HITOTSUBASHI UNIVERSITY

Global COE Hi-Stat Discussion Paper Series 210

Research Unit for Statistical
and Empirical Analysis in Social Sciences (Hi-Stat)

The Wavelet-based Estimation for Long Memory Signal Plus Noise Models

Kei Nanamiya

December 2011

Hi-Stat Discussion Paper

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

The wavelet-based estimation for long memory signal plus noise models

Kei Nanamiya *

Graduate school of Economics, Hitotsubashi University

November 30, 2011

Abstract

We propose new wavelet-based procedure to estimate the memory parameter of an unobserved process from an observed process affected by noise in order to improve the performance of the estimator by taking into account the dependency of the wavelet coefficients of long memory processes. In our procedure, using the AR(1) approximation for the wavelet transformed long memory processes which is introduced by Craigmile, Guttorp, and Percival (2005), we apply the ARMA(1,1) approximation to the wavelet coefficients of the observed process at each level. We also compare this procedure to the usual wavelet-based procedure by numerical simulations.

key words: wavelet; long memory process; measurement error problem.

1 Introduction

We consider the estimation problem of the memory parameter of signals in noisy environments via wavelets by numerical studies. This is one of the measurement error problems as follows; when we have the observable process $\{Y_t\}$ which is

$$Y_t = X_t + u_t,$$

*e-mail: ed061005@g.hit-u.ac.jp

where $\{X_t\}$ is an unobservable long memory process, $\{u_t\}$ is also an unobservable white noise process, and $\{X_t\}$ and $\{u_t\}$ are assumed mutually independent. Then our purpose is estimating the memory parameter of the unobservable process $\{X_t\}$. This model has also been considered as the long memory stochastic volatility (LMSV) model in financial econometrics (Breidt, Crato, and de Lima (1998), Deo and Hurvich (2003) and Teyssière and Abry (2007)).

In the wavelet domain, Wornell and Oppenheim (1992), Kaplan and Kuo (1993) and Zhang, Bao, and Wu (2004) use EM algorithm to recover long memory signals, $\{X_t\}$, from noise, and Tanaka (2004) tries to estimate the memory parameter by the white-noise (WN) approximation wavelet-based maximum likelihood (ML) estimation with Haar wavelet, and compares the performance of this wavelet-based method to that of other methods in the time domain and the frequency domain by numerical simulations.

To improve the estimation of the memory parameter, d, in this measurement error problem, we propose new wavelet-based estimation procedure by taking into account the point that the wavelet coefficients of long memory processes have some correlations. Using the AR(1) approximation introduced by Craigmile, Guttorp, and Percival (2005) for the wavelet transformed long memory process at each level, we approximate the wavelet coefficients of the unobservable long memory process $\{X_t\}$ at each level as AR(1), and then we have the ARMA(1,1) representation for the wavelet transformed observed process $\{Y_t\}$ at each level. We utilize this ARMA(1,1) representation to improve the estimator for the memory parameter of the unobservable process $\{X_t\}$, and compare the performance of this new procedure with that of the usual WN approximation procedure in numerical simulations.

This article is organized as follows; in section 2, the discrete wavelet transform (DWT) and the AR(1) approximation method (Craigmile, Guttorp, and Percival (2005)) are briefly reviewed; in section 3, we explain the ARMA(1,1) approximation method which is an extension of this AR(1) approximation method to the signal plus noise problem; in section 4, we have numerical simulations and compare the ARMA(1, 1) approximation method with the white-noise approximation method; conclusions are in section 5. All numerical simulations are done by Ox (Doornik (2007)), and all figures are drawn by R.

2 Preparations

In this section, we review the DWT and the AR(1) approximation wavelet-based likelihood function procedure for long memory processes in Craigmile, Guttorp, and Percival (2005), because we use this procedure to construct the ARMA(1,1) approximation wavelet-based likelihood function procedure for long memory processes plus noise model at the next section.

2.1 The discrete wavelet transform

The wavelet method we use in this article is called "the first generation wavelet[1]," which is introduced by Daubechies (1992), and we use mainly the DWT. The wavelet analysis can be defined in terms of an orthogonal analysis to represent a function by the weighted sum of wavelet and scaling functions. These weights of the summation are called wavelet and scaling coefficients. Using these coefficients, this method can decompose time series into subsequences which are the motion of the original time series in time and scales.

Following the definition of the DWT in Percival and Walden (2000), let $\{h_l\}_{l=0}^{L-1}$ denote the wavelet filter coefficients of a Daubechies compactly supported wavelet with the length L and let $\{g_l\}_{l=0}^{L-1}$ be the corresponding scaling filter coefficients which are defined by

$$g_l \equiv (-1)^{l+1} h_{L-1-l}.$$

Let's set $h_{1,l} = h_l$, $g_{1,l} = g_l$ and $L_1 = L$, then we can obtain the j-th level wavelet and scaling filter coefficients, $\{h_{j,l}\}_{l=0}^{L_j-1}$ and $\{g_{j,l}\}_{l=0}^{L_j-1}$, for $j \geq 2$, by

$$h_{j,l} = \sum_{k=0}^{L-1} h_{1,k} g_{j-1,l-2^{j-1}k} \quad \text{and} \quad g_{j,l} = \sum_{k=0}^{L-1} g_{1,k} g_{j-1,l-2^{j-1}k}, \tag{1}$$

[1]"The second generation wavelet" is based on the lifting schemes introduced by Sweldens (1996). This method is mainly used for irregularly spaced data (see Cohen (2003), Jensen and Oonicx (2005) and Mallat (2009))

where $L_j = (2^j - 1)(L - 1) + 1$. These filters have the following properties;

$$\sum_{l=0}^{L_j-1} h_{j,l} = 0, \qquad \sum_{l=0}^{L_j-1} h_{j,l}^2 = 1,$$
$$\sum_{l=0}^{L_j-1} g_{j,l} = 2^{j/2}, \qquad \sum_{l=0}^{L_j-1} g_{j,l}^2 = 1,$$

and

$$\sum_{l=0}^{L_j-1} g_{j,l} h_{j,l} = 0.$$

The (j,k)-th wavelet and scaling coefficients of the series $\{a_t\}_{t=0}^{T-1}$ are computed as

$$W_{a,j,k} = \sum_{l=0}^{L_j-1} h_{j,l} a_{2^j(k+1)-l-1 \bmod N}, \tag{2}$$

and

$$V_{a,j,k} = \sum_{l=0}^{L_j-1} g_{j,l} a_{2^j(k+1)-l-1 \bmod N}. \tag{3}$$

where $L_1 = L$ and $L_j = (2^j - 1)(L - 1) + 1$.

When we apply the DWT to time series analysis, one of the important points is "non-boundary coefficients." Non-boundary coefficients are the ones that are not affected by periodic filter operation, i.e., mod operation in eq.(2) and eq.(3). For example, let $\{X_t\}_{t\in\mathbb{Z}}$ and $\{\tilde{X}_t\}_{t=0}^{N-1}$ be a true sequence and its sub-process of $\{X_t\}_{t\in\mathbb{Z}}$, i.e., $\tilde{X}_t = X_t$ for $t = 0,\ 1,\ 2,\ \ldots,\ N-1$. Then their wavelet and scaling coefficients are defined; for $j = 1, 2, \ldots,$ and $k \in \mathbb{Z}$,

$$W_{X,j,k} = \sum_{l=0}^{L_j-1} h_{j,l} X_{2^j(k+1)-l-1}, \tag{4}$$
$$V_{X,j,k} = \sum_{l=0}^{L_j-1} g_{j,l} X_{2^j(k+1)-l-1}, \tag{5}$$

and for $j = 0, 1, \ldots, \log_2 N$, and $k = 0, 1, 2, \ldots, N_j - 1$,

$$W_{\tilde{X},j,k} = \sum_{l=0}^{L_j-1} h_{j,l} \tilde{X}_{2^j(k+1)-l-1 \bmod N},$$
$$V_{\tilde{X},j,k} = \sum_{l=0}^{L_j-1} g_{j,l} \tilde{X}_{2^j(k+1)-l-1 \bmod N}.$$

Because $\tilde{X}_t = X_t$ for $t = 0,\ 1,\ 2,\ \ldots,\ N-1$, $W_{X,j,k} = W_{\tilde{X},j,k}$ and $V_{X,j,k} = V_{\tilde{X},j,k}$ are non-boundary coefficients for $j = \{j : N_j - L'_j > 0\}$ and $k = L'_j, \ldots, N_j - 1$ with $L'_j \equiv \lceil (L-2)(1-2^{-j}) \rceil$.

2.2 The AR(1) approximation

Now let observable process $\{X_t\}$ be the fractional integrated process, I(d), with $d \in \mathbb{R}$;

$$(1-B)^d X_t = \varepsilon_t, \tag{6}$$

where we assume $\varepsilon_t \overset{i.i.d.}{\sim} N(0, \sigma_\varepsilon^2)$. For the (j,k)-th non-boundary wavelet coefficient, we set

$$W_{X,j,k} = \phi_j(d) W_{X,j,k-1} + (Z_{n,b})_{j,k}, \tag{7}$$

where

$$(Z_{n,b})_{j,k} \overset{i.i.d.}{\sim} N(0, \eta_j(d)\sigma_\varepsilon^2), \tag{8}$$

with

$$\phi_j(d) = s_{X,j,1}(d)/s_{X,j,0}(d), \tag{9}$$

$$\eta_j(d) = s_{X,j,0}(d)(1-\phi_j^2(d)). \tag{10}$$

If $d < (L+1)/2$, $\{W_{X,j,k}\}_{k \in \mathcal{B}_|}$ is stationary process, so $s_{X,j,0}(d)$ and $s_{X,j,1}(d)$ are defined

$$s_{X,j,\tau}(d) = \mathrm{Cov}(W_{X,j,k}, W_{X,j,k+\tau})/\sigma_\varepsilon^2, \quad \tau = 0, 1, \tag{11}$$

and are calculated by

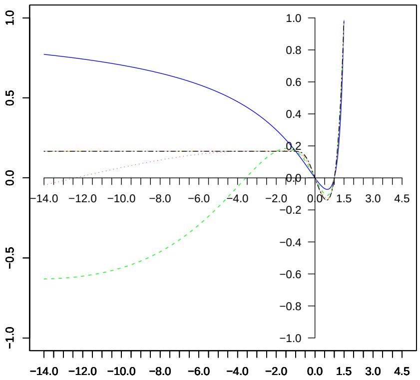
$$s_{X,j,\tau}(d) = \int_{-1/2}^{1/2} e^{i2\pi f\tau} S_{X,j}(f) df, \tag{12}$$

where

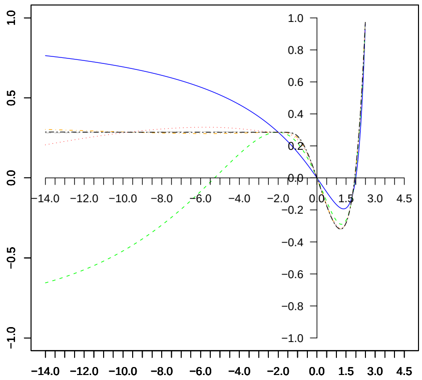
$$S_{X,j}(f) = 2^{-j} \sum_{k=0}^{2^j-1} \mathcal{H}_{j,L}(2^{-j}(f+k))(2\sin(\pi 2^{-j}(f+k)))^{-2d}. \tag{13}$$

Figure 1 displays the functions $\phi_j(d)$ for $j = 1, 2, \ldots, 6$ and $L = 2$, 4 and 8, where $\phi_j(d)$ for $j = 1, 2, \ldots, 6$ are drawn by solid, dashed, dotted, dot-dash, long-dash and two-dash lines. These lines are simultaneously zero when $d = 0$ and divergence when $d = (L+1)/2$. On the other hand, they seem different values at d except these points, and they are in $(-1, 1)$. When the length of filter is larger, the shaped width of them becomes larger.

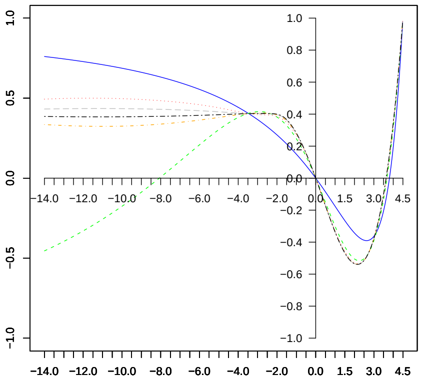
From eq.(7) and eq.(8), the wavelet-based likelihood function for long memory



(a) $L = 2$



(b) $L = 4$



(c) $L = 8$

Figure 1: $\phi_j(d)$, $D(L)$

process $\{X_t\}$ is

$$
\begin{aligned}
&L(d, \sigma_\varepsilon^2 | \{W_{X,j,k}\}_{(j,k)\in\mathcal{B}}) \qquad (14)\\
&= \prod_{j\in\mathcal{B}} \left(2\pi \frac{\eta_j(d)\sigma_\varepsilon^2}{1-\phi_j^2(d)}\right)^{-\frac{1}{2}} \exp\left\{-\frac{W_{x,j,L_j'}^2}{2\eta_j(d)\sigma_\varepsilon^2/(1-\phi_j^2(d))}\right\}\\
&\times(2\pi\eta_j(d)\sigma_\varepsilon^2)^{-\frac{M_j-1}{2}} \exp\left\{-\frac{\sum_{k=L_j'+1}^{N_j-1}(W_{X,j,k}-\phi_j(d)W_{X,j,k})^2}{2\eta_j(d)\sigma_\varepsilon^2}\right\}.
\end{aligned}
$$

3 The likelihood function for long memory process plus noise

We extend the AR(1) approximation procedure for wavelet transformed long memory process in the previous section to ARMA(1,1) representation for wavelet transformed long memory plus noise model. We observe only $\{Y_t\}_{t=0}^{N-1}$, which is

$$Y_t = X_t + u_t, \tag{15}$$

where

$$(1-B)^{d_0} X_t = \varepsilon_t, \tag{16}$$

$\varepsilon_t \overset{i.i.d.}{\sim} N(0, \sigma_\varepsilon^2)$, and

$$u_t \overset{i.i.d.}{\sim} N(0, \rho\sigma_\varepsilon^2), \tag{17}$$

with $\rho > 0$. We assume that $\{X_t\}$ and $\{u_t\}$ are independent.

Since the model is eq.(15), the (j,k)-th non-boundary wavelet coefficient is

$$W_{Y,j,k} = W_{X,j,k} + W_{u,j,k}, \tag{18}$$

where, from the remark in Chapter 4,

$$W_{u,j,k} \overset{i.i.d.}{\sim} N(0, \rho\sigma_\varepsilon^2). \tag{19}$$

Using the AR(1) approximation in the previous section for $\{W_{X,j,k}\}$,

$$W_{Y,j,k} = \phi_j(d) W_{Y,j,k-1} + (Z_{n,b})_{j,k} + W_{u,j,k} - \phi_j(d) W_{u,j,k-1}, \tag{20}$$

and then we have the ARMA(1,1) representation;

$$W_{Y,j,k} = \phi_j(d) W_{Y,j,k-1} + W_{v,j,k} + \beta_j(d,\rho) W_{v,j,k-1}, \tag{21}$$

where

$$W_{v,j,k} \overset{i.i.d.}{\sim} N(0, \alpha_j(d,\rho)\sigma_\varepsilon^2), \tag{22}$$

with

$$\beta_j(d,\rho) = \begin{cases} -\frac{\eta_j(d)+(1+\phi_j^2(d))\rho}{2\phi_j(d)\rho} + \sqrt{\left(\frac{\eta_j(d)\ +(1+\phi_j^2(d))\rho}{2\phi_j(d)\rho}\right)^2 - 1}, & \text{if} \quad 0 < \phi_j(d) < 1 \\ 0, & \text{if} \quad \phi_j(d) = 0 \\ -\frac{\eta_j(d)+(1+\phi_j^2(d))\rho}{2\phi_j(d)\rho} - \sqrt{\left(\frac{\eta_j(d)+(1+\phi_j^2(d))\rho}{2\phi_j(d)\rho}\right)^2 - 1}, & \text{if} \quad -1 < \phi_j(d) < 0, \end{cases} \tag{23}$$

and

$$\alpha_j(d,\rho) = \begin{cases} 1+\rho & \text{if} \quad \phi_j(d) = 0, \\ -\phi_j(d)\rho/\beta_j(d,\rho) & \text{if} \quad \phi_j(d) \neq 0. \end{cases} \tag{24}$$

Next we build up the likelihood function. Given $\{W_{y,j,k}\}_{(j,k)\in\mathcal{B}}$ with the set of the non-boundary wavelet coefficients, $\mathcal{B} = \{(j,k)|j = j_1,\ldots,j_2, k = L'_j,\ldots,N_j-1\}$, then the likelihood function is

$$\begin{aligned} &L(d,\rho,\sigma_\varepsilon^2|\{W_{y,j,k}\}_{(j,k)\in\mathcal{B}}) \\ &\quad = \prod_{j=j_1}^{j_2}(2\pi\alpha_j(d,\rho)\sigma_\varepsilon^2)^{-M_j/2}\prod_{k=L'_j}^{N_j-1} r_{j,k-1}^{-1/2}(d,\rho) \\ &\quad \times \exp\left\{-\frac{1}{2\alpha_j(d,\rho)\sigma_\varepsilon^2}\sum_{k=L'_j}^{N_j-1}\frac{(W_{y,j,k}-\hat{W}_{y,j,k}(d,\rho))^2}{r_{j,k-1}(d,\rho)}\right\}, \end{aligned} \tag{25}$$

where $M_j = N_j - L'_j$,

$$\begin{aligned} r_{j,L'_j-1}(d,\rho) &= (1+2\beta_j(d,\rho)\phi_j(d)+\beta_j^2(d,\rho))/(1-\phi_j^2(d)), && (26) \\ r_{j,k}(d,\rho) &= 1+\beta_j^2(d,\rho)-\beta_j^2(d,\rho)/r_{j,k-1}(d,\rho), \quad k \geq L'_j, && (27) \\ \beta_{j,k}(d,\rho) &= \beta_j(d,\rho)/r_{j,k-1}(d,\rho), && (28) \end{aligned}$$

and

$$\begin{aligned} \hat{W}_{y,j,L'_j}(d,\rho) &= 0, && (29) \\ \hat{W}_{y,j,k+1}(d,\rho) &= \phi_j(d)W_{y,j,k}+\beta_{j,k}(d,\rho)(W_{y,j,k}-\hat{W}_{y,j,k}(d,\rho)), \quad k \geq L'_j. && (30) \end{aligned}$$

From the log-likelihood function of eq.(25),

$$\widehat{\sigma}_\varepsilon^2(d,\rho|\{W_{y,j,k}\}_{(j,k)\in\mathcal{B}}) = \frac{1}{M}\sum_{j=j_1}^{j_2}\frac{1}{\alpha_j(d,\rho)}\sum_{k=L'_j}^{N_j-1}\frac{(W_{y,j,k}-\hat{W}_{y,j,k}(d,\rho))^2}{r_{j,k-1}(d,\rho)}, \tag{31}$$

where $M = \sum_{j=j_1}^{j_2} M_j$, and then its negative reduced log-likelihood function becomes

$$\begin{aligned} &\widetilde{\ell}_{ARMA}(d,\rho|\{W_{y,j,k}\}_{(j,k)\in\mathcal{B}}) \\ &\quad = M\log\left(\frac{1}{M}\sum_{j=j_1}^{j_2}\frac{1}{\alpha_j(d,\rho)}\sum_{k=L'_j}^{N_j-1}\frac{(W_{y,j,k}-\hat{W}_{y,j,k}(d,\rho))^2}{r_{j,k-1}(d,\rho)}\right) \\ &\quad + \sum_{j=j_1}^{j_2} M_j\log\alpha_j(d,\rho) + \sum_{j=j_1}^{j_2}\sum_{k=L'_j}^{N_j-1}\log r_{j,k-1}(d,\rho). \end{aligned} \tag{32}$$

Remark: When $d = 0$ and $s_{j,1}(d) = 0$, $\phi_j(d) = 0$, $\alpha_j(d, \rho) = 1+\rho$ and $\beta_j(d, \rho) = 0$. These make $\beta_{j,k}(d, \rho) = 0$ and $r_{j,k}(d, \rho) = 1$, and then $\hat{W}_{j,k}(d, \rho) = 0$. The log-likelihood function becomes

$$
\begin{aligned}
&\widetilde{\ell}_{ARMA}(\rho|d = 0, \{W_{y,j,k}\}_{(j,k)\in\mathcal{B}}) \\
&= M \log\left(\frac{1}{M(1+\rho)}\sum_{j=j_1}^{j_2}\sum_{k=L'_j}^{N_j-1} W_{y,j,k}^2\right) + M\log(1+\rho) \\
&= M \log\left(\frac{1}{M}\sum_{j=j_1}^{j_2}\sum_{k=L'_j}^{N_j-1} W_{y,j,k}^2\right).
\end{aligned}
$$

This equation is independent from ρ, so we can not identify ρ. In other words, because when $d = 0$ $I(d)$ signal becomes white noise and the observation $\{Y_t\}$ consists of two white noise, we can not divide the observation into the signal and noise processes.

4 Simulations

In these numerical simulations, we compare the two wavelet-based estimation procedures; one is based on the WN approximation for $\{X_t\}$, and the other is our ARMA(1,1) approximation. For convenience, we call them "WN-type" and "ARMA-type". The ARMA-type estimator is computed as in the previous section. On the other hand, when we calculate the WN-type estimator, the wavelet coefficients of the observation $\{Y_t\}$ is

$$W_{Y,j,k} = W_{X,j,k} + W_{u,j,k},$$

where, when $d < (L+1)/2$, $W_{X,j,k} \overset{i.i.d.}{\sim} N(0, \sigma_\varepsilon^2 s_{X,j,0}(d))$[2]. Then the likelihood function of WN-type is

$$L_{WN}(d, \rho, \sigma_\varepsilon^2) = \prod_{(j,k)\in\mathcal{B}} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2(s_{X,j,0}(d) + \rho)}} \exp\left\{-\frac{W_{Y,j,k}^2}{2\sigma_\varepsilon^2(s_{X,j,0}(d) + \rho)}\right\},$$

and the estimators of (d, ρ) are computed by

$$(\widehat{d}_{WN}, \widehat{\rho}_{WN}) = \arg\min \widetilde{\ell}_{WN}(d, \rho),$$

[2] Although Tanaka (2004) approximated the squared gain function of Haar filter in eq.(13) to the ideal band-pass filter in order to avoid the numerical integration and to evaluate the wavelet variance quickly, we evaluate the eq.(13) exactly.

where

$$\widetilde{\ell}_{WN}(d,\rho) = M\log\left(\frac{1}{M}\sum_{(j,k)\in\mathcal{B}}\frac{W_{Y,j,k}^2}{s_{X,j,0}(d)+\rho}\right) + \sum_{j\in\mathcal{J}} M_j \log\left(s_{X,j,0}(d)+\rho\right).$$

To compare these estimators, we consider the four sets of parameter combination as in Tanaka (2004);

Case 1: $(d,\rho) = (0.1, 0.1)$;

Case 2: $(d,\rho) = (0.45, 0.5)$;

Case 3: $(d,\rho) = (0.8, 10)$;

Case 4: $(d,\rho) = (1.8, 900000)$.

We set $\sigma_\varepsilon^2 = 1$ for all cases, and Haar, D(4) and D(8) wavelet filters are used. We also compare the two kinds of the set of scales, $\mathcal{J}_1 = \{j = 1, 2, \ldots, 6\}$ and $\mathcal{J}_2 = \{j = 2, \ldots, 6\}$. Sample sizes are $N = 512$, 1024, 2048 and 4096. Although we replicate 600 times for each cases, some replications are failed at numerical maximization, so we report about 500 succeeded replications. Numerical integrations for the maximization of likelihood functions are used to compute the wavelet autocovariance sequences.

Since we can not identify the signal and noise when $\widehat{d} = 0$ as noted in the previous section, we need to divide the parameter space of d into three spaces; $D_1 = (\epsilon, (L+1)/2)$, $D_2 = \{d = 0\}$, and $D_3 = (-\infty, -\epsilon)$, where ϵ is small positive constant. In the simulations, we set that $\epsilon = 10^{-5}$, the lower bound of D_3 is -100 instead of $-\infty$, and $\widehat{\Theta} = (\widehat{d}, \widehat{\rho}) = (0, 0)$ when $\widehat{d} \in D_2$. Then the algorithm of estimation becomes as follows;

Step 1 Compute $\widehat{\Theta} = (\widehat{d}, \widehat{\rho})$ and the negative reduced log likelihood function $\widetilde{\ell}(\widehat{\Theta})$ for D_1, D_2 and D_3, say $\widehat{\Theta}_i$ and $\widetilde{\ell}(\widehat{\Theta}_i)$ for $i = 1, 2, 3$.

Step 2 We elect $\widehat{\Theta}_i$ which takes the lowest value in $\widetilde{\ell}(\widehat{\Theta}_i)$ for $i = 1, 2, 3$ as the estimator $\widehat{\Theta}$ i.e.,

$$\widehat{\Theta} = \widehat{\Theta}_i = \arg\min_{\widehat{\Theta}_i, i=1,2,3}\{\widetilde{\ell}(\widehat{\Theta}_i)\}.$$

The results are drawn in Figures 2-45. This is because it is difficult to report them as tables. The eight small figures in the left side of each figure are the histograms of $\widehat{d}$ and those in the right side are plots of $(\widehat{d}, \widehat{\rho})$, where the horizonal line is $\widehat{d}$ and the vertical line is $\widehat{\rho}$. The upper group is the result of the WN-type, and the lower is that of the ARMA-type. The first, second, third and fourth row in each group are when $N = 512$, 1024, 2048 and 4096.

Figures 2-13 show the results of $(d, \rho) = (0.1, 0.1)$. In this case, because the long memory signal is stationary, we can calculate its signal-noise-ratio (SNR) and $SNR = 10.195$. Each histogram of $\widehat{d}$ has two peaks, which seems to be affected the $\widehat{\rho}$ simultaneously estimated. The std. of $\widehat{d}$ in the ARMA-type. is equally or smaller than that in the WN-type.

Figures 14-25 are the results of $(d, \rho) = (0.45, 0.5)$. We can also have SNR, and $SNR = 7.28486$. The shape of the histograms of the ARMA-type seems better than that of the WN-type, and there seems positive correlation between $\widehat{d}$ and $\widehat{\rho}$.

Figures 26-37 show the results of $(d, \rho) = (0.8, 10)$. Because the long memory signal in this case is nonstationary, we can not calculate SNR. Although there are few differences between two types, the range of $\widehat{d}$ in the ARMA-type. is a bit smaller than that in the WN-type.

Figures 38-45 are the results of $(d, \rho) = (1.8, 900000)$. This case is that the long memory signal is also nonstationary, but the noise is awfully heavy. Because the signal is covered by the heavy noise, we have many failures in estimations, for example, $\widehat{d}$ is estimated as zero or an incredibly negative value. The ARMA-type. seems better than the WN-type Each peak is larger than $d = 1.8$.

Throughout these four cases, $\mathcal{J}_1$ seems better than $\mathcal{J}_2$.

5 Conclusions

Using the AR(1) approximation for the wavelet coefficients of the unobservable long memory process at each scale, we have proposed the ARMA(1,1) approximation procedure to improve the estimator of the memory parameter in the long memory process plus noise model. Although the ARMA-type procedure spends more time to computing estimators, this procedure produced better results than the usual WN-type procedure used by Tanaka (2004) in numerical simulations. For further study, we need to explore the asymptotic theory for the both procedures.

ACKNOWLEDGMENT

The author thanks to Katsuto Tanaka, Toshiaki Watanabe and the members of their seminars. Financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through the Global COE program "Research Unit for Statistical and Empirical Analysis in Social Sciences" at Hitotsubashi University is gratefully acknowledged. This research is partially supported by the Global COE program "Research Unit for Statistical and Empirical Analysis in Social Sciences" at Hitotsubashi University.

References

Breidt, J., N. Crato, and P. de Lima (1998): "On the detection and estimation of long memory in stochastic volatility," *Journal of Econometrics*, 83, 325–348.

Cohen, A. (2003): *Numerical Analysis of Wavelet Methods*, vol. 32. North-Holland Publishing Co.

Craigmile, P. F., P. Guttorp, and D. B. Percival (2005): "Wavelet-Based Parameter Estimation for Polynomial Contaminated Fractionally Differenced Processes," *IEEE Transactions on Signal Processing*, 53, 3151–3161.

Daubechies, I. (1992): *Ten Lectures on Wavelets.* Philadelphia: Society for Industrial and Applied Mathematics.

Deo, R. S., and C. M. Hurvich (2003): "Estimation of long-memory in volatility," in *Long-Range Dependence: Theory and Applications*, ed. by P. Doukhan, G. Oppenheim, and M. S. Taqqu, pp. 313–324. Birkhauser, Boston.

Doornik, J. (2007): *Object -Oriented Matrix Programming Using Ox.* Timberlake Consultants Press and Oxford, London, 3rd edn.

Jensen, M., and P. Oonicx (2005): *Second Generation Wavelets and Applications.* Springer.

Kaplan, L. M., and C.-C. J. Kuo (1993): "Fractal estimation from noisy data via discrete fractional gaussian noise (DFGN) and the Haar basis," *IEEE Transactions on Signal Processing*, 41, 3554–3564.

Mallat, S. (2009): *a wavelet tour of signal processing The Sparse Way*. Academic Press, 3 edn.

Percival, D., and A. Walden (2000): *Wavelet Methods for Time Series Analysis*. Cambridge University Press.

Sweldens, W. (1996): "The lifting scheme: a custum-design construction of biorthogonal wavelets," *Applied and Computational Harmonic Analysis*, 3, 186–200.

Tanaka, K. (2004): "Frequency domain and wavelet-based estimation for long-memory signal plus noise models," in *State Space and Unobserved Component Models Theory and Applications*, ed. by A. Harvey, S. J. Koopman, and N. Shephard, pp. 75–91. Cambridge University Press, Cambridge.

Teyssière, G., and P. Abry (2007): "Wavelet Analysis of Nonlinear Long-Range Dependent Processes. Applications to Finincail Time Series," in *Long Memory in Economics*, ed. by G. Teyssière, and A. P. Kirman. Springer, Heidelberg.

Wornell, G. W., and A. V. Oppenheim (1992): "Estimation of fractinal signals from noisy measurements using wavelets," *IEEE Transactions on Signal Processing*, 40, 611–623.

Zhang, L., P. Bao, and X. Wu (2004): "Wavelet Estimation of Fractional Brownian Motion Embedded in a Noisy Environment," *IEEE Transactions on Information Theory*, 50, 2194–2200.

Figure 2: Case 1: Haar, WN-type, $j = 1, 2, \ldots, 6$

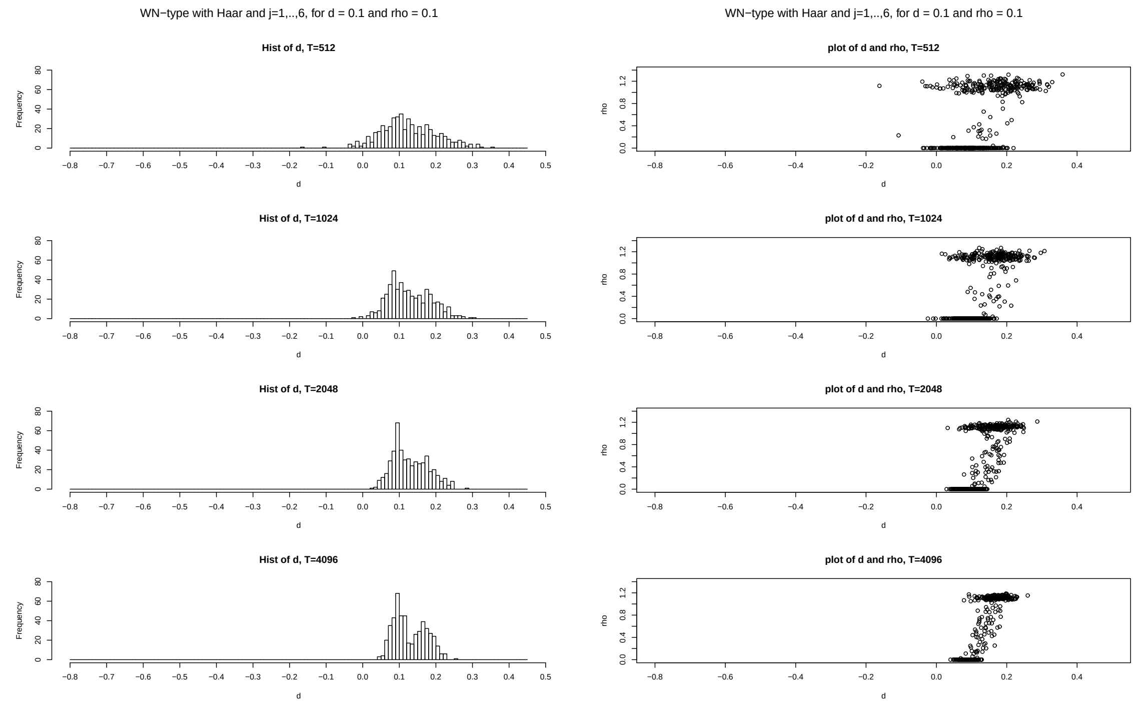


Figure 3: Case 1: Haar, ARMA-type, $j = 1, 2, \ldots, 6$

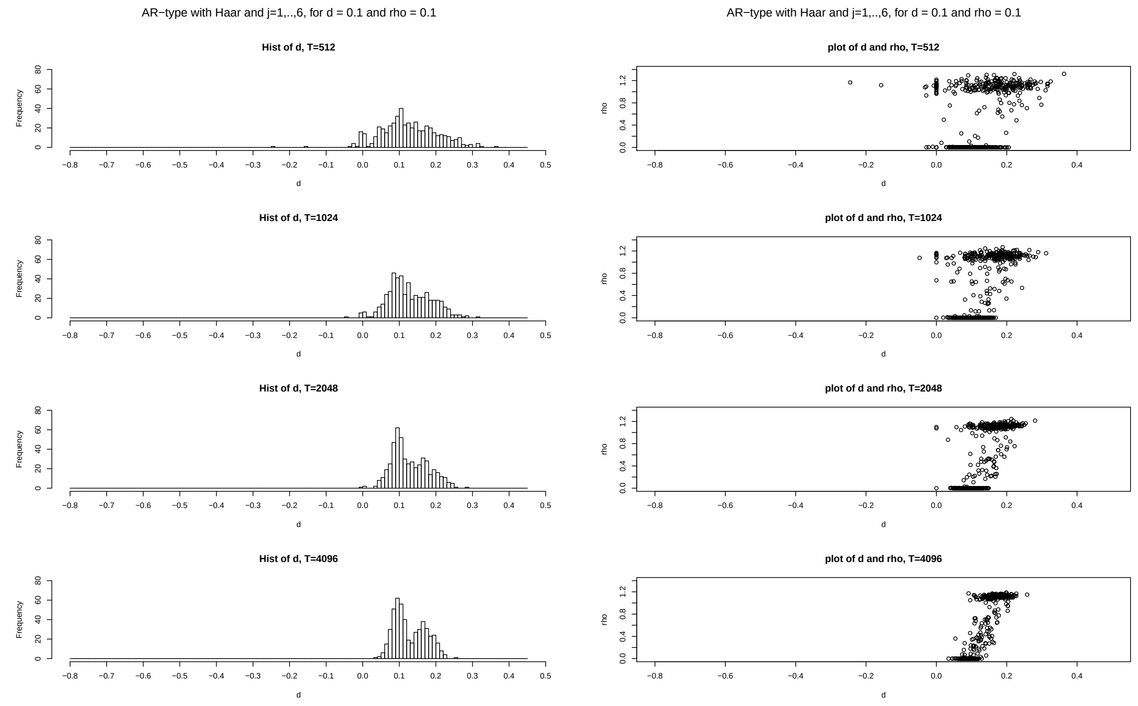


Figure 4: Case 1: Haar, WN-type, $j = 2, \ldots, 6$

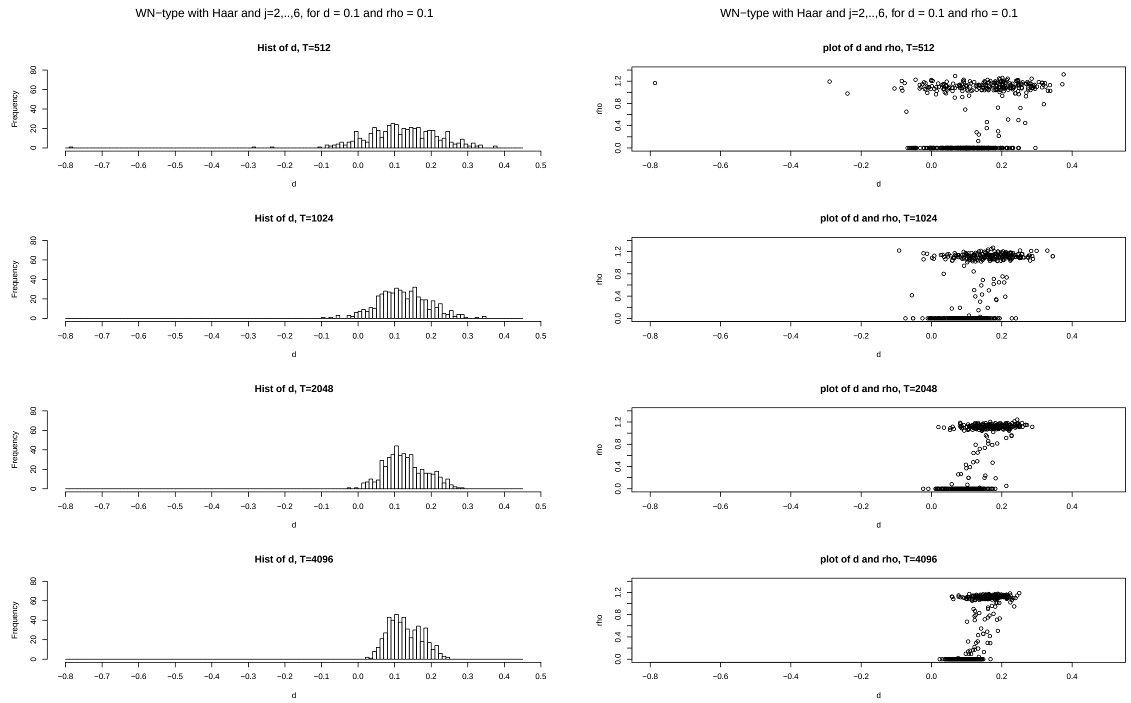


Figure 5: Case 1: Haar, ARMA-type, $j = 2, \ldots, 6$

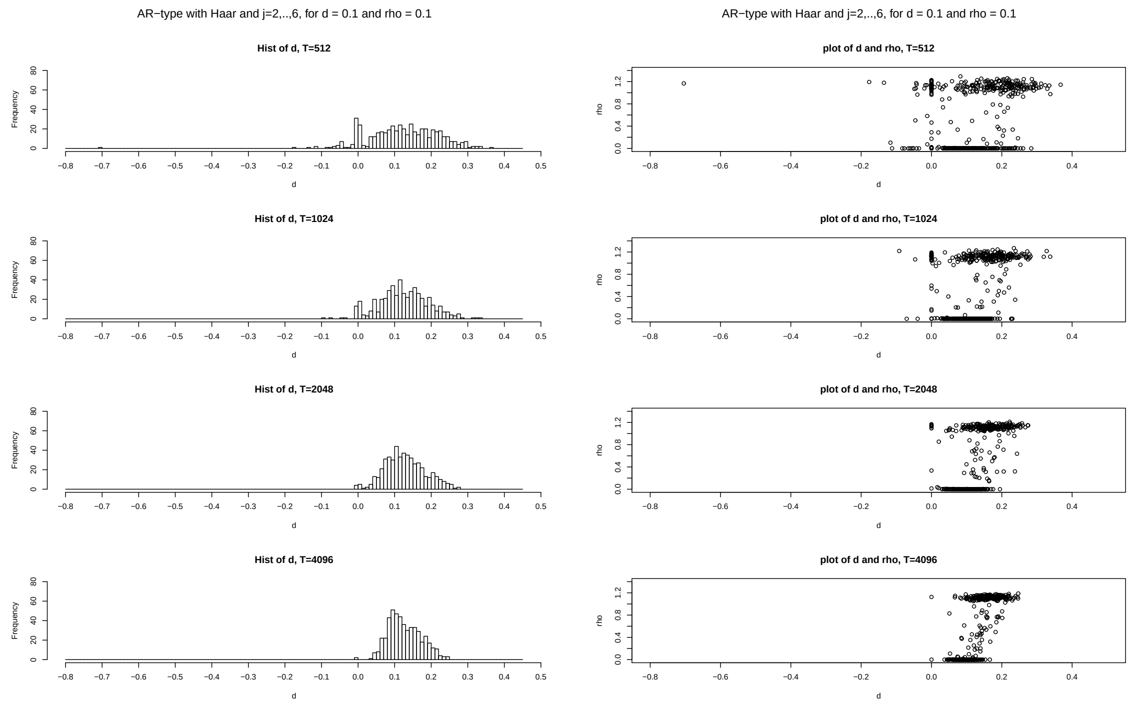


Figure 6: Case 1: D(4), WN-type, $j = 1, 2, \ldots, 6$

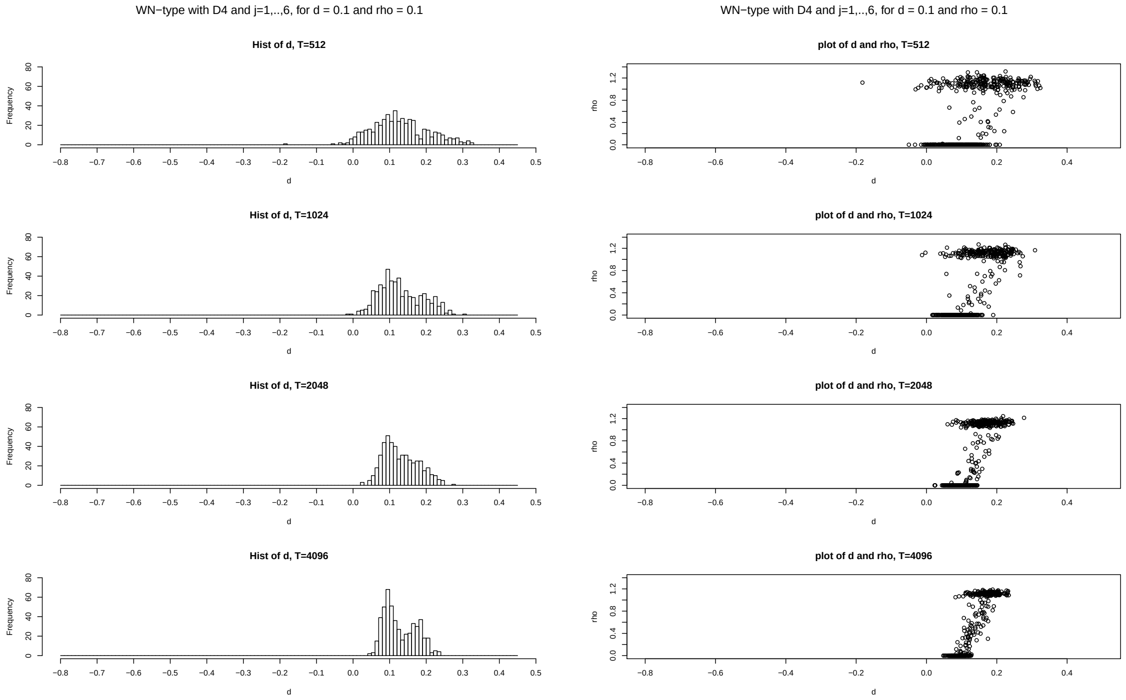


Figure 7: Case 1: D(4), ARMA-type, $j = 1, 2, \ldots, 6$

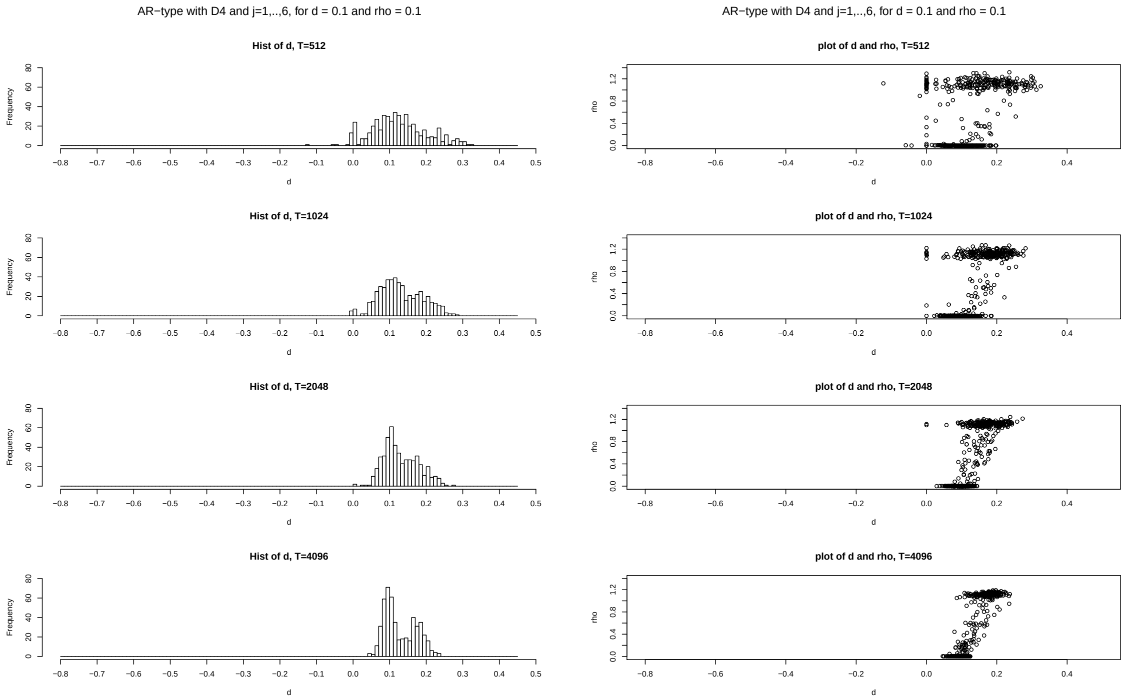


Figure 8: Case 1: D(4), WN-type, $j = 2, \ldots, 6$

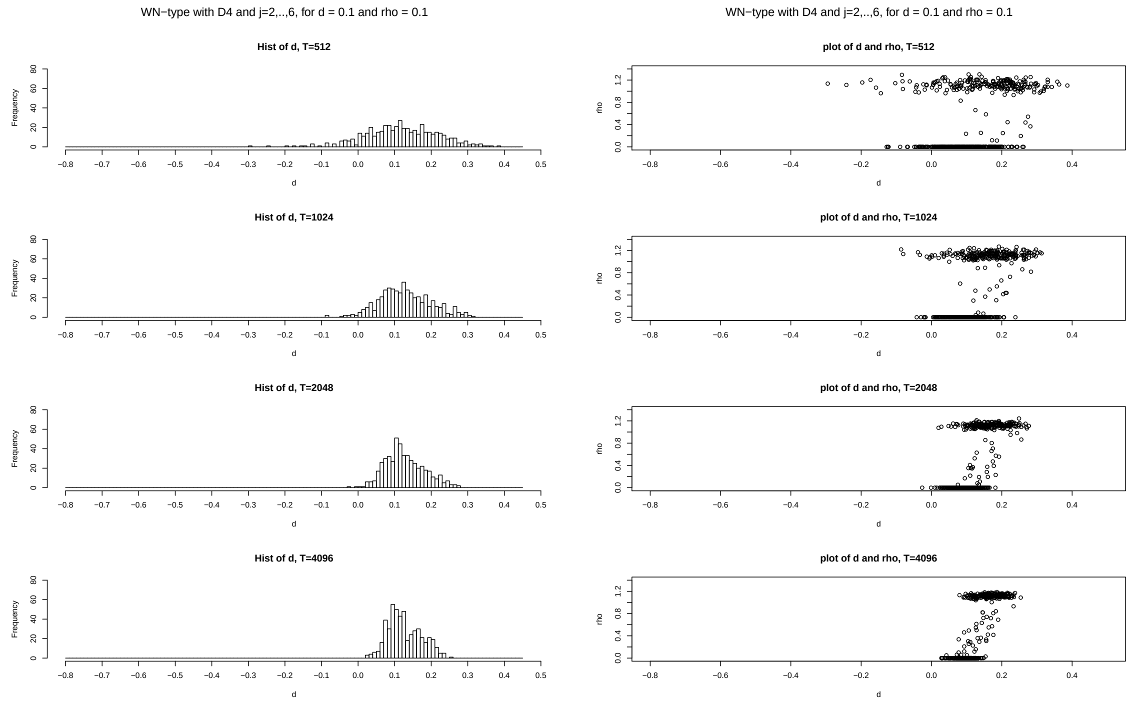


Figure 9: Case 1: D(4), ARMA-type, $j = 2, \ldots, 6$

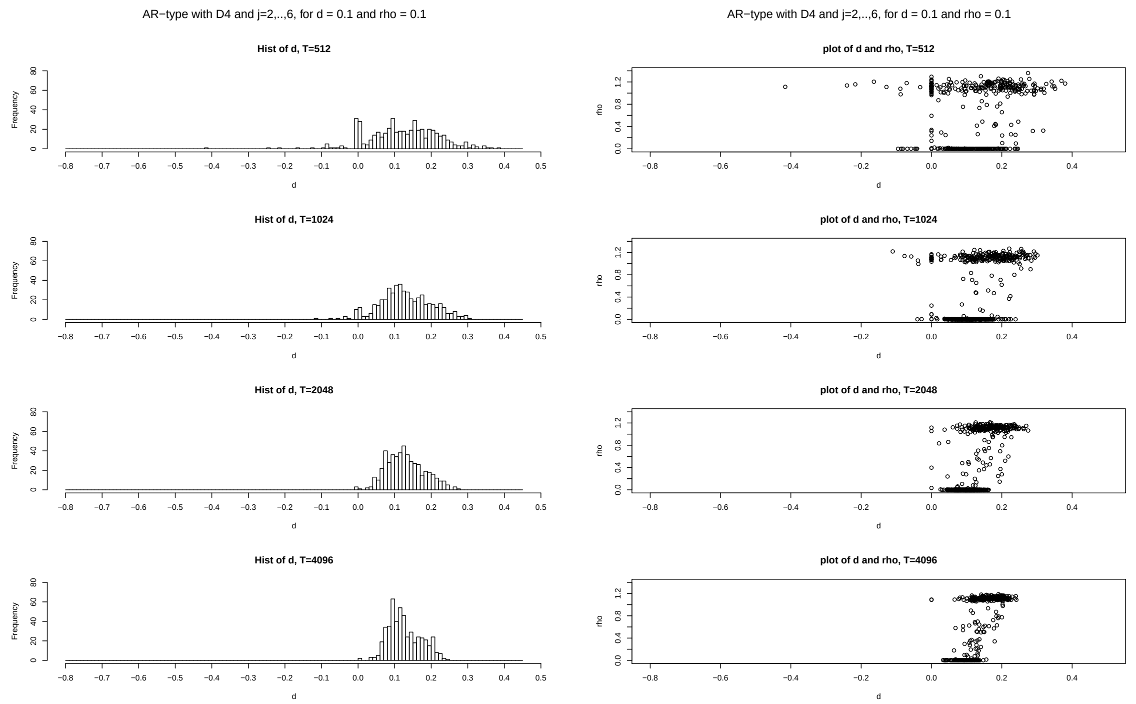


Figure 10: Case 1: D(8), WN-type, $j = 1, 2, \ldots, 6$

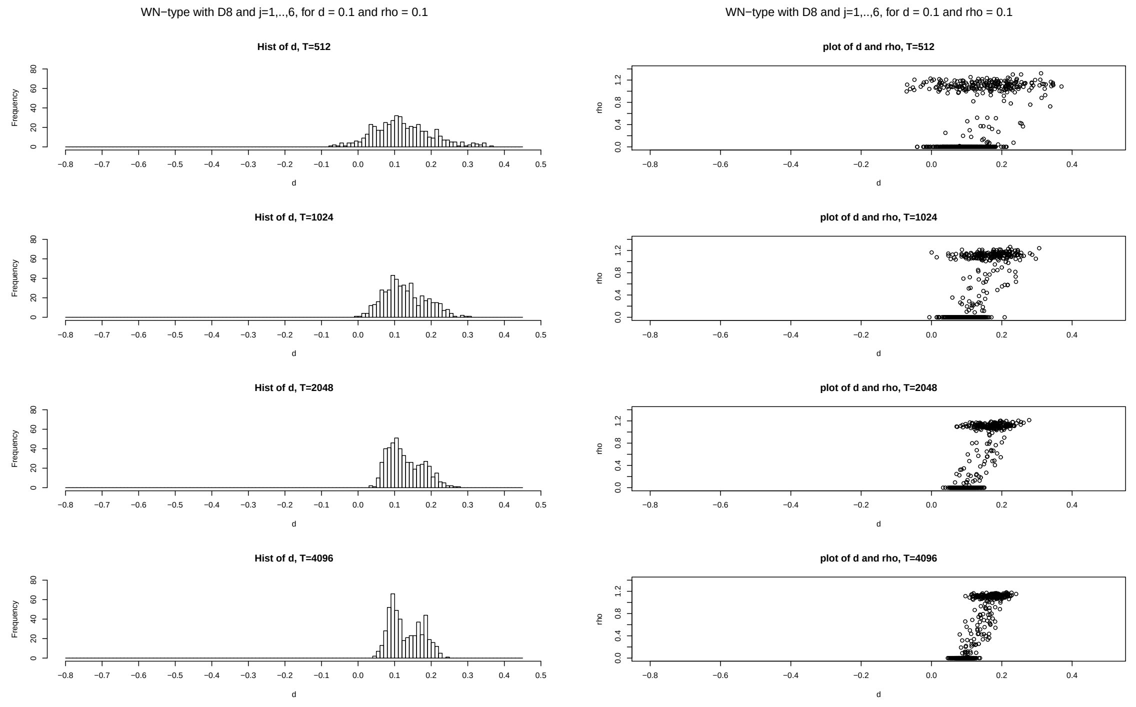


Figure 11: Case 1: D(8), ARMA-type, $j = 1, 2, \ldots, 6$

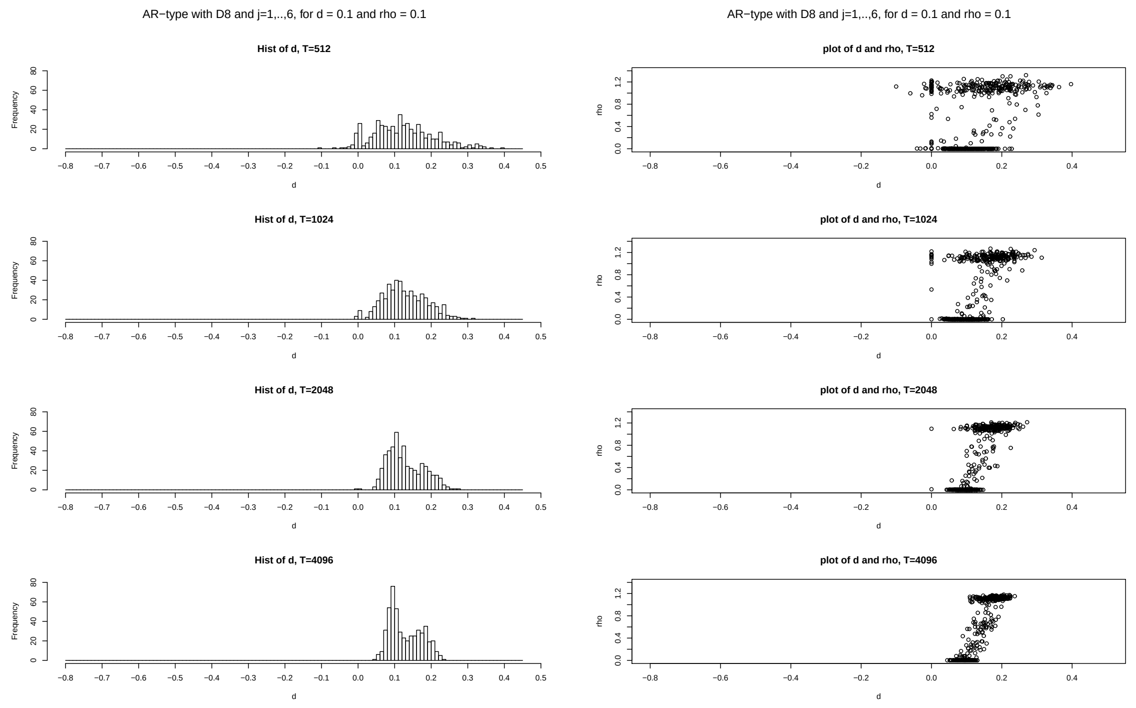


Figure 12: Case 1: D(8), WN-type, $j = 2, \ldots, 6$

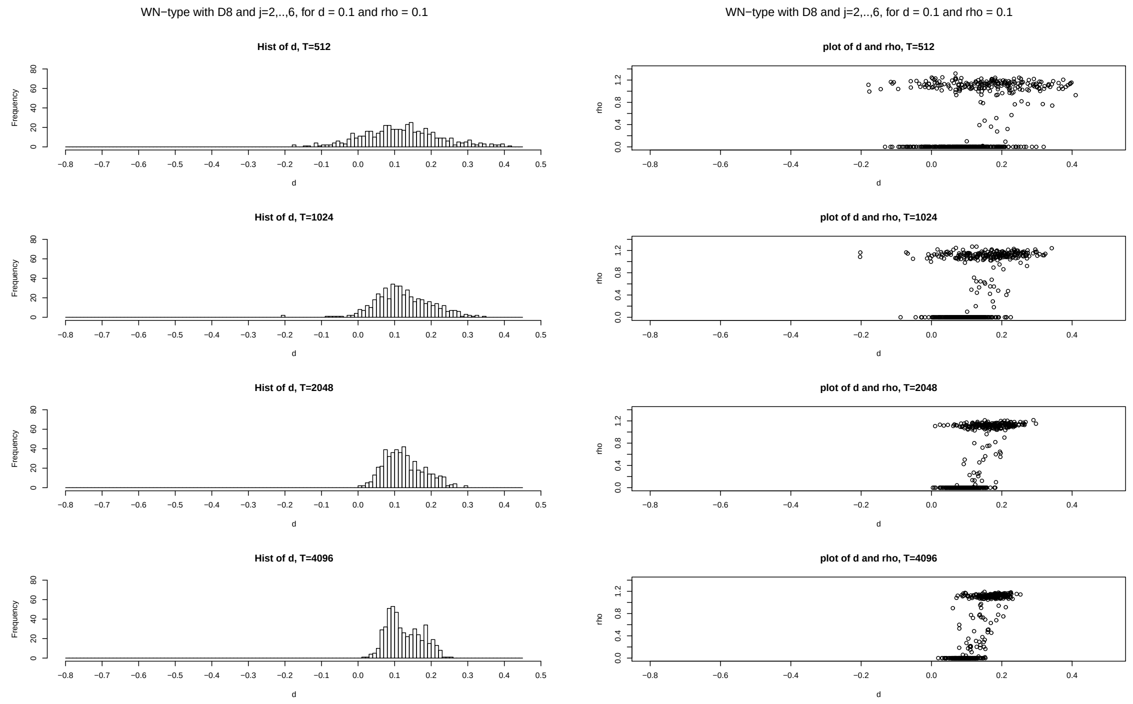


Figure 13: Case 1: D(8), ARMA-type, $j = 2, \ldots, 6$

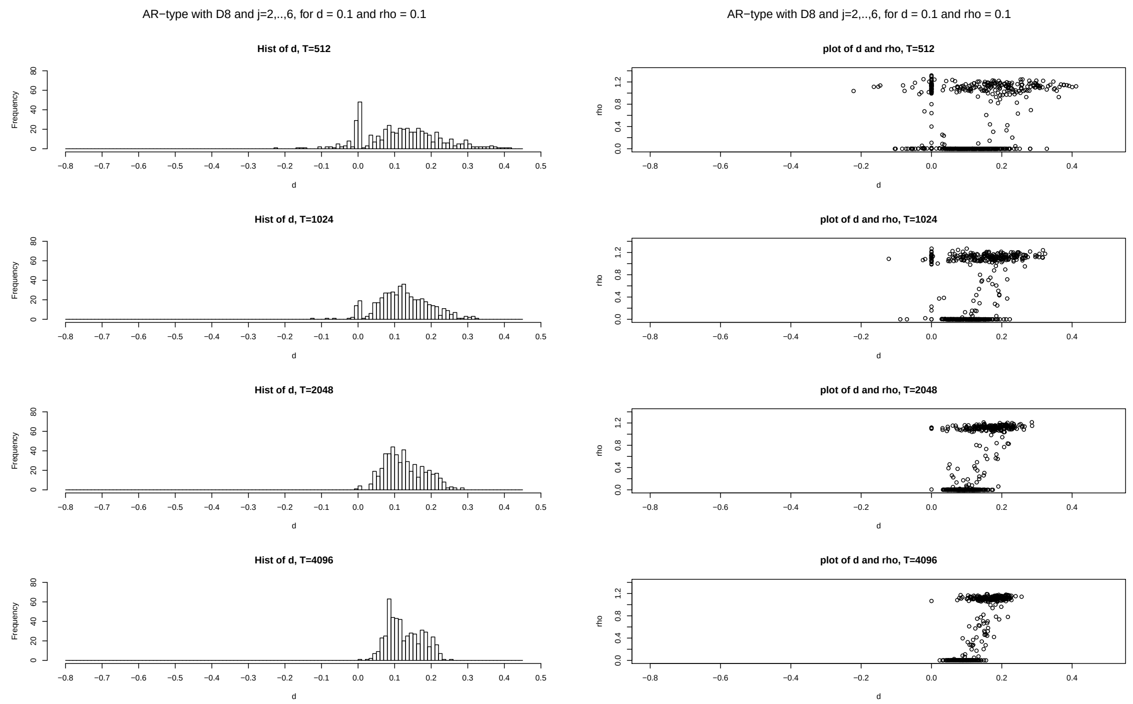


Figure 14: Case 2: Haar, WN-type, $j = 1, 2, \ldots, 6$

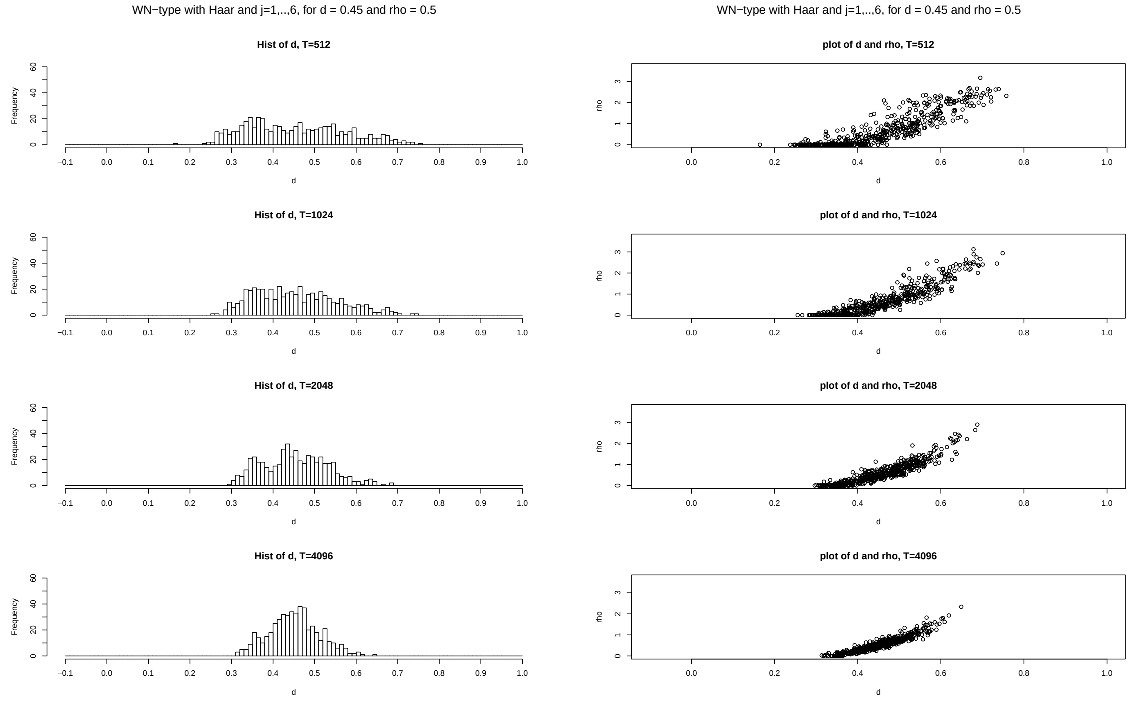


Figure 15: Case 2: Haar, ARMA-type, $j = 1, 2, \ldots, 6$

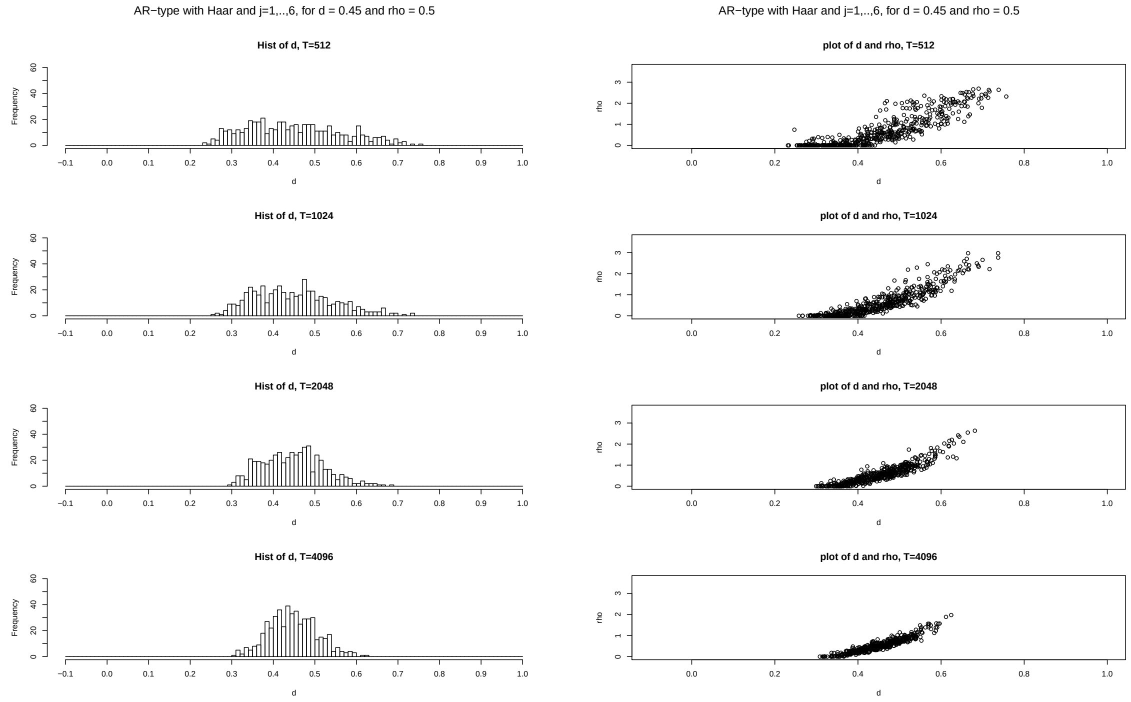


Figure 16: Case 2: Haar, WN-type, $j = 2, \ldots, 6$

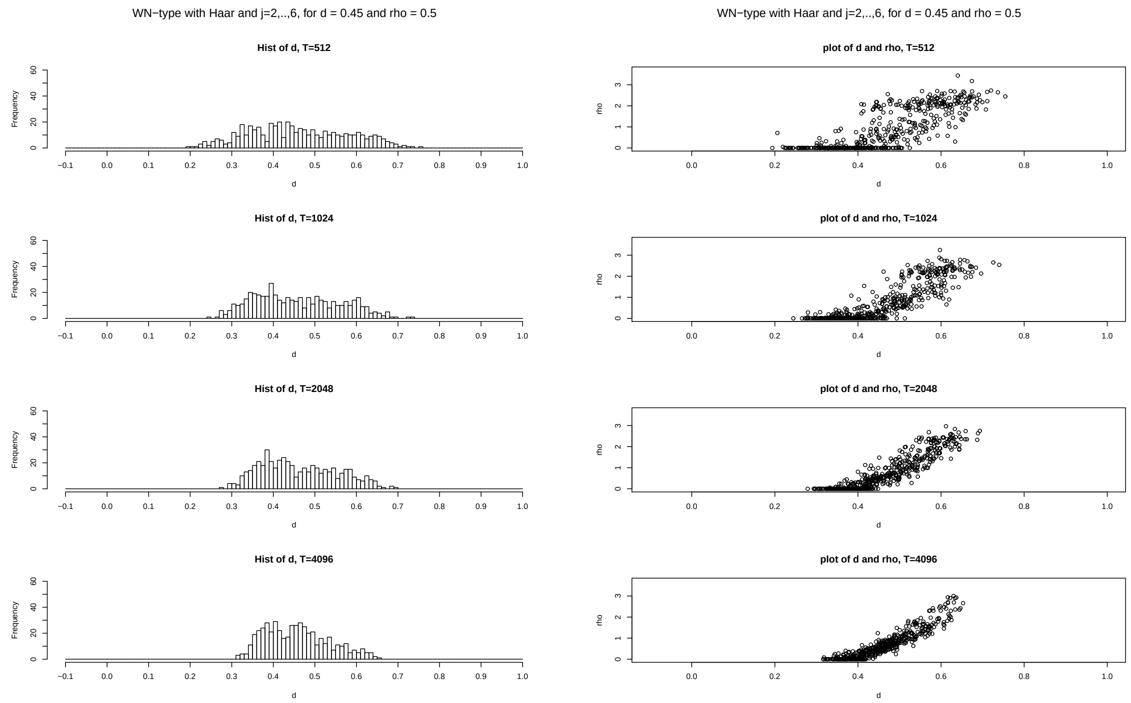


Figure 17: Case 2: Haar, ARMA-type, $j = 2, \ldots, 6$

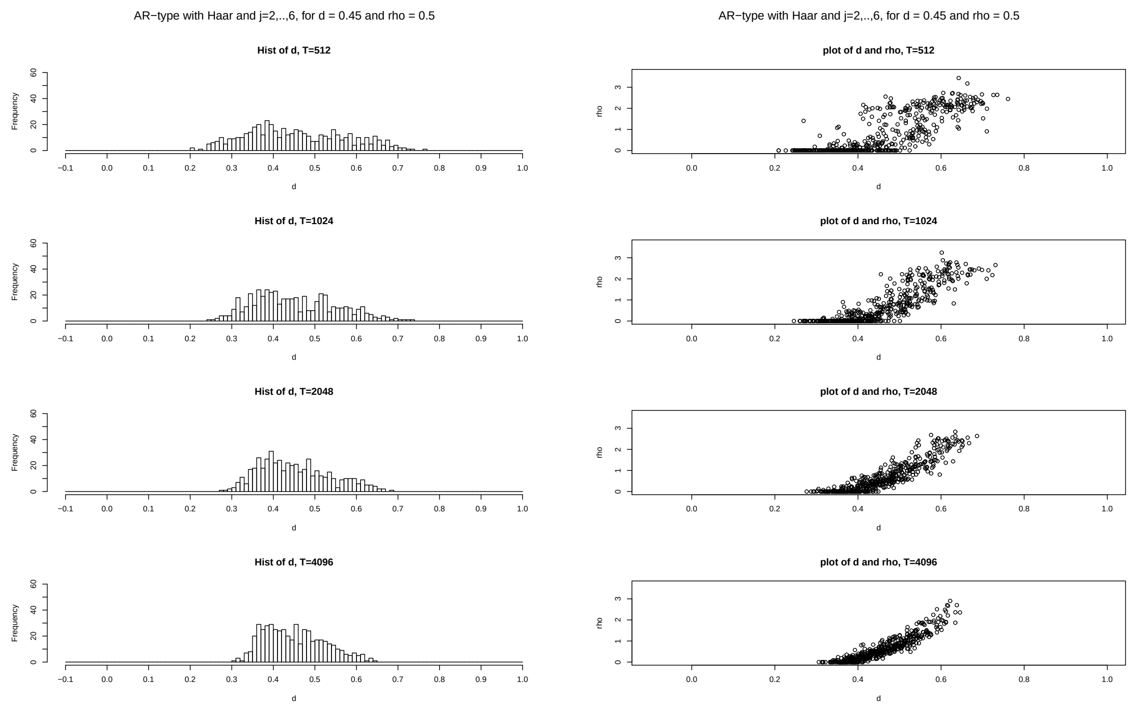


Figure 18: Case 2: D(4), WN-type, $j = 1, 2, \ldots, 6$

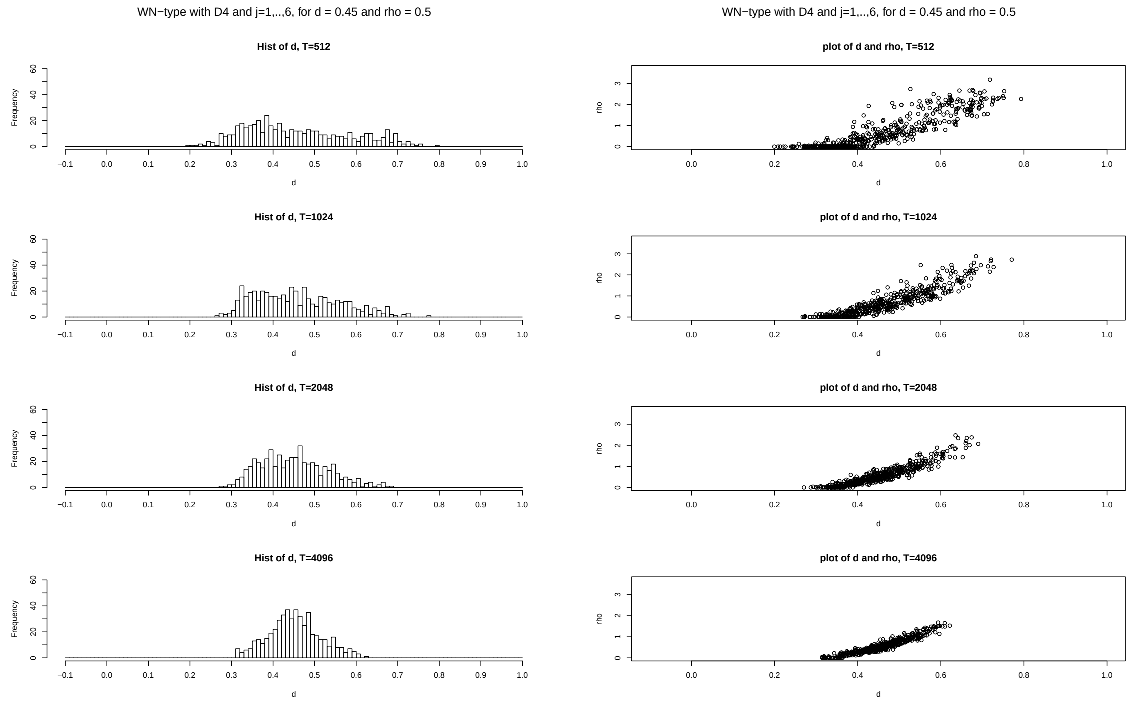


Figure 19: Case 2: D(4), ARMA-type, $j = 1, 2, \ldots, 6$

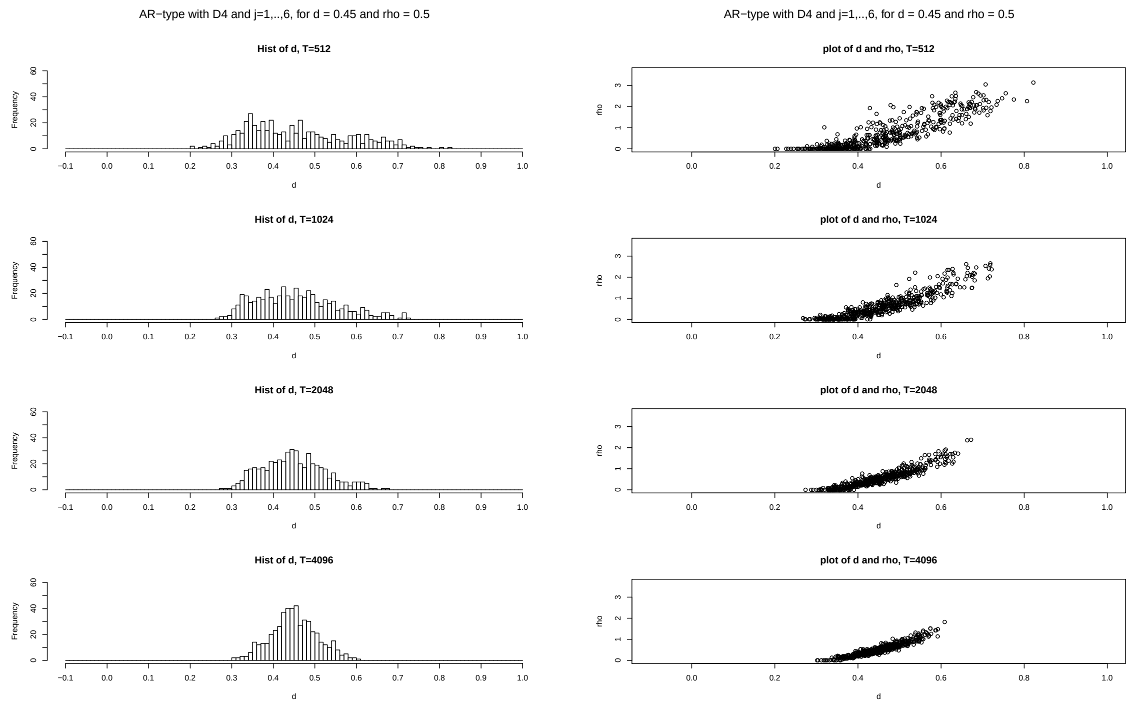


Figure 20: Case 2: D(4), WN-type, $j = 2, \dots, 6$

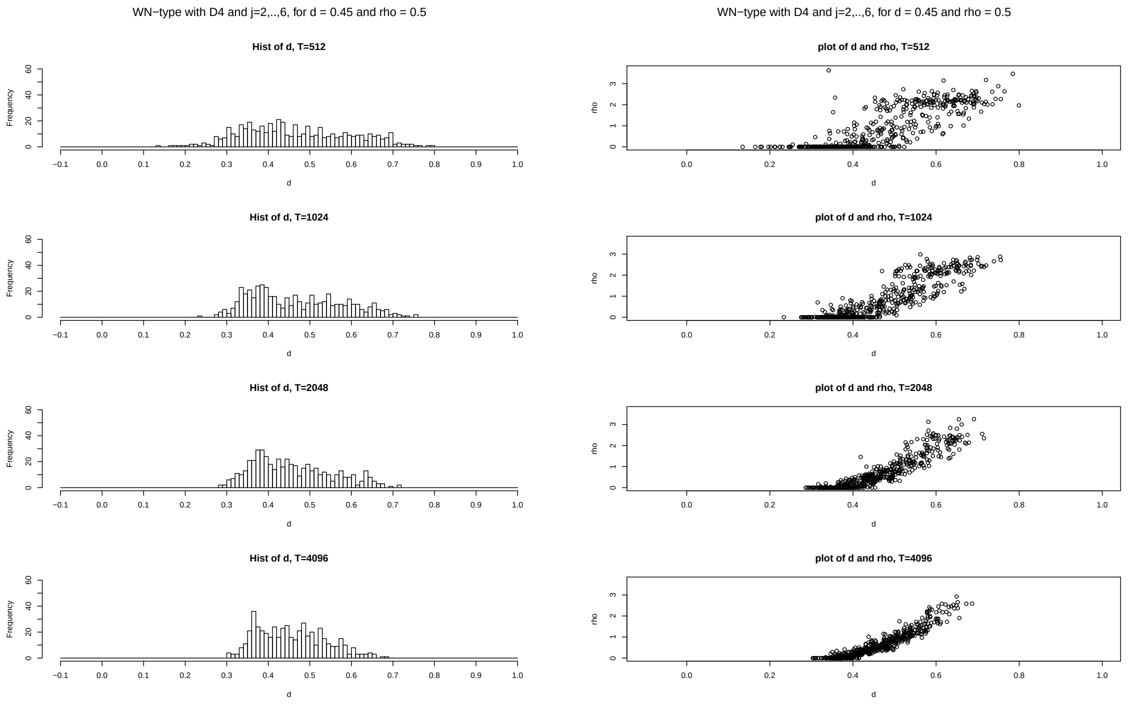


Figure 21: Case 2: D(4), ARMA-type, $j = 2, \dots, 6$

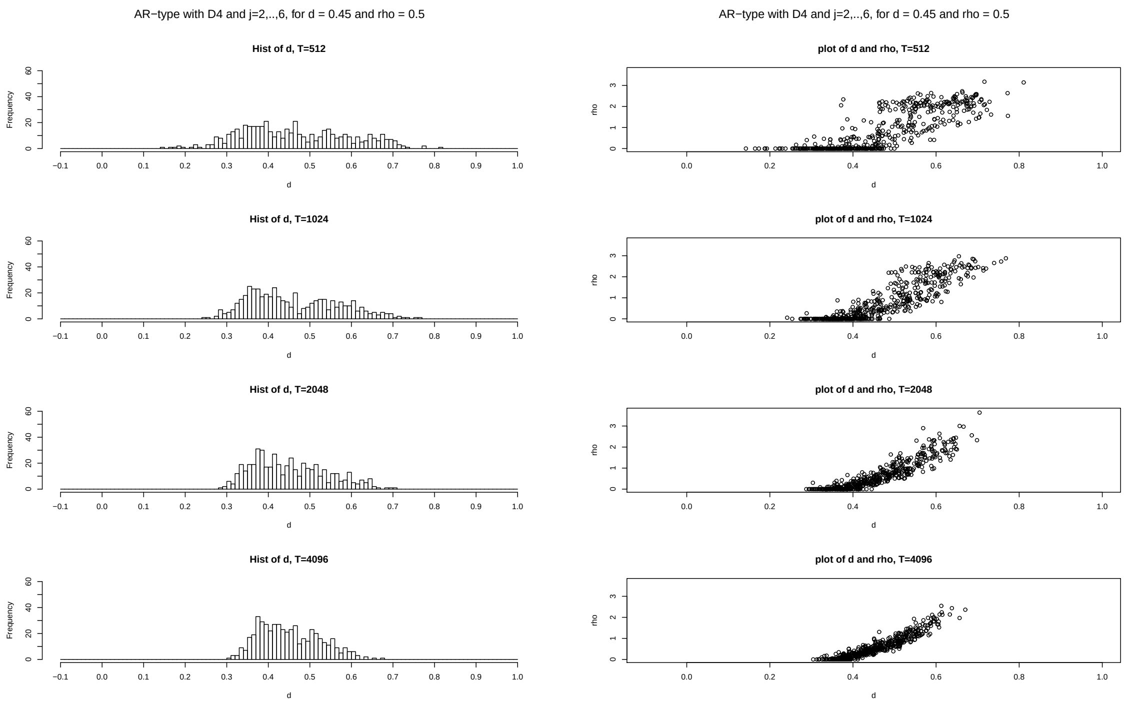


Figure 22: Case 2: D(8), WN-type, $j = 1, 2, \ldots, 6$

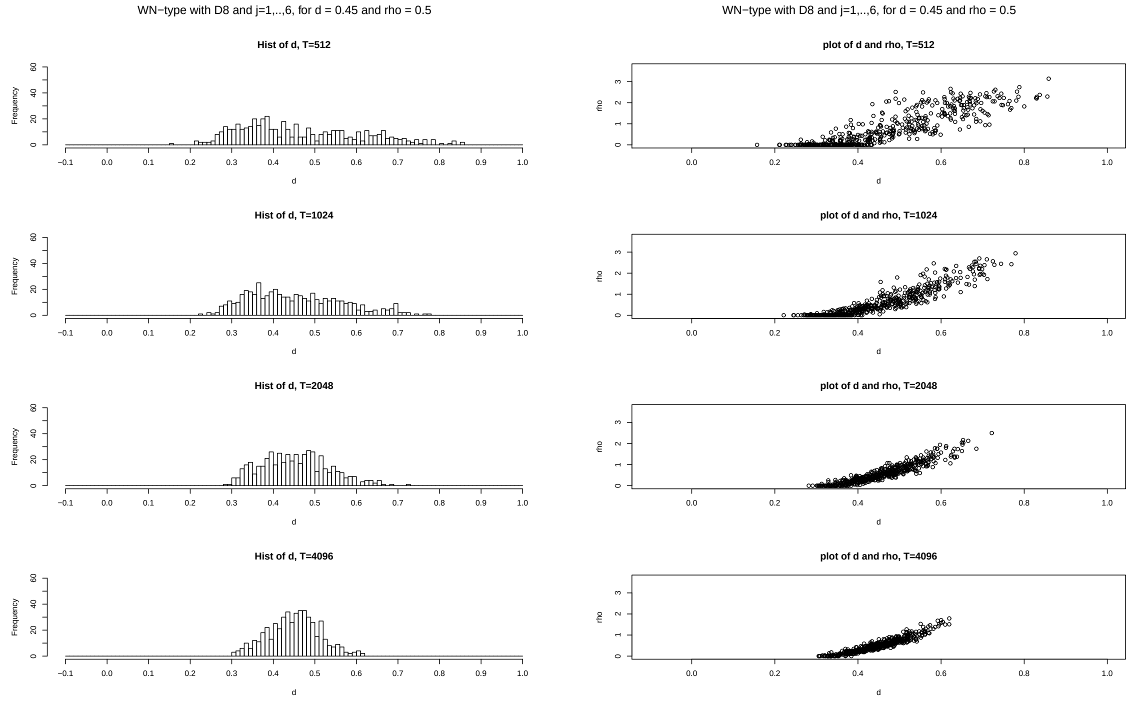


Figure 23: Case 2: D(8), ARMA-type, $j = 1, 2, \ldots, 6$

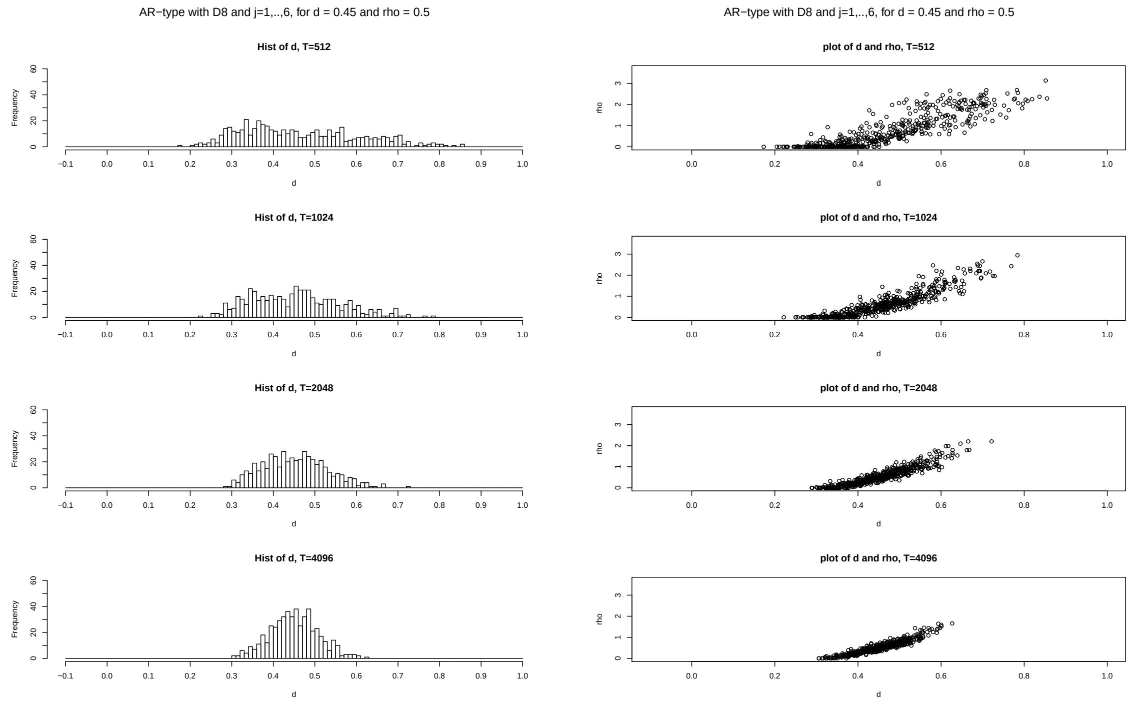


Figure 24: Case 2: D(8), WN-type, $j = 2, \ldots, 6$

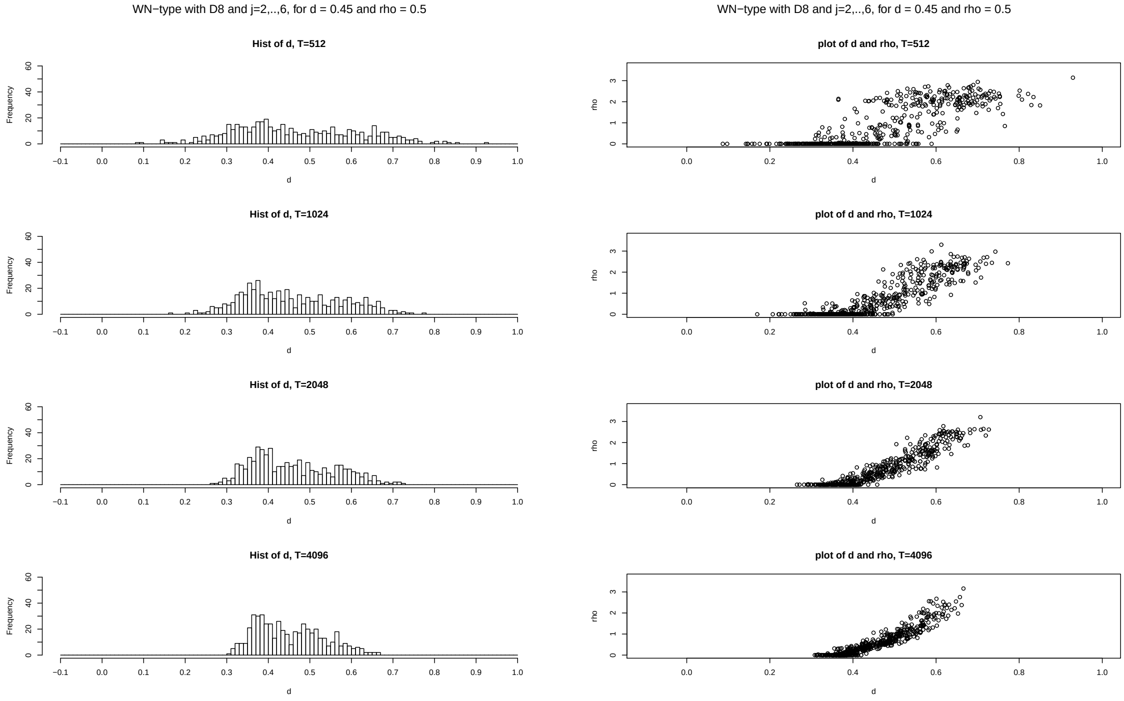


Figure 25: Case 2: D(8), ARMA-type, $j = 2, \ldots, 6$

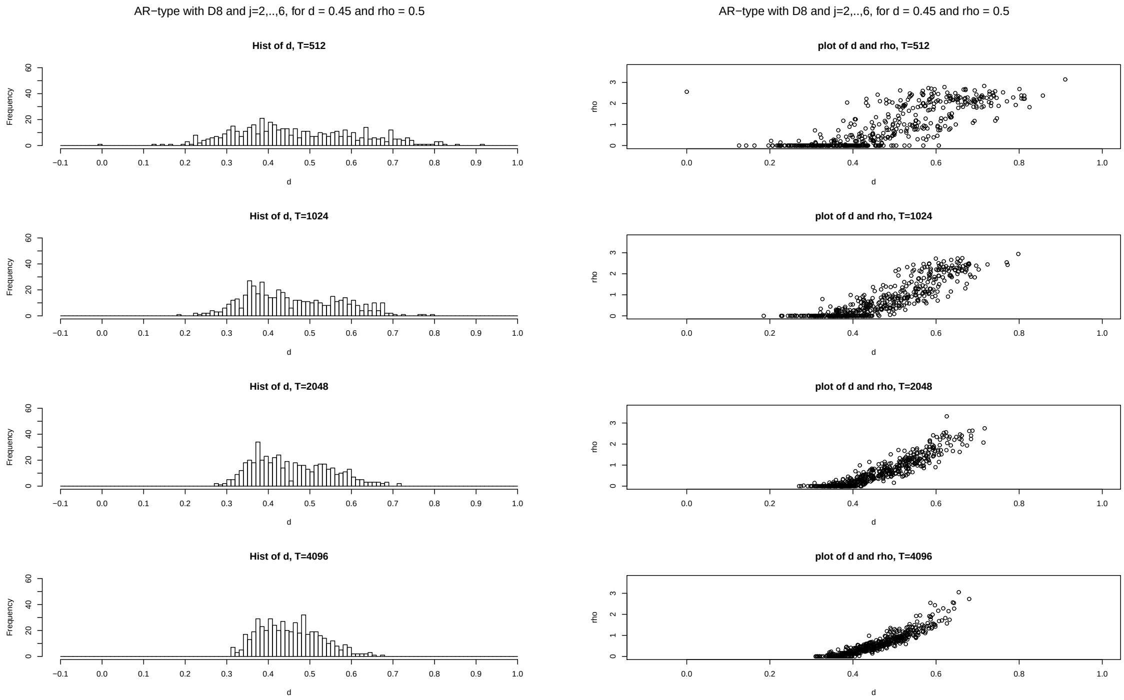


Figure 26: Case 3: Haar, WN-type, $j = 1, 2, \ldots, 6$

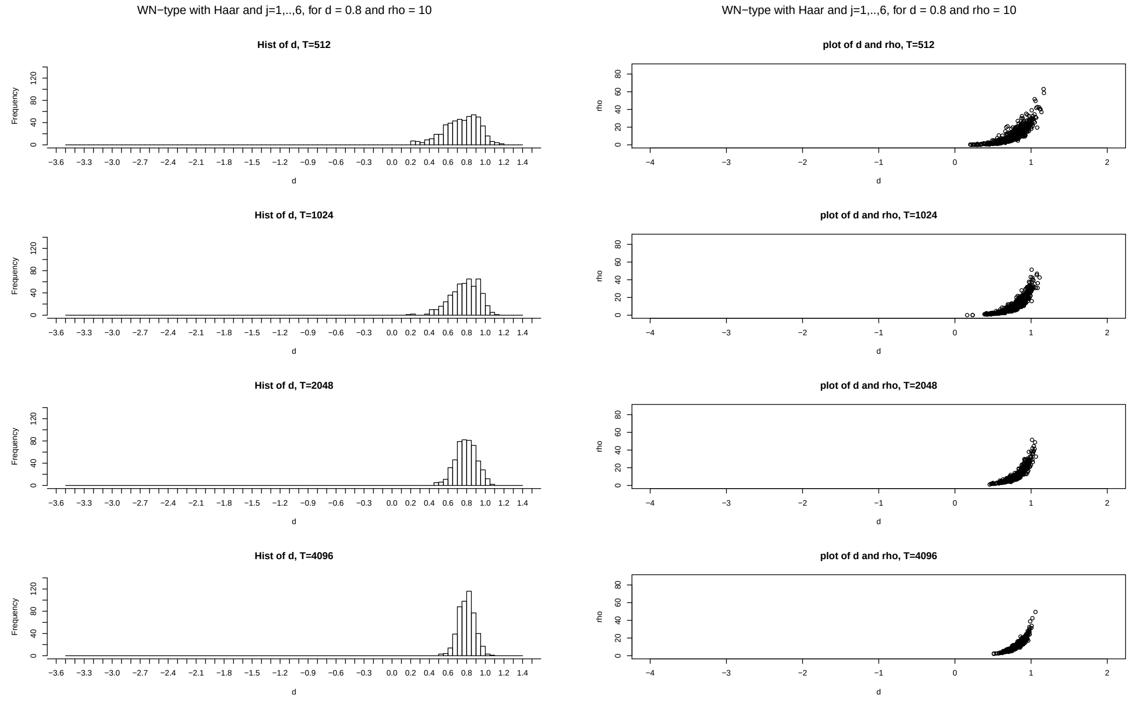


Figure 27: Case 3: Haar, ARMA-type, $j = 1, 2, \ldots, 6$

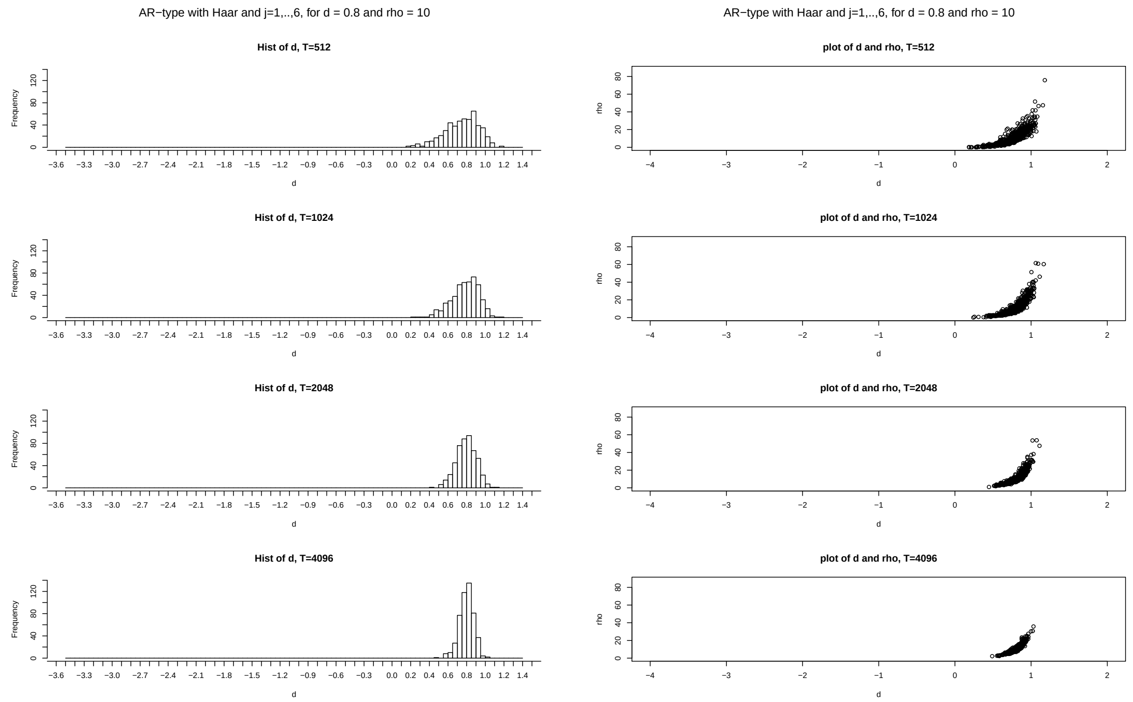


Figure 28: Case 3: Haar, WN-type, $j = 2, \ldots, 6$

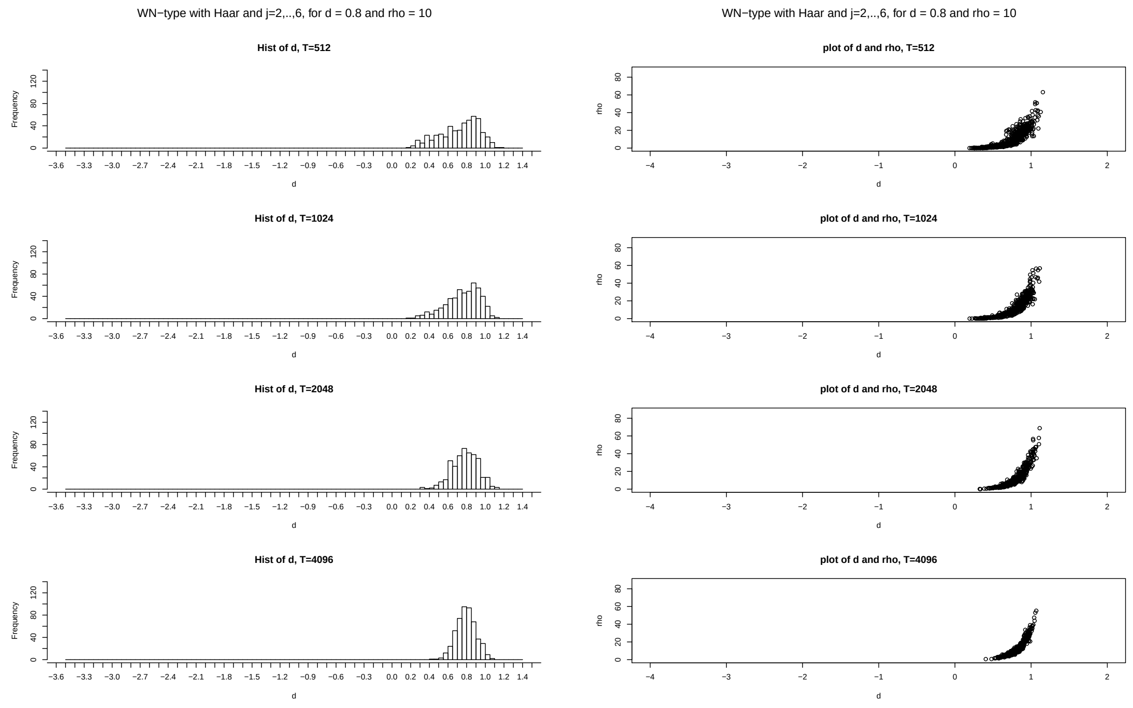


Figure 29: Case 3: Haar, ARMA-type, $j = 2, \ldots, 6$

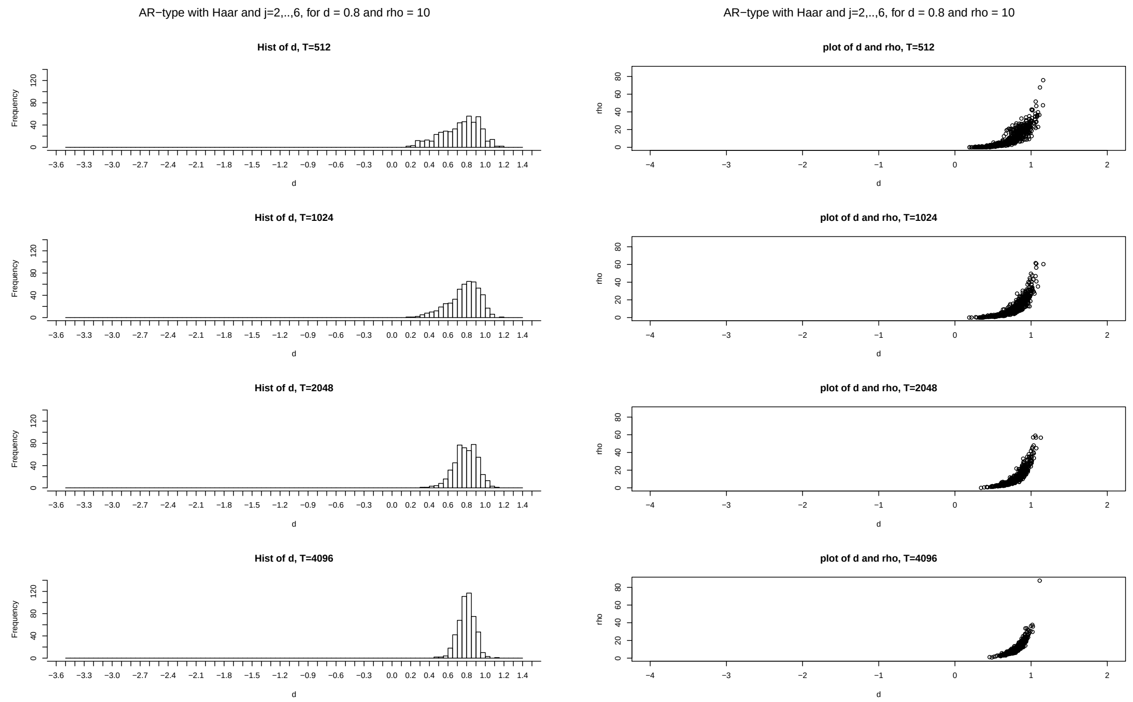


Figure 30: Case 3: D(4), WN-type, $j = 1, 2, \ldots, 6$

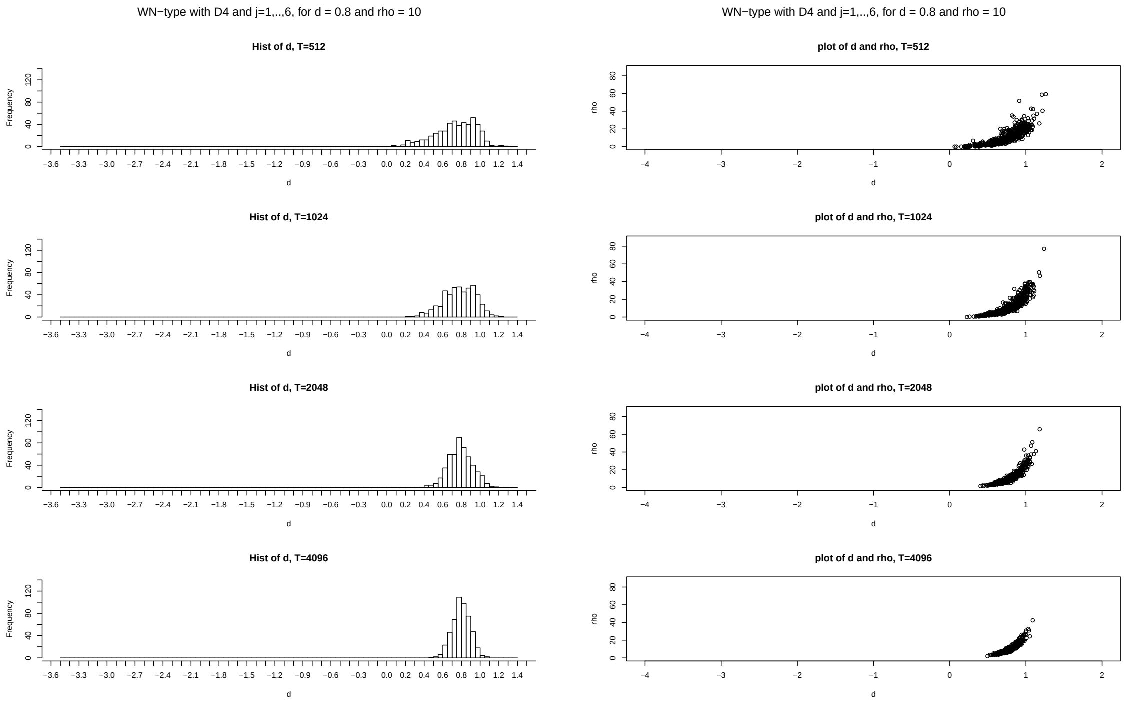


Figure 31: Case 3: D(4), ARMA-type, $j = 1, 2, \ldots, 6$

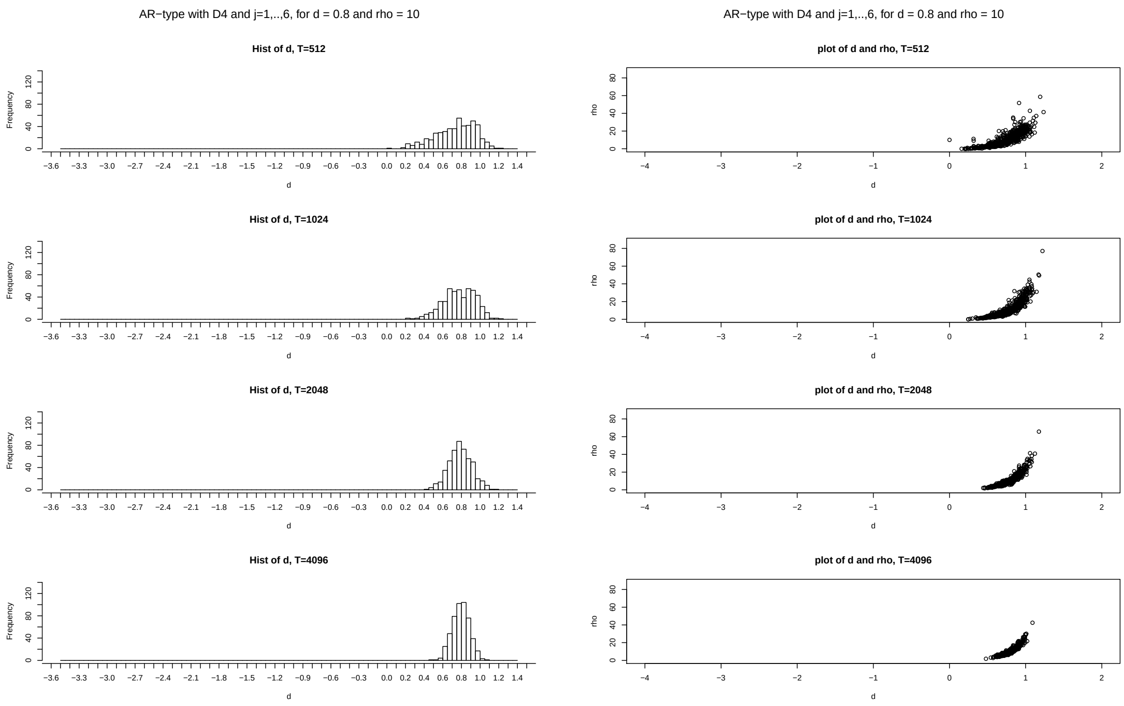


Figure 32: Case 3: D(4), WN-type, $j = 2, \ldots, 6$

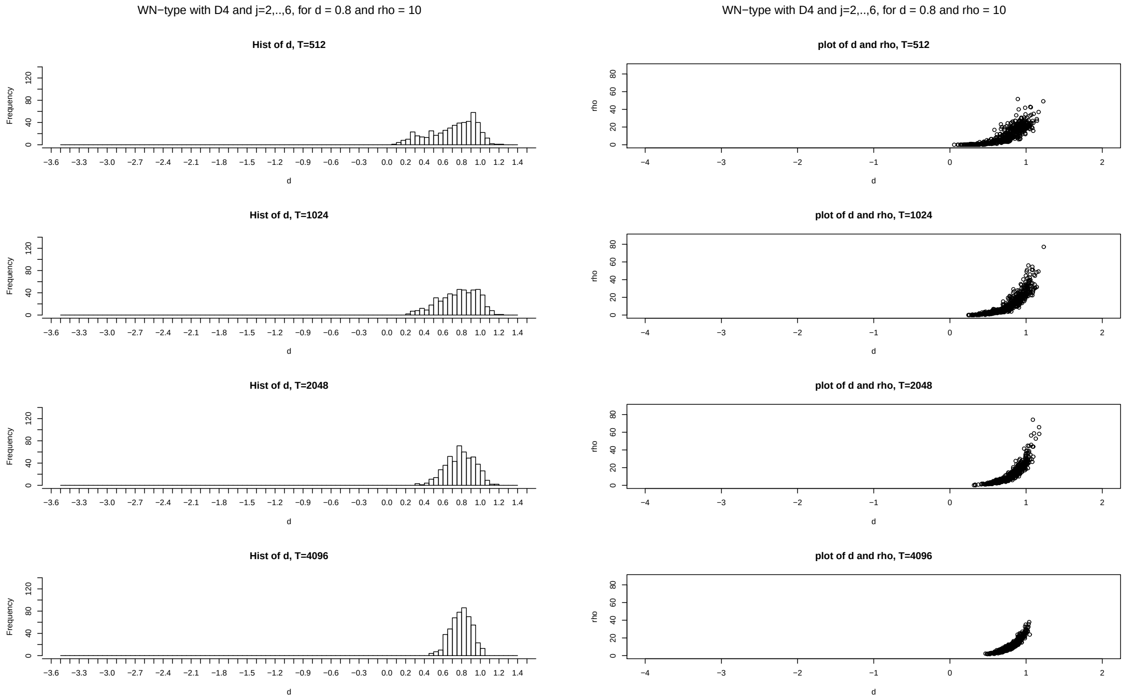


Figure 33: Case 3: D(4), ARMA-type, $j = 2, \ldots, 6$

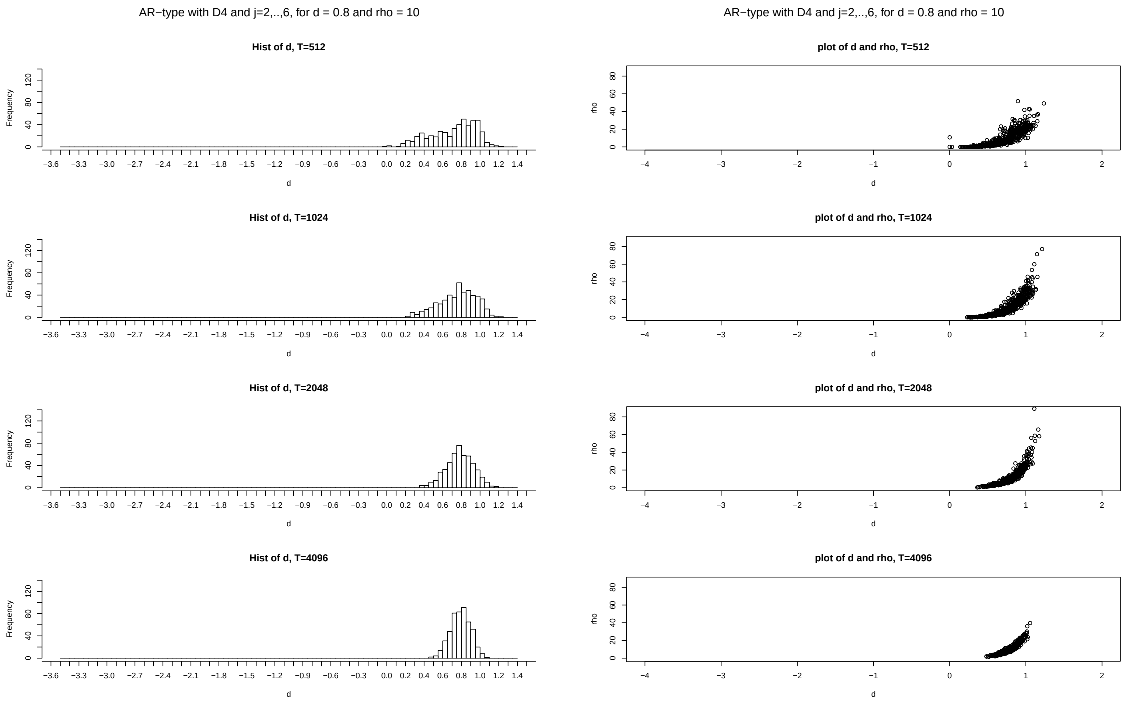


Figure 34: Case 3: D(8), WN-type, $j = 1, 2, \ldots, 6$

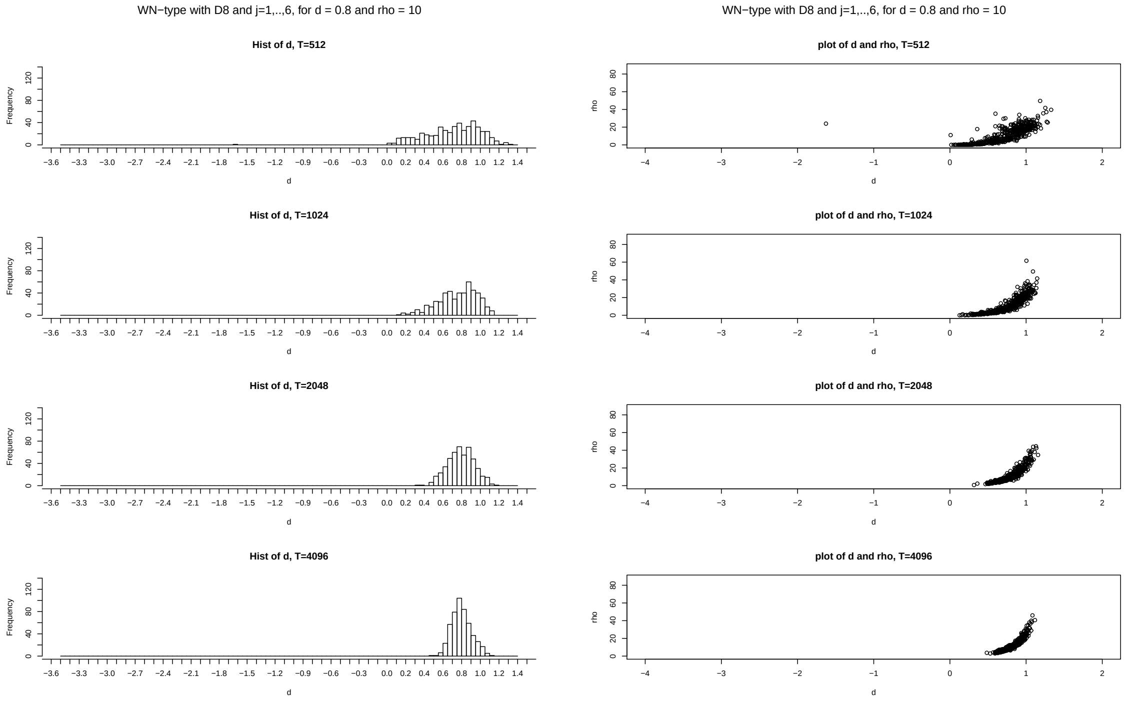


Figure 35: Case 3: D(8), ARMA-type, $j = 1, 2, \ldots, 6$

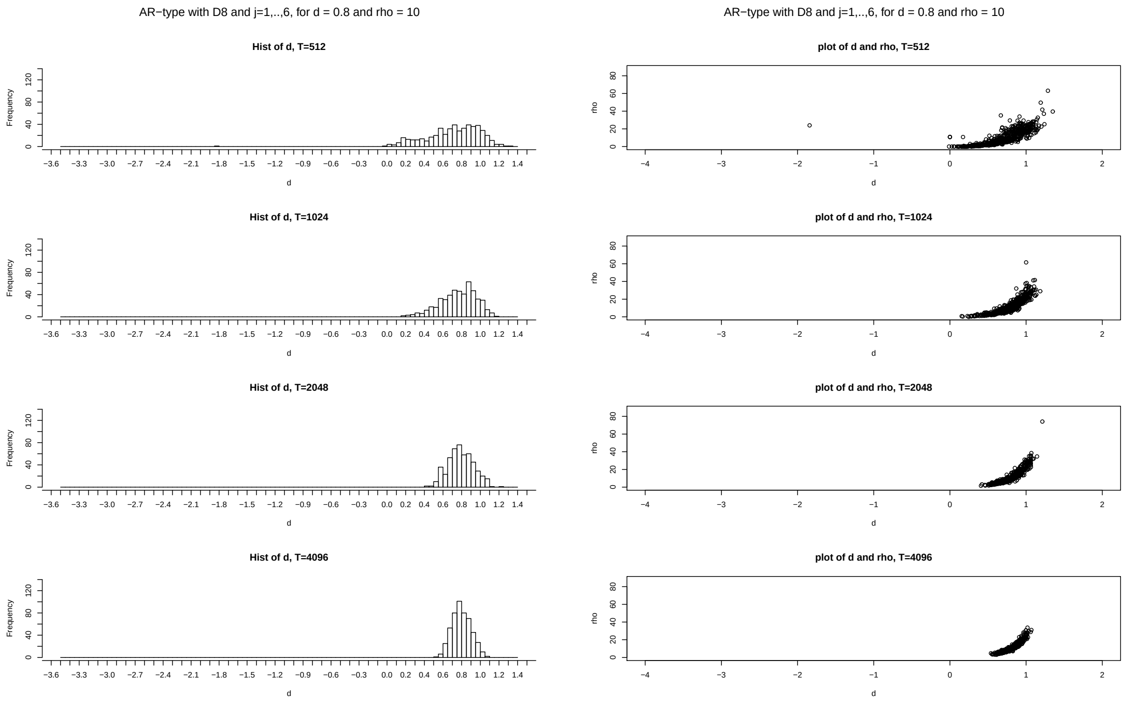


Figure 36: Case 3: D(8), WN-type, $j = 2, \ldots, 6$

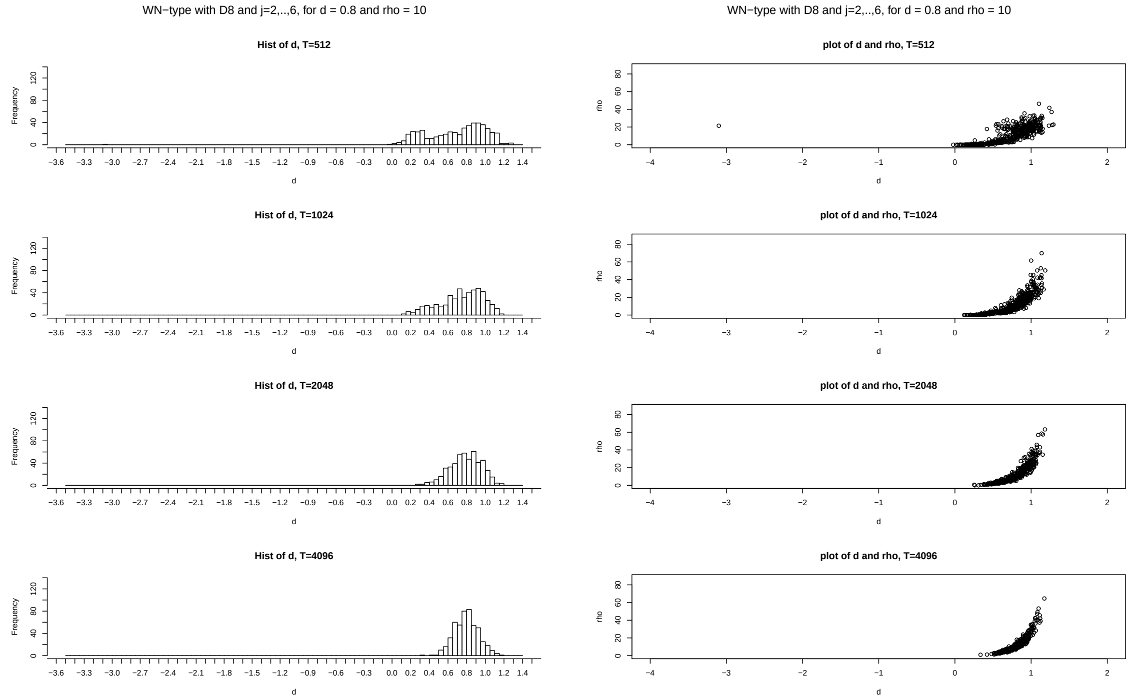


Figure 37: Case 3: D(8), ARMA-type, $j = 2, \ldots, 6$

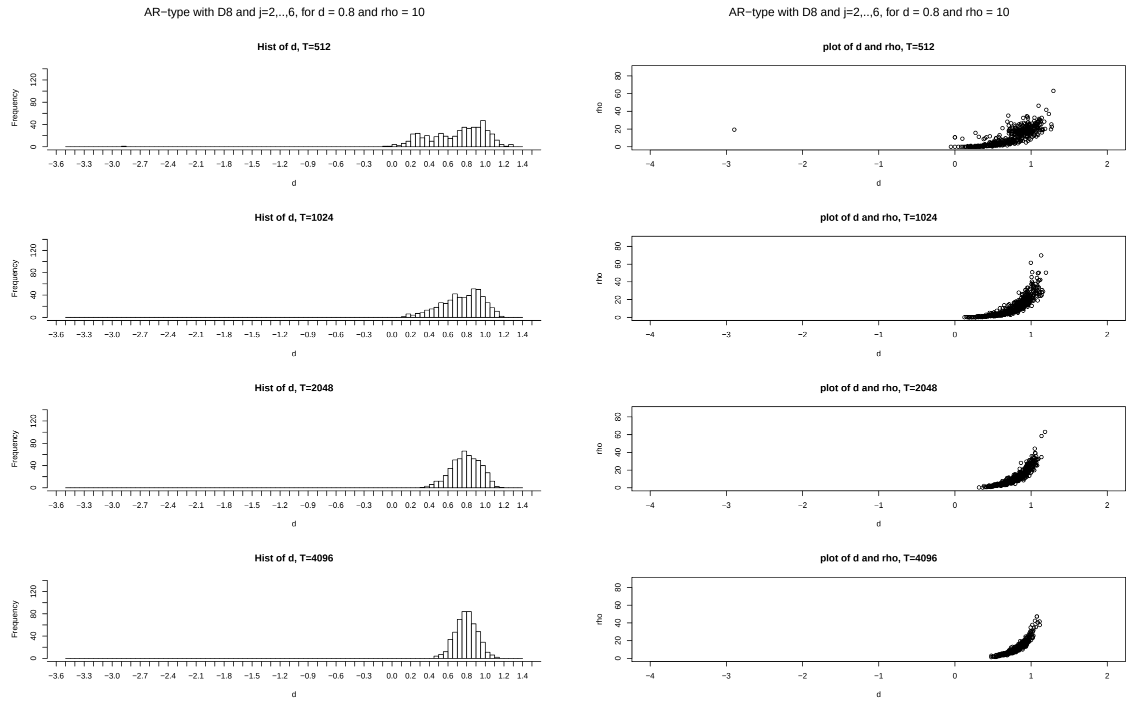


Figure 38: Case 4: D(4), WN-type, $j = 1, 2, \ldots, 6$

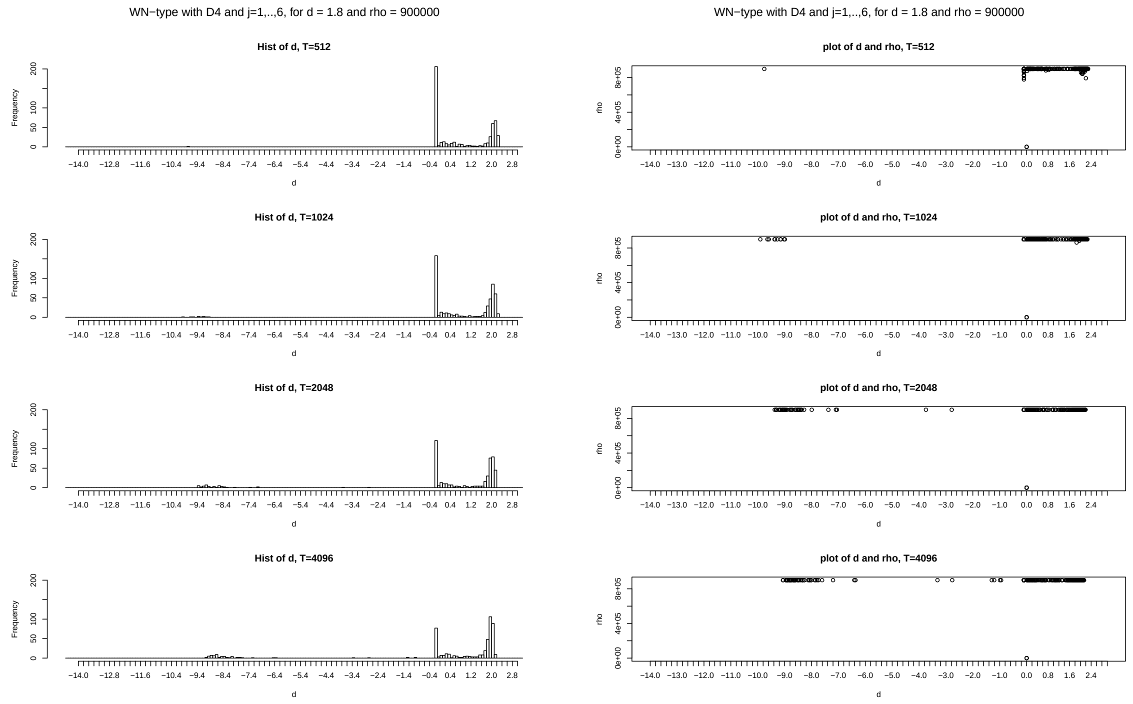


Figure 39: Case 4: D(4), ARMA-type, $j = 1, 2, \ldots, 6$

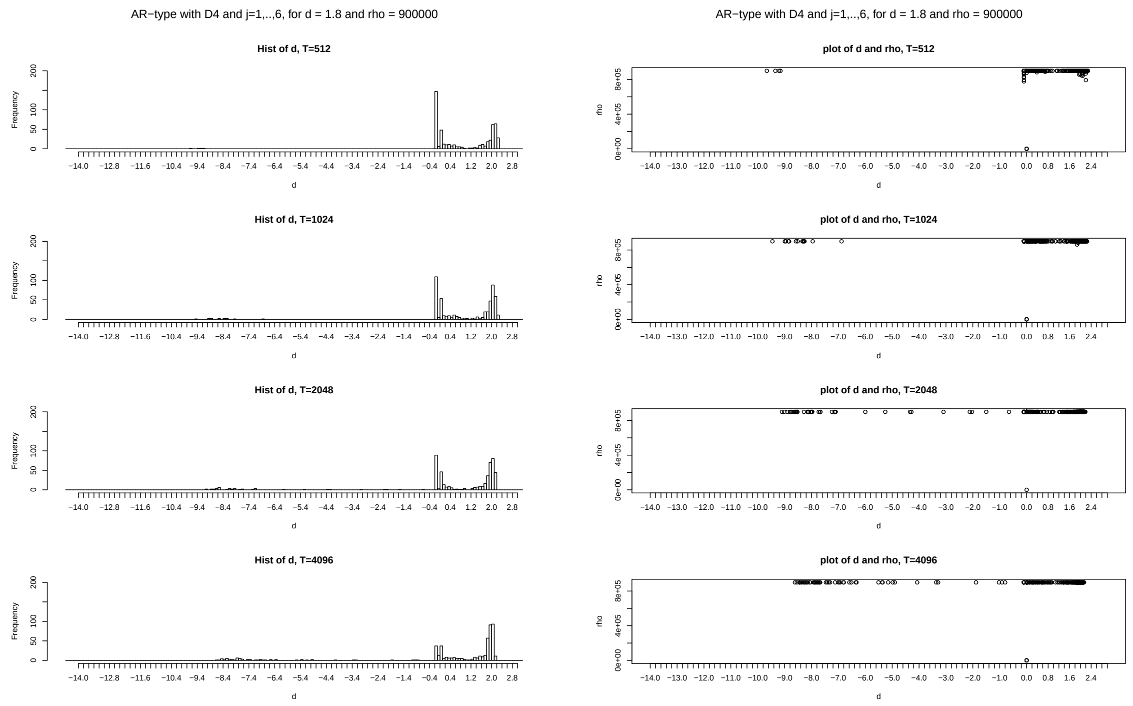


Figure 40: Case 4: D(4), WN-type, $j = 2, \ldots, 6$

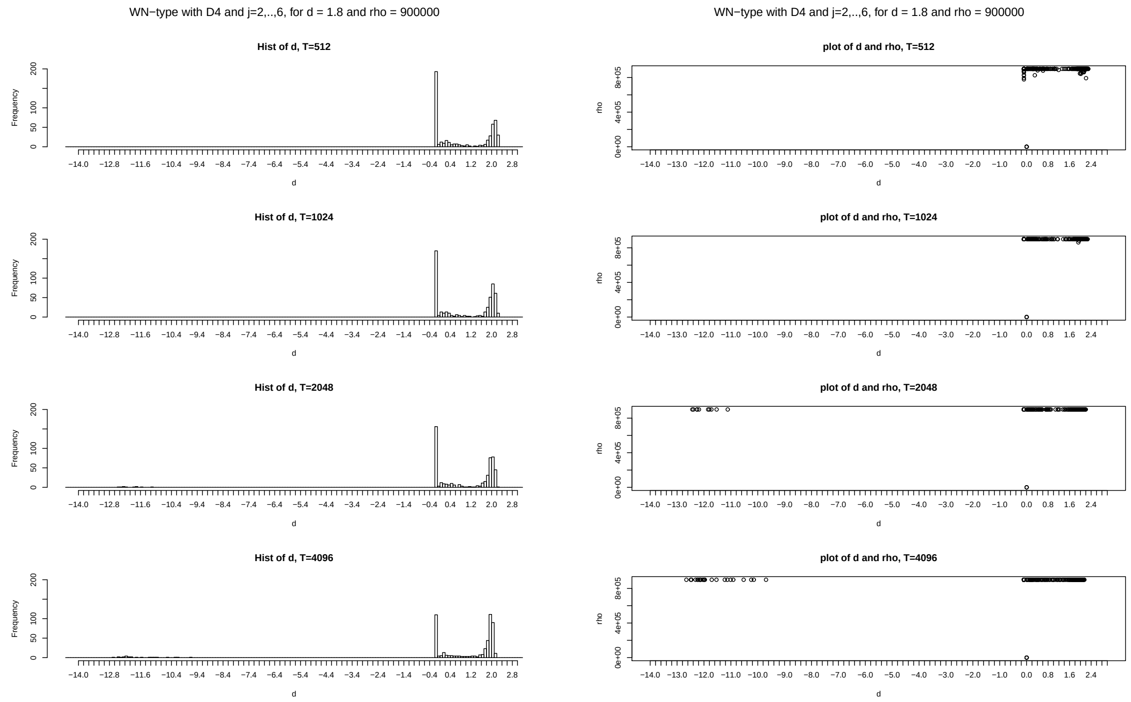


Figure 41: Case 4: D(4), ARMA-type, $j = 2, \ldots, 6$

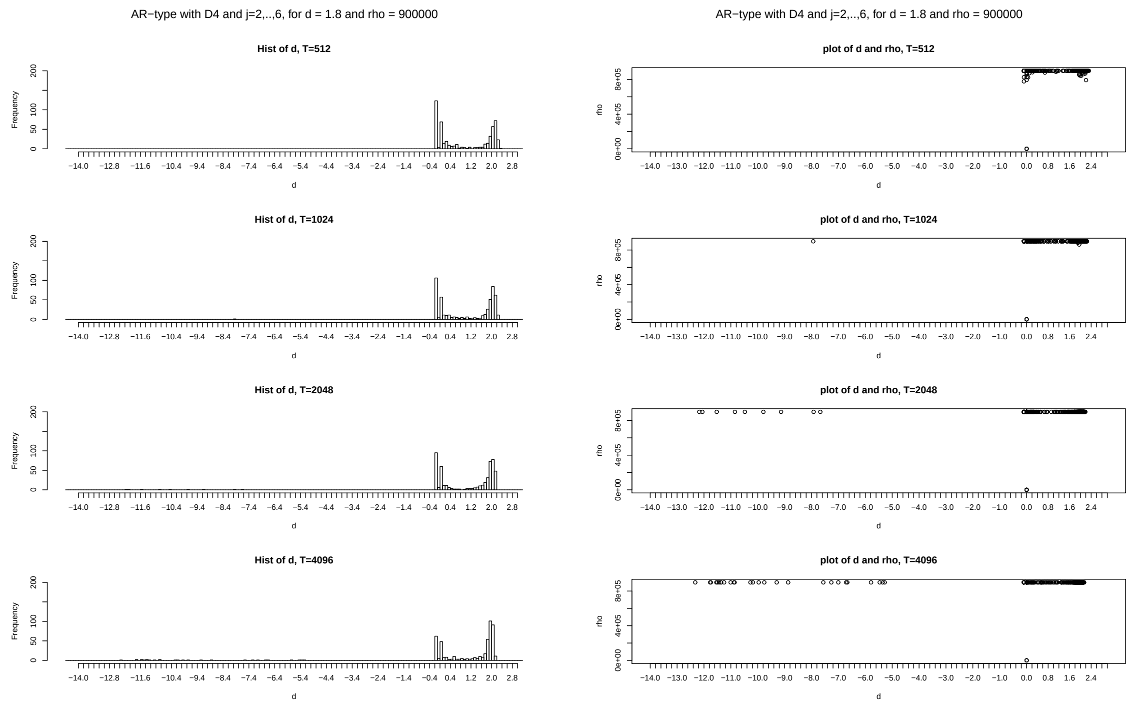


Figure 42: Case 4: D(8), WN-type, $j = 1, 2, \ldots, 6$

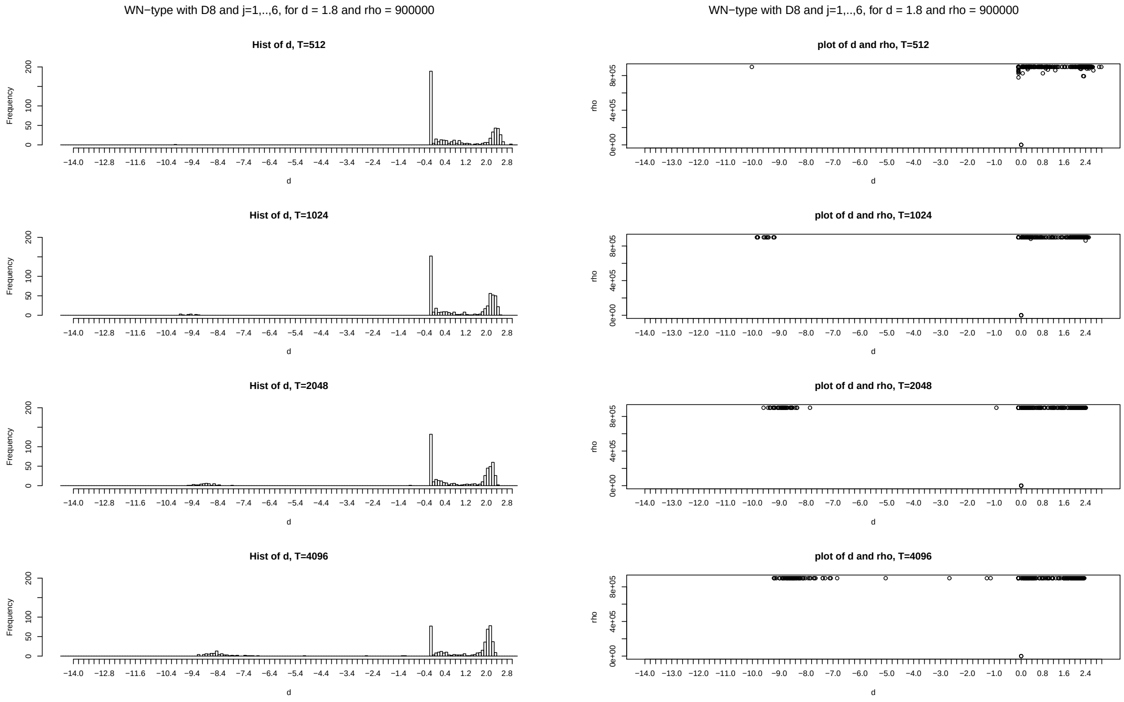


Figure 43: Case 4: D(8), ARMA-type, $j = 1, 2, \ldots, 6$

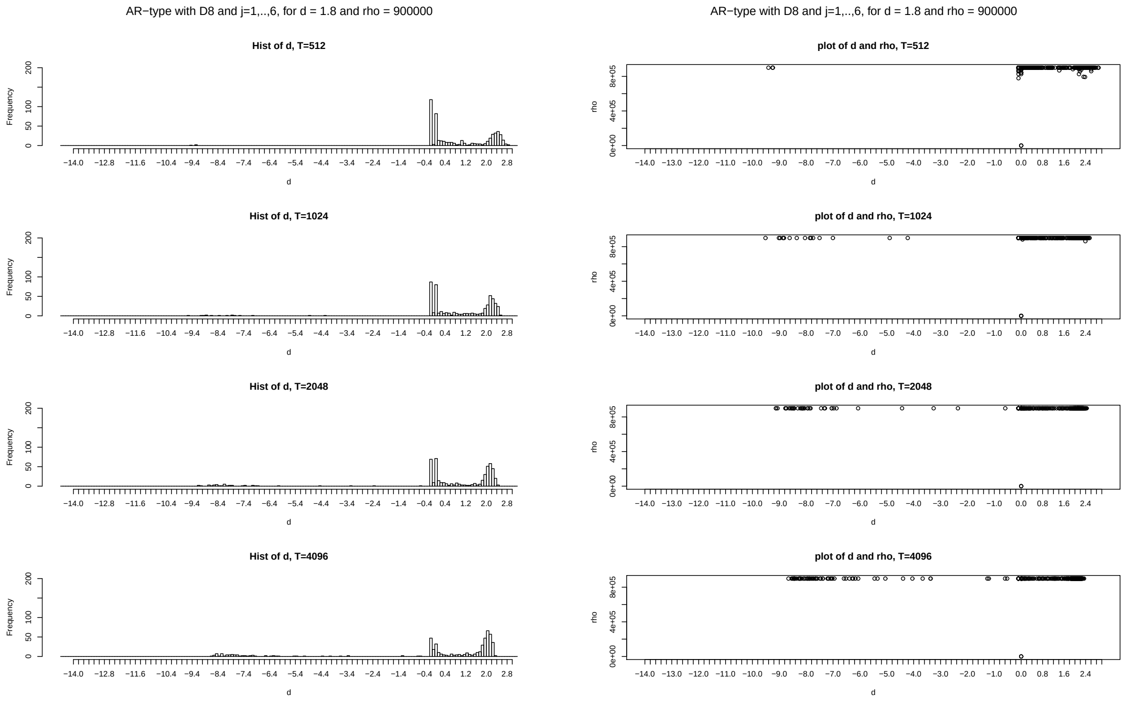


Figure 44: Case 4: D(8), WN-type, $j = 2, \ldots, 6$

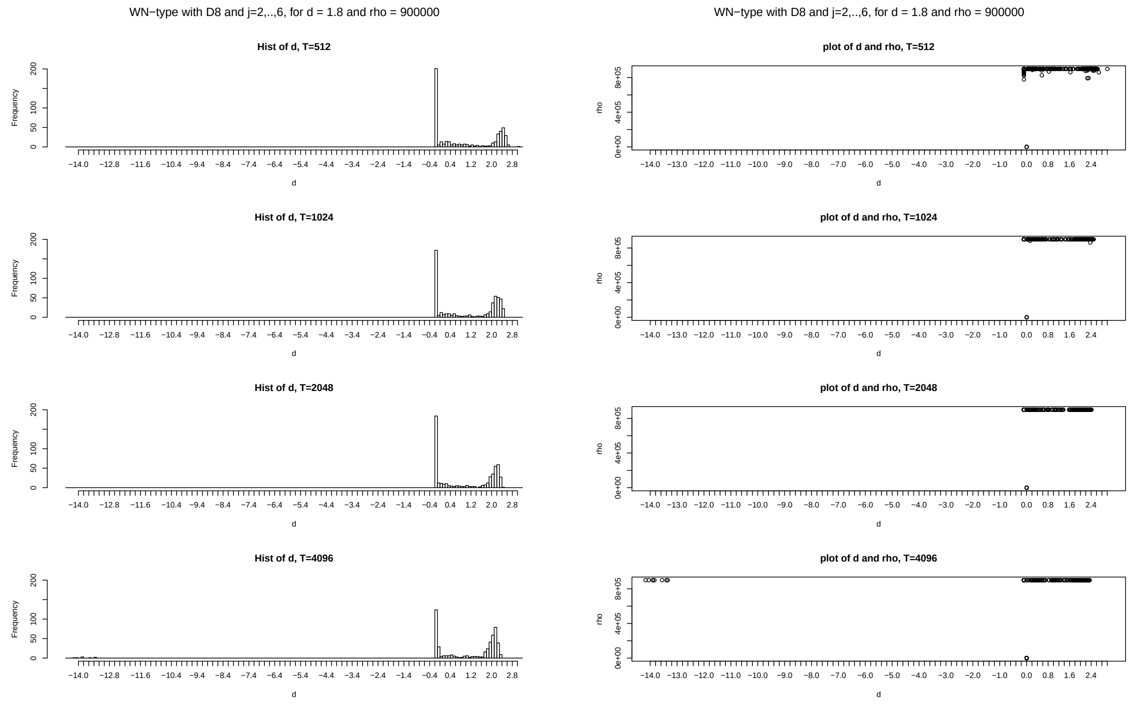


Figure 45: Case 4: D(8), ARMA-type, $j = 2, \ldots, 6$