function of $(\mathbf{d}_0, \ldots, \mathbf{d}_{J-K}^{'})'$ is given by (after dropping the constant term)

$$\ell(d, \sigma^2) = -\frac{1}{2} \sum_{j=0}^{J-K} \left[2^j \log (2^{2(J-j)d} \sigma^2) + \sum_{k=1}^{2^j} \frac{d_{jk}^2}{2^{2(J-j)d} \sigma^2} \right]$$

The approximate MLE of d and σ^2 can be obtained by maximizing $\ell(d, \sigma^2)$. In the next section we report the results of a Monte Carlo experiment on the finite sample distribution of the approximate MLE of d.

4. Monte Carlo Results

We generated $\{Y_t\}$ from an ARFIMA(p, d, q) process with (p, q) = (0, 0), (1, 0) and (0, 1). N is taken to be 1024 and 2048 (i.e., J is 10 and 11, respectively). To examine the effects of truncating the wavelet coefficients, we let $K = 2$, 3 and 4. The approximate MLE of d is then denoted as $W(2)$, $W(3)$ and $W(4)$, respectively. As a comparison, we also calculated the frequency-domain MLE (denoted as MLE for short) of d.[2] Estimates of the bias and root mean squared errors (RMSE) of each estimate are summarized in Tables 1, 2 and 3. The results are based on Monte Carlo samples of 1000 each.[3]

From Table 1 (for (p, q) = (0, 0)) we observe that $W(2)$ performs quite well relative to the MLE. For $d \geq 0.2$, truncating more wavelet coefficients reduces the biases of W, but this is achieved at the expense of increasing the RMSE.

From Table 2 (for (p, q) = (1, 0)) we can see that $W(2)$ performs quite badly for $\phi = 0.4$ or 0.6. Indeed for $\phi = 0.6$, even $W(3)$ gives rise to rather large RMSE. $W(4)$, however, appears to be satisfactory compared to the MLE. The same conclusion applies to the results in Table 3 for (p, q) = (0, 1).

Overall the results suggest that while $W(4)$ has the largest RMSE for pure fractional processes, it performs favourably against the MLE and W estimates with fewer truncations when the serial correlation in the short-memory component is present. As the W estimate has an advantage over the MLE in that the structure of the short-memory component need not be specified, it provides a robust estimate for d.

[2]See Beran (1994, Chapter 6) for the details of the frequency domain MLE.

[3]All computations performed in this paper were coded in GAUSS with the application modules TSM and MAXLIK. TSM contains procedures for the generation of observations following an ARFIMA process and the computation of the frequency-domain MLE of the fractional differencing parameter d.

References

[1] Johnstone, J.M. and B.W. Silverman, 1997, "Wavelet Threshold Estimators for Data with Correlated Noise", *Journal of the Royal Statistical Society, Series B*, **59**, 319 - 351.

[2] McCoy, E.J. and A.T. Walden, 1996, "Wavelet Analysis and Synthesis of Stationary Long-Memory Processes", *Journal of Computational and Graphical Statistics*, **5**, 26 - 56.

[3] Beran, J., 1994, *Statistics for Long-Memory Processes*, New York: Chapman & Hall.

Table 1: Monte Carlo Results of the Wavelet Approximate MLE and the Frequency-Domain MLE for ARFIMA(0, d, 0) Model

d	Estimation Method	Sample Size 1024	2048
0.1	W(2)	−0.0069	−0.0047
		0.0339	0.0249
	W(3)	−0.0063	−0.0035
		0.0484	0.0336
	W(4)	−0.0125	−0.0058
		0.0732	0.0513
	MLE	−0.0005	0.0007
		0.0245	0.0173
0.2	W(2)	−0.0106	−0.0095
		0.0348	0.0243
	W(3)	−0.0061	−0.0058
		0.0523	0.0322
	W(4)	−0.0097	−0.0058
		0.0801	0.0513
	MLE	0.0023	0.0009
		0.0246	0.0173
0.3	W(2)	−0.0108	−0.0104
		0.0349	0.0266
	W(3)	−0.0027	−0.0023
		0.0470	0.0332
	W(4)	−0.0035	0.0017
		0.0722	0.0510
	MLE	0.0062	0.0021
		0.0272	0.0174
0.4	W(2)	−0.0118	−0.0090
		0.0379	0.0261
	W(3)	−0.0015	0.0017
		0.0548	0.0361
	W(4)	−0.0053	0.0040
		0.0795	0.0511
	MLE	0.0096	0.0048
		0.0281	0.0180

Notes: The first figure in each cell is the estimated bias and the second figure is the estimated root mean squared error (RMSE). For example, −0.0069 is the estimated bias, while 0.0339 is the estimated RMSE for $d = 0.1$, $W(2)$ and $N = 1024$.

Table 2: Monte Carlo Results of the Wavelet Approximate MLE and the Frequency-Domain MLE for ARFIMA(1, d, 0) Model

ϕ	d	Estimation Method	Sample Size 1024	2048
0.2	0.2	W(2)	0.0512	0.0524
			0.0626	0.0570
		W(3)	0.0194	0.0195
			0.0536	0.0372
		W(4)	0.0006	0.0051
			0.0755	0.0461
		MLE	0.0017	0.0016
			0.0529	0.0332
0.4	0.2	W(2)	0.1445	0.1450
			0.1485	0.1471
		W(3)	0.0612	0.0630
			0.0797	0.0719
		W(4)	0.0171	0.0175
			0.0754	0.0539
		MLE	−0.0020	0.0040
			0.0742	0.0471
0.6	0.2	W(2)	0.2932	0.2894
			0.2957	0.2906
		W(3)	0.1480	0.1469
			0.1562	0.1509
		W(4)	0.0580	0.0568
			0.0957	0.0750
		MLE	0.0009	−0.0004
			0.0969	0.0632
0.2	0.4	W(2)	0.0494	0.0468
			0.0620	0.0528
		W(3)	0.0215	0.0209
			0.0605	0.0405
		W(4)	0.0064	0.0123
			0.0821	0.0552
		MLE	0.0211	0.0096
			0.0595	0.0359
0.4	0.4	W(2)	0.1350	0.1315
			0.0585	0.0564
		W(3)	0.0217	0.0207
			0.0795	0.0662
		W(4)	0.0217	0.0207
			0.0817	0.0541
		MLE	0.0215	0.0105
			0.0752	0.0459
0.6	0.4	W(2)	0.2732	0.2713
			0.2756	0.2726
		W(3)	0.1425	0.1372
			0.1513	0.1419
		W(4)	0.0553	0.0562
			0.0973	0.0759
		MLE	0.0304	0.0147
			0.0976	0.0672

Notes: The data generation process is $(1-\phi L)(1-L)^d Y_t = \epsilon_t$. See Notes to Table 1 for more details.

Table 3: Monte Carlo Results of the Wavelet Approximate MLE and the Frequency-Domain MLE for ARFIMA(0, d, 1) Model

θ	d	Estimation Method	Sample Size 1024	2048
0.2	0.2	W(2)	−0.0738	−0.0673
			0.0809	0.0716
		W(3)	−0.0288	−0.0241
			0.0568	0.0410
		W(4)	−0.0169	−0.0064
			0.0754	0.0494
		MLE	0.0040	0.0060
			0.0511	0.0337
0.4	0.2	W(2)	−0.1664	−0.1595
			0.1703	0.1617
		W(3)	−0.0809	−0.0773
			0.0940	0.0847
		W(4)	−0.0386	−0.0263
			0.0824	0.0547
		MLE	0.0123	0.0057
			0.0738	0.0462
0.6	0.2	W(2)	−0.2952	−0.2822
			0.2979	0.2838
		W(3)	−0.1817	−0.1722
			0.1887	0.1759
		W(4)	−0.0835	−0.0776
			0.1094	0.0923
		MLE	0.0228	0.0062
			0.1021	0.0651
0.2	0.4	W(2)	-0.0704	-0.0674
			0.0791	0.0724
		W(3)	-0.0237	-0.0216
			0.0562	0.0420
		W(4)	-0.0075	-0.0024
			0.0785	0.0539
		MLE	0.0275	0.0110
			0.0637	0.0390
0.4	0.4	W(2)	−0.1617	−0.1537
			0.1666	0.1563
		W(3)	−0.0689	−0.0640
			0.0869	0.0741
		W(4)	−0.0222	−0.0187
			0.0793	0.0543
		MLE	0.0382	0.0181
			0.0957	0.0513
0.6	0.4	W(2)	−0.2915	−0.2774
			0.2942	0.2792
		W(3)	−0.1683	−0.1565
			0.1761	0.1612
		W(4)	−0.0758	−0.0619
			0.1070	0.0827
		MLE	0.0350	0.0237
			0.1055	0.0759

Notes: The data generation process is $(1-L)^d Y_t = (1-\theta L)\epsilon_t$. See notes to Table 1 for more details.