac{\cos(k\nu)}{(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda})^d}\, d\nu\\
&= \frac{\sigma^2}{2\pi} \int_{0}^{2\pi} \frac{(-1)^k \cos(k\nu')}{(1 + 2e^{-\lambda}\cos\nu' + e^{-2\lambda})^d}\, d\nu' \ [\nu' := \nu + \pi]\\
&= \sigma^2 \frac{e^{-\lambda k}\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)}\ {}_2F_1(d; k+d; k+1; e^{-2\lambda}),
\end{aligned} \tag{5.14}$$

where we applied the integral formula (see [9], 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 (5.14) into (5.13) it follows that (2.9) holds.

To complete the proof, we need to justify the interchange of the sum and integral in (5.13). Observe that for any $d \notin \mathbb{Z}$ and any $|\rho| < 1$ we have

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

where we applied (5.14) for $k = 0$. □

Proof of Theorem 3.2: First we note that the ARTFIMA(p, d, λ, q) time series is ergodic since it is an infinite moving average with ergodic innovations $\{Z_t\}_{t\in\mathbb{Z}}$, see [27, Corollarly 2.1.8]. Then we can apply Theorem 8.2.1 in [8] (see also Theorem 1 in Hannan [11]). 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 (5.3).

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 (5.5) it follows that $|a_j^{-d,\lambda}| \leq C_3$ for all $j \geq 0$. Then by a similar argument to that of (5.5) it can be shown that

$$\sum_{j=0}^{\infty} j (a_j^{-d,\lambda})^2 \leq C_3 \sum_{j=0}^{\infty} j |a_j^{-d,\lambda}| < \infty. \tag{5.15}$$

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. □.

Proof of Theorem 3.3. The result follows from Theorem 2 in Hannan [11] or Theorem 8.3.1 in Giraitis et al. [8], 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 (5.15) 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. □

Proof of Theorem 3.4 : The matrix $\mathbf{I}$ is the same as the matrix $\mathbf{W}$ in Brockwell and Davis [3, Theorem 10.8.2, p. 386] for the asymptotic covariance of the ARMA(p,q) parameter estimates, see also Box and Jenkins [2, p.240] or Li and McLeod [14, Section 3]. This proves part (1).

To prove the remaining parts, write

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

Observe that

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

and then

$$\begin{aligned}\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial\phi_j} &= \frac{\partial \log(\left|\Phi(e^{-i\nu})\right|^2)}{\partial\phi_j} = \frac{\partial}{\partial\phi_j}\big\{\log(\Phi(e^{-i\nu})) + \log(\Phi(e^{i\nu}))\big\}\\ &= \frac{\partial}{\partial\phi_j}\log(1+\phi_1 e^{-i\nu} + \ldots + \phi_j e^{-i\nu j} + \ldots + \phi_p e^{-i\nu p})\\ &+ \frac{\partial}{\partial\phi_j}\log(1+\phi_1 e^{i\nu} + \ldots + \phi_j e^{i\nu j} + \ldots + \phi_p e^{i\nu p})\\ &= \frac{e^{-i\nu j}}{1+\phi_1 e^{-i\nu} + \ldots + \phi_j e^{-i\nu j} + \ldots + \phi_p e^{-i\nu p}}\\ &+ \frac{e^{i\nu j}}{1+\phi_1 e^{i\nu} + \ldots + \phi_j e^{i\nu j} + \ldots + \phi_p e^{i\nu p}}\\ &= e^{-i\nu j}\Phi^{-1}(e^{-i\nu}) + e^{i\nu j}\Phi^{-1}(e^{i\nu}).\end{aligned} \tag{5.16}$$

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

$$\begin{aligned}\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial \theta_j} &= \frac{\partial \log(\big|\Theta(e^{-i\nu})\big|^2)}{\partial \theta_j} = \frac{\partial}{\partial \theta_j}\big\{\log(\Theta(e^{-i\nu})) + \log(\Theta(e^{i\nu}))\big\}\\ &= \frac{\partial}{\partial \theta_j}\log(1+\theta_1 e^{-i\nu}+\ldots+\theta_j e^{-i\nu j}+\ldots+\theta_q e^{-i\nu q})\\ &+ \frac{\partial}{\partial \theta_j}\log(1+\theta_1 e^{i\nu}+\ldots+\theta_j e^{i\nu j}+\ldots+\theta_q e^{i\nu q})\\ &= \frac{e^{-i\nu j}}{1+\theta_1 e^{-i\nu}+\ldots+\theta_j e^{-i\nu j}+\ldots+\theta_q e^{-i\nu q}}\\ &+ \frac{e^{i\nu j}}{1+\theta_1 e^{i\nu}+\ldots+\theta_j e^{i\nu j}+\ldots+\theta_q e^{i\nu q}}\\ &= e^{-i\nu j}\Theta^{-1}(e^{-i\nu}) + e^{i\nu j}\Theta^{-1}(e^{i\nu}).\end{aligned} \tag{5.17}$$

It is also easy to see that

$$\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial d} = -\log(1-2e^{-\lambda}\cos\nu+e^{-2\lambda}) \tag{5.18}$$

and

$$\frac{\partial \log K(\nu,\boldsymbol{\theta}_0)}{\partial \lambda} = -2de^{-\lambda}\Big(\frac{\cos\nu-e^{-\lambda}}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big) \tag{5.19}$$

Then part (2) follows easily.

As for (3), the formula for $v_{1,1}$ follows immediately from (5.18). A useful approximation will be provided in a remark following the proof. Next write

$$\begin{aligned}v_{2,2} &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\Big(-2de^{-\lambda}\Big(\frac{\cos\nu-e^{-\lambda}}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big)\Big)^2 d\nu\\ &= \frac{d^2}{4\pi}\int_{-\pi}^{\pi}\Big(1+\frac{e^{-2\lambda}-1}{1-2e^{-\lambda}\cos\nu+e^{-2\lambda}}\Big)^2 d\nu\\ &= \frac{d^2}{4\pi}\int_{-\pi}^{\pi}\Big[1+\frac{(e^{-2\lambda}-1)^2}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)^2}+\frac{2(e^{-2\lambda}-1)}{1+e^{-2\lambda}-2e^{-\lambda}\cos\nu}\Big]d\nu.\end{aligned} \tag{5.20}$$

Observe that

$$\begin{aligned}\int_{-\pi}^{\pi}\frac{d\nu}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)^2} &= 2\int_0^{\pi}\frac{d\nu}{(1+e^{-2\lambda}-2e^{-\lambda}\cos\nu)^2}\\ &= \frac{2\pi(1+e^{-2\lambda})}{(1-e^{-2\lambda})^3}\end{aligned} \tag{5.21}$$

where we used the standard formula

$$\int_0^{\pi}\frac{d\nu}{(1+a^2-2a\cos\nu)^n} = \frac{\pi}{(1-a^2)^n}\sum_{k=0}^{n-1}\frac{(n+k-1)!}{(k!)^2(n-k-1)!}\Big(\frac{a^2}{1-a^2}\Big)^k \tag{5.22}$$

for $a^2 < 1$ (see [9, p. 608]). By applying (5.22) for $n = 1$, we also have

$$\int_{-\pi}^{\pi} \frac{d\nu}{(1 + e^{-2\lambda} - 2e^{-\lambda}\cos\nu)} = \frac{2\pi}{1 - e^{-2\lambda}}. \tag{5.23}$$

Therefore, from (5.20), (5.21) and (5.23):

$$v_{2,2} = \frac{d^2}{4\pi}\Big[2\pi + \frac{2\pi(1 + e^{-2\lambda})}{(1 - e^{-2\lambda})} - 4\pi\Big] = \frac{d^2 e^{-2\lambda}}{1 - e^{-2\lambda}} \tag{5.24}$$

as claimed. Finally, write

(5.25)

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

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

$$\begin{aligned} I_1 &= \nu\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda})\big]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{2e^{-\lambda}\nu\sin\nu}{1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}}\, d\nu \\ &= 2\pi\ln(1 + e^{-\lambda})^2 - 4e^{-\lambda}\int_0^{\pi} \frac{\nu\sin\nu}{1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}}\, d\nu \\ &= 4\pi\ln(1 + e^{-\lambda}) - 4\pi\ln(1 + e^{-\lambda}) = 0, \end{aligned} \tag{5.26}$$

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 [9, 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 [9, No. 16, p. 925]). In our case $a = b = e^{-\lambda}$. Therefore

$$\begin{aligned} I_2 &= (e^{-2\lambda} - 1) \int_{-\pi}^{\pi} \frac{\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda})}{1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}}\, d\nu \\ &= 2(e^{-2\lambda} - 1) \int_{0}^{\pi} \frac{\ln(1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda})}{1 - 2e^{-\lambda}\cos\nu + e^{-2\lambda}}\, d\nu \\ &= \frac{4\pi \ln(1 - e^{-2\lambda})}{1 - e^{-2\lambda}}(e^{-2\lambda} - 1) = -4\pi \ln(1 - e^{-2\lambda}). \end{aligned} \tag{5.27}$$

Now, from (5.25), (5.26) and (5.27), we have

$$v_{1,2} = v_{2,1} = \frac{-d}{4\pi}\big[- 4\pi \ln(1 - e^{-2\lambda})\big] = d \ln(1 - e^{-2\lambda}), \tag{5.28}$$

and this completes the proof of (3). □

Proof of Theorem 3.6. The proof follows from Theorem 1 in [11]. □

Proof of Theorem 3.7. This follows immediately from Theorem 3.3, Theorem 3.4, and Theorem 10.8.2 in Brockwell and Davis [3]. □

6. Acknowledgments

We would like to thank Mantha Phanikumar, Michigan State University, for access to the water velocity data, and Joe Klewicki, University of Delaware, for the wall turbulence data, as well as helpful discussions. We would also like to thank two anonymous referees for their insightful comments.

References

[1] O.E. Barndorff-Nielsen (1998) Processes of normal inverse Gaussian type. *Finance Stoch.* **2**, 41–68.
[2] G.E.P. Box and G.M. Jenkins (1976) *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
[3] P.J. Brockwell and R.A. Davis (1991) *Time Series: Theory and Methods.* 2nd Ed., Springer-Verlag, New York.
[4] P. Carr, H. Geman, D. B. Madan, and M. Yor (2002) The fine structure of asset returns: An empirical investigation. *J. Business* **75**, 303–325.
[5] P. Carr, H. Geman, D. B. Madan, and M. Yor (2003) Stochastic volatility for Lévy processes. *Math. Finance* **13**, 345–382.
[6] H.C. Fritts, T.J. Blasing, B.P. Hayden, and J.E. Kutzbach (1971) Multivariate Techniques for Specifying Tree-growth and Climatic Relationships and for Reconstructing Anomalies in Paleoclimate. *J. Appl. Meteorology* **10**, 845–864.
[7] L. Giraitis, P. Kokoszka and R. Leipus (2000) Stationary ARCH models: dependence structure and central limit theorem. *Econometric Th.* **16**, 3–22.
[8] L. Giraitis, H.L. Koul, and D. Surgailis (2012) *Large Sample Inference for Long Memory Processes*. World Scientific, Singapore.
[9] I.S. Gradshteyn and I.M. Ryzhik (2007) *Table of Integrals, Series and Products*. 7th Ed., Academic Press, New York.

[10] J.D. Hamilton (1994) *Time Series Analysis*. Princeton University Press, New Jersey.
[11] E.J. Hannan (1973) The asymptotic theory of linear time series models. *J. Appl. Probab.* **10**, 130–145.
[12] K.W. Hipel and A.I. McLeod (1994) *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, New York.
[13] E. Hille (1959) *Analytic Function Theory*. Vol. 1, Ginn, Boston.
[14] W.K. Li and A.I. McLeod (1986) Fractional time series modeling. *Biometrika* **73**, 217–221.
[15] G. Liu, J. J. Butler Jr, G. C. Bohling, E. Reboulet, S. Knobbe, and D. W. Hyndman (2009) A new method for high-resolution characterization of hydraulic conductivity. *Water Resour. Res.*, **45**(8), W08202.
[16] A.I. McLeod (1993) Parsimony, model adequacy and periodic correlation in forecasting time series. *International Statist. Rev.* **61**, 387–393.
[17] M.M. Meerschaert, Y. Zhang, and B. Baeumer (2008). Tempered anomalous diffusion in heterogeneous systems. *Geophys. Res. Lett.* **35**, L17403.
[18] M.M. Meerschaert and A. Sikorskii (2012) *Stochastic Models for Fractional Calculus*. De Gruyter, Berlin/Boston.
[19] M.M. Meerschaert and F. Sabzikar, Stochastic integration for tempered fractional Brownian motion. Stoch. Proc. Appl. 124 (2014), 2363–2387.
[20] M.M. Meerschaert, F. Sabzikar, M.S. Phanikumar and A. Zeleke (2014) Tempered fractional time series model for turbulence in geophysical flows. *J. Stat. Mech.* **2014**, P09023.
[21] M.M. Meerschaert, M. Dogan, R.L. Van Dam, D.W. Hyndman, and D.A. Benson (2013) Hydraulic conductivity fields: Gaussian or not? *Water Resour. Res.* **49**, 4730–4737.
[22] F.J. Molz, H. Rajaram, and S. Lu (2004) Stochastic fractal-based models in subsurface hydrology: Origins, applications, limitations and future research questions. *Rev. Geophys.* **42**, RG1002.
[23] C. Morrill-Winter, J. Klewicki, R. Baidya, and I. Marusic (2015) Temporally optimized spanwise vorticity sensor measurements in turbulent boundary layers. *Exp. Fluids* **56**, 216.
[24] Oldham, K.B., Spanier, J. (1974). The Fractional Calculus. Academic Press.
[25] Sabzikar, F. Meerschaert, M. M. and Chen, J. (2014). Tempered Fractional Calculus. *J. Comput. Phys.* 293 (2015), 14–28.
[26] Samko, S.G, Kilbas, A.A., Marichev, O.I. (1993). Fractional Integrals and Derivatives. Gordon and Breach.
[27] Samorodnitsky, G. *Stochastic Processes and Long Range Dependence*. Spring Series in Operations Research and Financial Engineering. Spring International Publishing, 2016. ISBN 978-3-319-45575-4.
[28] F. Sowell (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *J. Econometrics* **53**, 165–183.
[29] M.R. Schofield. R.J. Barker, A. Gelman , E.R. Cook, and K.R. Briffa (2016) Model-Based Approach to Climate Reconstruction Using Tree-Ring Data. *J. Amer. Statist. Assoc.* **111**, 93–106.
[30] Y. Zhang and M.M. Meerschaert (2011) Gaussian setting time for solute transport in fluvial systems. *Water Resour. Res.* **47**, W08601.

Farzad Sabzikar, Department of Statistics, Iowa State University, Ames, IA 50011
E-mail address: sabzikar@iastate.edu
URL: http://sabzikar.public.iastate.edu/

Ian McLeod, Department of Statistics, Western University, Canada
E-mail address: aim@stats.uwo.ca

Mark M. Meerschaert, Department of Statistics and Probability, Michigan State University, East Lansing MI 48823
E-mail address: mcubed@stt.msu.edu
URL: http://www.stt.msu.edu/users/mcubed/