AUTOREGRESSIVE MODELS WITH EPSILON-SKEW-NORMAL INNOVATIONS

Pascal Bondon

CNRS UMR 8506, Université Paris-Sud, 91192 Gif-sur-Yvette, France.

ABSTRACT

We consider the problem of modelling asymmetric near-Gaussian correlated signals by autoregressive models with epsilon-skew normal innovations. Moments and maximum likelihood estimators of the parameters are proposed and their limit distributions are derived. Monte Carlo simulation results are analyzed and the model is fitted to a real time series.

Index Terms— Non-Gaussian, skewness, autoregressive model, maximum likelihood estimation.

1. INTRODUCTION

In the modelling of non-Gaussian time series, two strategies may be adopted. We may either retain the general autoregressive moving average (ARMA) framework and allow the white noise to be non-Gaussian, or we may completely abandon the linearity assumption, see e.g. [1], [2]. In the former case, the difficulty is to choose the distribution of the white noise appropriately so that the ARMA time series exhibits a specified non-Gaussian feature. In the latter case, one has to find an adequate explicit model among infinitely many nonlinear forms that typically express the time series as a nonlinear function of its lagged values.

In this paper, we are interested in correlated data exhibiting asymmetry and we follow the first strategy. The data are short-range dependent in the sense that their autocorrelations decay to zero exponentially, and their distributions are near-Gaussian. We study the problem of fitting an AR model to these data. Indeed, AR models are very popular in the signal processing community. They are used for instance for spectral analysis, for modelling speech and audio signals, and for identifying systems in control engineering. A causal AR model with a martingale difference sequence as innovations has the nice property that optimal (in the mean square sense) nonlinear infinite past predictors reduce to linear finite past predictors. Furthermore, the so-called Yule-Walker estimator of the parameters of an AR model can be easily calculated using the well-known Levinson-Durbin algorithm, leading to fast implementations. The Yule-Walker estimator is strongly consistent and asymptotically efficient when the innovations are Gaussian, see for instance [3, Chap. 8].

Many non-Gaussian AR models were proposed in the literature. For instance, integer-valued AR processes have been introduced to model weakly dependent sequences of counts, see e.g. [4], [5] and [6] for pioneer works. AR models with exponentially distributed innovations were studied by [7], [8] and [9]. The problem of ARMA modelling with non-Gaussian innovations was addressed by [10]. The authors established general results on maximum likelihood estimates (MLE) of the ARMA parameters and as real examples, they fitted ARMA models with log-normal and gamma innovations to the sunspot and the Canadian lynx data respectively, demonstrating that linear time series model with non-Gaussian innovations can be a useful tool in time series modelling. The estimation of AR models with symmetric innovations that follow a shift-scaled Student's t distribution was considered by [11, 12]. AR models with asymmetric innovations distributed according to gamma and generalized logistic distributions were studied by [13], [14] and [15]. These authors derived modified MLE of the parameters that are easy to compute. On the other hand, a different approach that consists in modelling a non-Gaussian time series as a nonlinear instantaneous transformation of a Gaussian ARMA time series, the nonlinear transformation being determined from the first-order marginal distribution of the data was proposed by [16].

Here, we consider the statistical estimation of an AR model with iid epsilon-skew-normal (ESN) innovations. The ESN distribution, the origin of which can be traced back to [17], was defined by [18] and has been used in regression problems by [19]. Its main advantage is its flexibility since it is analytically tractable, it accommodates practical values of skewness and kurtosis, and it strictly includes the Gaussian distribution. It is therefore of interest to investigate the use of the ESN distribution to model correlated asymmetric data and the purpose of this paper is to consider AR modelling with ESN innovations. The Yule-Walker estimator of the AR parameters is strongly consistent but it is not asymptotically efficient since the innovations are non-Gaussian.

The rest of this paper is organized as follows. The AR model with ESN innovations is presented in Section 2. Moments estimates (ME) and MLE of the parameters are established in Section 3. Numerical simulation results for finite samples are presented in Section 4, and a real data modelling is considered in Section 5. Concluding remarks can be found in Section 6.

2. MODEL DESCRIPTION

Let $f(x) = \exp(-x^2/2)/\sqrt{2\pi}$ be the standard Gaussian density. The ESN distribution is the skewed version of f defined by

$$f_\epsilon(x) = f\left[\frac{x}{1+\epsilon}\right]\mathbb{1}_{\{x<0\}} + f\left[\frac{x}{1-\epsilon}\right]\mathbb{1}_{\{x\geq 0\}},$$

where $\epsilon \in (-1,1)$ is the skew parameter. We extend the family f_ϵ to include location and scale parameters. Let Y be a random variable with density f_ϵ. The family of location-scale ESN distributions is defined as the distribution of $Z = \mu + \sigma Y$ for $\mu \in \mathbb{R}$ and $\sigma > 0$. The corresponding density is given by

$$f_\theta(x) = \frac{1}{\sigma} f_\epsilon\left[\frac{x-\mu}{(1+\epsilon)\sigma}\right]\mathbb{1}_{\{x<\mu\}} + \frac{1}{\sigma} f_\epsilon\left[\frac{x-\mu}{(1-\epsilon)\sigma}\right]\mathbb{1}_{\{x\geq\mu\}},$$

where $\theta = (\epsilon, \mu, \sigma)$, and we denote $Z \sim \mathrm{ESN}(\epsilon,\mu,\sigma)$. The distribution of Z is unimodal with mode at μ and it has probability mass $(1+\epsilon)/2$ below the mode. If $Z \sim \mathrm{ESN}(\epsilon, 0, 1)$, then Z has the same distribution as UV where U and V are two independent random variables, U is discrete with $\mathbb{P}(U = 1-\epsilon) = (1-\epsilon)/2$, $\mathbb{P}(U = -1-\epsilon) = (1+\epsilon)/2$, and V is absolutely continuous with the density $2f(x)\,\mathbb{1}_{\{x\geq 0\}}$ (V has the same distribution as $|S|$ where S is a standard Gaussian random variable), see [20]. This stochastic representation of Z is useful for generating realizations of Z and for calculating the moments of Z. Specifically, if $Z \sim \mathrm{ESN}(\epsilon,\mu,\sigma)$, $\mathrm{E}(Z) = \mu - 4\sigma\epsilon/\sqrt{2\pi}$ and the kth central moment of Z is

$$\mathrm{E}(Z - \mathrm{E}\,Z)^k = \frac{\sigma^k}{\sqrt{2\pi}} \sum_{l=0}^{k} \binom{k}{l} \left(\frac{4\epsilon}{\sqrt{2\pi}}\right)^{k-l} \left((-1)^l I_l(-\epsilon) + I_l(\epsilon)\right),$$

where

$$I_l(\epsilon) = \begin{cases} \sqrt{\frac{\pi}{2}}(1-\epsilon)^{l+1}\prod_{i=1}^{m}(2i-1) & \text{if } l = 2m, \\ 2^m(1-\epsilon)^{l+1}m! & \text{if } l = 2m+1. \end{cases}$$

Therefore, $(-1)^l I_l(-\epsilon) + I_l(\epsilon)$ is a polynomial of degree l and the kth central moment of Z takes the form $\sigma^k P_k(\epsilon)$ where P_k is a polynomial of degree k. For $k > 1$, the kth cumulant $c_{k,Z}$ of Z is obtained from the lth central moments for $l \leq k$ by means of well-known polynomial relations, see for instance [21, eqn (3.43)]. It follows from these relations that $c_{k,Z}$, $k > 1$, takes also the form $\sigma^k P'_k(\epsilon)$ where P'_k is a polynomial of degree k. The four firsts cumulants of Z are

$$\begin{aligned} c_{1,Z} &= \mu - 4\sigma\epsilon/\sqrt{2\pi}, \\ c_{2,Z} &= \frac{\sigma^2}{\pi}[(3\pi - 8)\epsilon^2 + \pi], \\ c_{3,Z} &= \frac{2\sqrt{2}\sigma^3\epsilon}{\pi^{3/2}}[(5\pi - 16)\epsilon^2 - \pi], \\ c_{4,Z} &= \frac{4\sigma^4\epsilon^2}{\pi^2}[(-3\pi^2 + 40\pi - 96)\epsilon^2 + \pi(3\pi - 8)]. \end{aligned} \tag{1}$$

Since $\epsilon \in (-1,1)$, $c_{3,Z}/c_{2,Z}^{3/2} \in (-c_0, c_0)$ where $c_0 = \sqrt{2}(4-\pi)(\pi-2)^{-3/2} = 0.995$, and $c_{4,Z}/c_{2,Z}^2 \in (0, 0.870)$. The ESN distribution is therefore useful for modelling asymmetric data with slight leptokurticity. Of course, the $\mathrm{ESN}(\epsilon,\mu,\sigma)$ distribution reduces to the Gaussian distribution with mean μ and variance σ^2 when $\epsilon = 0$.

An AR(p) model with ESN innovations is defined by the difference equation

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t, \tag{2}$$

where (Z_t) is a sequence of independent and identically distributed (iid) random variables with $Z_t \sim \mathrm{ESN}(\epsilon,\mu,\sigma)$, and polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ with real coefficients has no zeros in the closed unit disk $\{z \in \mathbb{C} : |z| \leq 1\}$. The (unique) stationary solution (X_t) of (2) has the MA representation $X_t = \sum_{i=0}^{\infty}\psi_i Z_{t-i}$, where (ψ_i) are the coefficients in the Taylor series expansion of $1/\phi(z)$ for $|z| \leq 1$. We have $\sum|\psi_i| < +\infty$, and then finiteness of $\mathrm{E}(|Z_t|^k)$ imply finiteness of $\mathrm{E}(|X_t|^k)$ for all $k \geq 1$.

3. PARAMETER ESTIMATION

Fitting model (2) to some data consists in choosing p and estimating the parameter vector $\eta = (\phi', \mu, \epsilon, \sigma^2)'$ where $\phi = (\phi_1, \ldots, \phi_p)'$ and ϕ' denotes the transpose of ϕ. In all the following, we assume that η lies in the open set $S = C \times \mathbb{R} \times (-1,1) \times (0,\infty)$, where C is the interior of the domain of vectors ϕ such that $\phi(z)$ has no zeros in the closed unit disk. We shall denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^{p+3}$, so that $\|\eta\|^2 = \eta'\eta$. We shall first propose ME which will be used as initial values in a quasi-Newton method to find MLE.

The standard Yule-Walker equations for model (2) are

$$M_2\phi = m_2,$$

where M_2 is the invertible covariance matrix $[m_{2,i-j}]_{i,j=1}^{p}$, $m_2 = (m_{2,1}, \ldots, m_{2,p})'$ and $m_{2,k} = \mathrm{E}(X_0 X_k) - \mathrm{E}(X_0)^2$. The Yule-Walker estimator $\widehat{\phi}_n$ of parameter vector ϕ based on observations $(X_t)_{t=1}^n$ is $\widehat{\phi}_n = \widehat{M}_2^{-1}\widehat{m}_2$, where $\widehat{M}_2$ is the sample covariance matrix $[\widehat{m}_{2,|i-j|}]_{i,j=1}^{p}$, $\widehat{m}_2 = (\widehat{m}_{2,1}, \ldots, \widehat{m}_{2,p})'$ with $\widehat{m}_{2,k} = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \overline{X})(X_{t+k} - \overline{X})$ and $\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$. According to [3, Theorem 8.1.1], $\widehat{\phi}_n \xrightarrow{a.s.} \phi$ and $\sqrt{n}(\widehat{\phi}_n - \phi) \xrightarrow{d} \mathcal{N}(0, c_{2,Z}M_2^{-1})$ as $n \to \infty$. Moreover, the covariance matrix $c_{2,Z}M_2^{-1}$ depends only on the parameter vector ϕ.

Let

$$\widehat{Z}_t = X_t - \widehat{\phi}_{n,1}X_{t-1} - \cdots - \widehat{\phi}_{n,p}X_{t-p}.$$

The ME of $(\mu, \epsilon, \sigma^2)$, denoted by $(\widehat{\mu}_n, \widehat{\epsilon}_n, \widehat{\sigma}_n^2)$ are obtained by solving the first three equations in (1) where the left-hand sides are replaced by the corresponding sample cumulants obtained from $(\widehat{Z}_t)$. Let $k = 2, 3$, and

$$\widehat{c}_{k,Z} = \frac{1}{n}\sum_{t=p+1}^{n}\left(\widehat{Z}_t - \overline{\widehat{Z}}\right)^k \quad \text{where} \quad \overline{\widehat{Z}} = \frac{1}{n}\sum_{t=p+1}^{n}\widehat{Z}_t,$$

we obtain,

$$\begin{aligned}\widehat{\epsilon}_n &= g^{-1}(\widehat{c}_{3,Z}/\widehat{c}_{2,Z}^{3/2}),\\ \widehat{\sigma}_n^2 &= \frac{\pi\widehat{c}_{2,Z}}{(3\pi-8)\widehat{\epsilon}_n^2+\pi},\\ \widehat{\mu}_n &= \overline{\widehat{Z}} + 4\widehat{\sigma}_n\widehat{\epsilon}_n/\sqrt{2\pi},\end{aligned} \tag{3}$$

where $g : (-1, 1) \to (-c_0, c_0)$ is defined by

$$g(x) = 2\sqrt{2}x\frac{(5\pi-16)x^2-\pi}{[(3\pi-8)x^2+\pi]^{3/2}}. \tag{4}$$

Function g is continuously differentiable on $(-1, 1)$ with derivative

$$g'(x) = 2\sqrt{2}\pi\frac{(21\pi-64)x^2-\pi}{[(3\pi-8)x^2+\pi]^{5/2}}.$$

On the interval $(-1, 1)$, $g' < 0$ and then g is strictly monotone which implies that g is an homeomorphism from $(-1, 1)$ onto $(-c_0, c_0)$. If $\widehat{c}_{3,Z}/\widehat{c}_{2,Z}^{3/2} \notin (-c_0, c_0)$, the ME $(\widehat{\mu}_n, \widehat{\epsilon}_n, \widehat{\sigma}_n^2)$ are not defined. Since

$$\widehat{Z}_t = Z_t + \sum_{i=1}^{p}(\phi_i - \widehat{\phi}_{n,i})X_{t-i},$$

and $\widehat{\phi}_n \xrightarrow{a.s.} \phi$, $\mathrm{E}\,|Z_t^k X_{t-1}^{k_1}\dots X_{t-p}^{k_p}| < \infty$ for all nonnegative integers $k, k_1, \dots, k_p$, we have $\overline{\widehat{Z}} \xrightarrow{a.s.} \mathrm{E}(Z_t)$ and $\widehat{c}_{k,Z} \xrightarrow{a.s.} \mathrm{E}(Z_t - \mathrm{E}(Z_t))^k$ as $n \to \infty$. Hence, the continuity of transformation (3) implies that $(\widehat{\mu}_n, \widehat{\epsilon}_n, \widehat{\sigma}_n^2) \xrightarrow{a.s.} (\mu, \epsilon, \sigma^2)$ as $n \to \infty$. Furthermore, it can be shown that the ME $\widehat{\eta}_n = (\widehat{\phi}_n', \widehat{\mu}_n, \widehat{\epsilon}_n, \widehat{\sigma}_n^2)'$ of η is asymptotically normal at the usual rate $\sqrt{n}$.

Since (Z_t) is non-Gaussian, $\widehat{\eta}_n$ is not asymptotically efficient in general. We now present the MLE of η, and in the following we suppose that η_0 is the true value of η. We consider the likelihood estimator based on maximization of the conditional likelihood of $(X_1, \dots, X_n)$ conditionally to $(X_1, \dots, X_p)$. According to (2), the logarithm of the conditional likelihood is

$$L_n(\eta) = \textstyle\sum_{t=p+1}^{n} \ln f_\theta(X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}).$$

Our main result is the following.

Theorem 1. Let (X_t) be defined by (2) where ϕ is replaced by ϕ_0 and (Z_t) are iid random variables with an $\mathrm{ESN}(\epsilon_0, \mu_0, \sigma_0)$ distribution, and let $\eta_0 = (\phi_0', \mu_0, \epsilon_0, \sigma_0^2)' \in S$. Then, there exists a sequence of estimators $(\widetilde{\eta}_n)$ such that, for any $\epsilon > 0$, there exists an event E with $\mathsf{P}(E) > 1 - \epsilon$ and an n_0 such that on E, for $n > n_0$, $\frac{\partial L_n}{\partial \eta}(\widetilde{\eta}_n) = 0$ and L_n attains a relative maximum at $\widetilde{\eta}_n$. Furthermore, $\widetilde{\eta}_n \xrightarrow{a.s.} \eta_0$ and $\sqrt{n}(\widetilde{\eta}_n - \eta_0) \xrightarrow{d} \mathrm{N}(0, \Sigma)$ as $n \to \infty$, where

$$\Sigma = \sigma_0^2(1-\epsilon_0^2)\begin{pmatrix} M_2^{-1} & -mM_2^{-1}e & 0 & 0\\ -me'M_2^{-1} & c_1 & c_2 & 0\\ 0 & c_2 & c_3 & 0\\ 0 & 0 & 0 & c_4\end{pmatrix}, \tag{5}$$

$$m = \mathrm{E}(X_t) = (\mu_0 - 4\sigma_0\epsilon_0/\sqrt{2\pi})/(1 - e'\phi_0),$$

$$e = (1, \dots, 1)', \quad c_1 = \frac{3\pi}{3\pi-8} + m^2 e' M_2^{-1} e,$$

$$c_2 = \frac{2\sqrt{2\pi}}{(3\pi-8)\sigma_0}, \quad c_3 = \frac{\pi}{(3\pi-8)\sigma_0^2}, \quad c_4 = \frac{2\sigma_0^2}{1-\epsilon_0^2},$$

and M_2 is calculated for $\eta = \eta_0$. The covariance matrix Σ can be estimated strongly consistently by replacing η_0 by $\widetilde{\eta}_n$ in its expression. One may also replace M_2 by the estimated covariance matrix $[\widehat{m}_{2,i-j}]_{i,j=1}^p$.

Remark 1. The MLE $\widetilde{\eta}_n$ is asymptotically efficient, i.e., Σ is the inverse of the Fisher information matrix of η evaluated at η_0.

Remark 2. The MLE $\widetilde{\sigma}_n^2$ and $\widetilde{\phi}_n$ are asymptotically independent of $(\widetilde{\phi}_n', \widetilde{\mu}_n, \widetilde{\epsilon}_n)$ and $\widetilde{\epsilon}_n$, respectively. This property does not hold for the ME. Furthermore, the asymptotic covariance of $\widetilde{\phi}_n$ is reduced compared to the asymptotic covariance of the Yule-Walker estimator $\widehat{\phi}_n$ by the factor

$$\frac{c_{2,Z}}{\sigma_0^2(1-\epsilon_0^2)} = \frac{(3\pi-8)\epsilon_0^2+\pi}{\pi(1-\epsilon_0^2)} \in [1, \infty).$$

This factor is a strictly increasing function of $|\epsilon_0|$ and is equal to 1 in the Gaussian case $\epsilon_0 = 0$. On the other hand, the asymptotic variances of $\widetilde{\phi}_n, \widetilde{\mu}_n, \widetilde{\epsilon}_n$ and $\widetilde{\sigma}_n^2$ depend on (ϕ_0, ϵ_0), $(\phi_0, \mu_0, \epsilon_0, \sigma_0^2)$, ϵ_0 and σ_0^2, respectively.

Remark 3. When the skewness ϵ_0 is known a priori and is not estimated, the asymptotic covariance Ψ of the MLE of $(\phi_0', \mu_0, \sigma_0^2)'$ is obtained by inverting the matrix $V = \Sigma^{-1}$ whose $(p+2)$th row and $(p+2)$th column have been deleted. We obtain that

$$\Psi = \sigma_0^2(1-\epsilon_0^2)\begin{pmatrix} M_2^{-1} & -mM_2^{-1}e & 0\\ -me'M_2^{-1} & c_8 & 0\\ 0 & 0 & c_4\end{pmatrix}, \tag{6}$$

where $c_8 = 1 + m^2 e' M_2^{-1} e$ and M_2, m are calculated for $\eta = \eta_0$. Therefore, the asymptotic variance of the MLE of the location μ_0 is reduced by the factor

$$\frac{c_1}{c_8} = \frac{\frac{3\pi}{3\pi-8} + m^2 e' M_2^{-1} e}{1 + m^2 e' M_2^{-1} e}$$

when the value of the skewness, if known a priori, is used. The asymptotic covariances of $\widetilde{\phi}_n$ and $\widetilde{\sigma}_n^2$ are unchanged.

Remark 4. When it is known that (X_t) is Gaussian and thus the skewness is not estimated, $(\widetilde{\phi}_n, \widetilde{\mu}_n)$ coincide with the usual least squares estimates (LSE) $(\phi_n^{\mathrm{LS}}, \mu_n^{\mathrm{LS}})$ obtained by minimizing the sum of squares

$$S(\phi, \mu) = \sum_{t=p+1}^{n} (X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} - \mu)^2,$$

and $\widetilde{\sigma}_n^2 = \sigma_n^{2,\mathrm{LS}} = (n-p)^{-1} S(\phi_n^{\mathrm{LS}}, \mu_n^{\mathrm{LS}})$. The corresponding asymptotic covariance is given by (6) where $\epsilon_0 = 0$.

Remark 5. It is instructive to study the properties of estimator $(\phi_n^{\mathrm{LS}}, \mu_n^{\mathrm{LS}}, \sigma_n^{2,\mathrm{LS}})$ when $\epsilon_0 \neq 0$. According to the standard theory, $(\phi_n^{\mathrm{LS}}, \mu_n^{\mathrm{LS}}, \sigma_n^{2,\mathrm{LS}}) \xrightarrow{a.s.} (\phi_0, c_{1,Z}, c_{2,Z})$ and $\sqrt{n}(\phi_n^{\mathrm{LS}} - \phi_0) \xrightarrow{d} \mathrm{N}(0, c_{2,Z} M_2^{-1})$. Therefore, when $\epsilon_0 \neq 0$, ϕ_n^{LS} is consistent but is not efficient, and we deduce from (1) that μ_n^{LS} tends to overestimate the location μ when $\epsilon_0 < 0$ and tends to underestimate μ when $\epsilon_0 > 0$, and $\sigma_n^{2,\mathrm{LS}}$ tends to overestimate σ^2.

4. MONTE CARLO SIMULATIONS

In this section, we illustrate the finite sample behaviors of the ME $\widehat{\eta}_n$ and the MLE $\widetilde{\epsilon}_n$ by Monte Carlo simulations. All the experiments are based on 1000 replications, and the number of data considered is $n = 1000$. For each realization, $\widehat{\eta}_n$ is used as initial value in a quasi-Newton method to find $\widetilde{\eta}_n$. The data generating process is a causal AR(1) model with $Z_t \sim \mathrm{ESN}(\epsilon, \mu, \sigma)$. We fix $\phi = 0.8$, $\mu = 10$, $\sigma^2 = 1$ and ϵ varies in $(-1, 1)$. In Figure 1, we plot the mean square errors (MSE) of $\widehat{\eta}_n$ and $\widetilde{\eta}_n$, and the asymptotic variances of $\widetilde{\eta}_n$ as a function of ϵ. For the four parameters, we observe that the MSE of $\widehat{\eta}_n$ are significantly larger than those of $\widetilde{\eta}_n$, and the differences between the MSE increase as the absolute value of ϵ increases. Furthermore, the MSE of $\widetilde{\eta}_n$ are close from the asymptotic variances given by (5).

5. REAL DATA EXAMPLE

We consider the Dow-Jones Utilities index between July 3, 1972 and December 20, 1972. The very slowly decaying positive sample autocorrelation function of this time series suggests differencing at lag one before attempting to fit a stationary model, see [22, Example 5.1.1]. Figure 2 shows that the differenced series is asymmetric and may be modelled by an AR(1) process.

We fit an AR(1) model with ESN innovations and we compare with an AR(1) model with Gaussian innovations. The results are given in Table 1 where the variances are calculated with (5) where η_0 is replaced by $\widetilde{\eta}_n$, and (6) where $\epsilon_0 = 0$ and $(\phi_0, \mu_0, \sigma_0^2)$ is replaced by $(\phi_n^{\mathrm{LS}}, \mu_n^{\mathrm{LS}}, \sigma_n^{2,\mathrm{LS}})$, respectively. The approximate 95% confidence interval for the skewness parameter ϵ deduced from Table 1 is

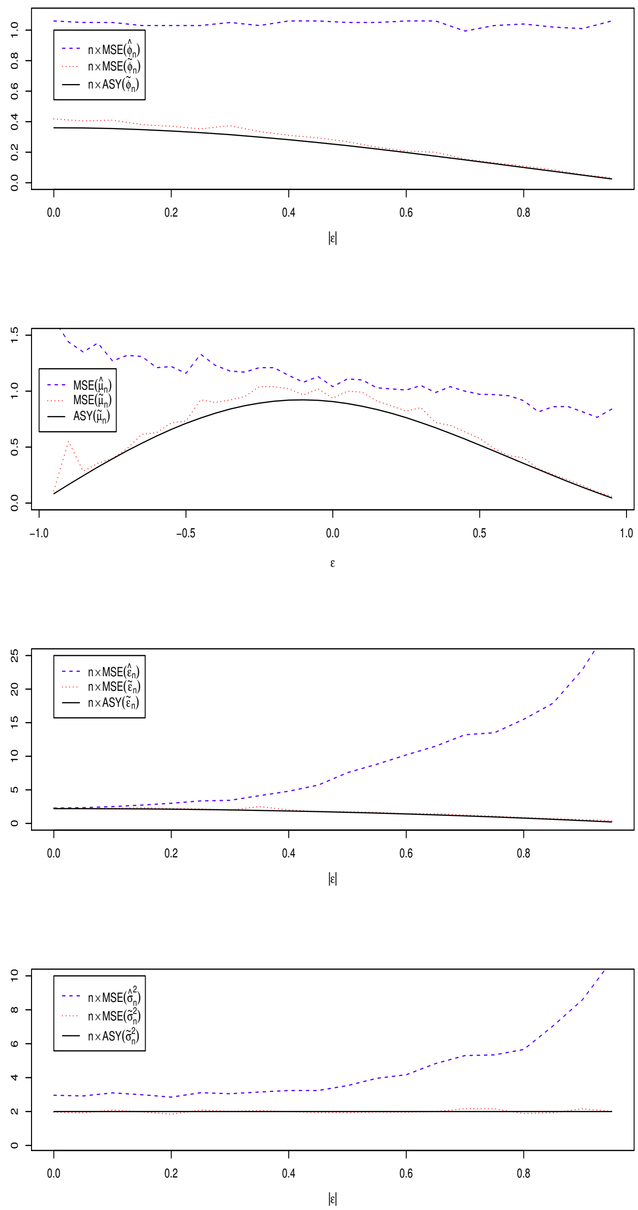

Fig. 1. MSE of $\widehat{\eta}_n$ and $\widetilde{\eta}_n$, and asymptotic variances (ASY) of $\widetilde{\eta}_n$ when $\phi = 0.8$, $\mu = 10$, $\sigma^2 = 1$ and $n = 1000$.

η	ESN		Gaussian	
	Estimate	Variance	Estimate	Variance
ϕ	0.47	5.8e-3	0.49	6.4e-3
μ	-0.11	6.3e-3	0.05	1.2e-3
σ^2	0.12	2.6e-4	0.13	2.8e-4
ϵ	-0.30	1.7e-2	-	-

Table 1. AR(1) models with ESN and Gaussian innovations fitted to the differenced series of the Dow-Jones Utilities index (Jul. 3 - Dec. 20, 1972).

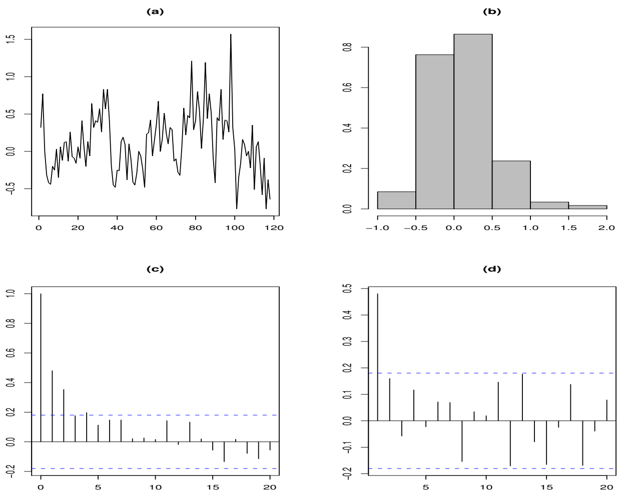

Fig. 2. Differenced series of the Dow-Jones Utilities index (Jul. 3 - Dec. 20, 1972): (a) Series, (b) Histogram, (c) Sample autocorrelation function, (d) Sample partial autocorrelation function.

$(-0.56, -0.04)$. Therefore, we reject at the 5% significance level the hypothesis that ϵ is zero.

Finally, we use the normality tests by Shapiro-Wilk and Jarque-Bera to check the residuals of the Gaussian AR(1) model. The corresponding p-values are $1.9 \cdot 10^{-2}$ and $4.6 \cdot 10^{-4}$. Therefore, both tests reject the null hypothesis of normality at the 95% confidence level.

6. CONCLUSIONS

We have proposed an AR process with ESN innovations to model near-Gaussian asymmetric correlated data. The great flexibility of the ESN distribution allows to model a large class of data with slight leptokurticity.

7. REFERENCES

[1] H. Tong, *Nonlinear Time Series*, vol. 6 of *Oxford Statistical Science Series*, The Clarendon Press Oxford University Press, New York, 1990.

[2] J. Fan and Q. Yao, *Nonlinear time series*, Springer Series in Statistics. Springer-Verlag, New York, 2003, Nonparametric and parametric methods.

[3] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer Verlag, New York, second edition, 1991.

[4] P. A. Jacobs and P. A. W. Lewis, "Stationary discrete autoregressive-moving average time series generated by mixtures," *J. Time Ser. Anal.*, vol. 4, no. 1, pp. 19–36, 1983.

[5] E. McKenzie, "Autoregressive moving-average processes with negative-binomial and geometric marginal distributions," *Adv. in Appl. Probab.*, vol. 18, no. 3, pp. 679–705, 1986.

[6] M. A. Al-Osh and A. A. Alzaid, "First-order integer-valued autoregressive (INAR(1)) process," *J. Time Ser. Anal.*, vol. 8, no. 3, pp. 261–275, 1987.

[7] J. Anděl, "On AR(1) processes with exponential white noise," *Comm. Statist. Theory Methods*, vol. 17, no. 5, pp. 1481–1495, 1988.

[8] B. Nielsen and N. Shephard, "Likelihood analysis of a first-order autoregressive model with exponential innovations," *J. Time Ser. Anal.*, vol. 24, no. 3, pp. 337–344, 2003.

[9] M. Ibazizen and H. Fellag, "Bayesian estimation of an AR(1) process with exponential white noise," *Statistics*, vol. 37, no. 5, pp. 365–372, 2003.

[10] W. K. Li and A. I. McLeod, "ARMA modelling with non-Gaussian innovations," *J. Time Ser. Anal.*, vol. 9, no. 2, pp. 155–168, 1988.

[11] M. L. Tiku, W.-K. Wong, and G. Bian, "Estimating parameters in autoregressive models in non-normal situations: symmetric innovations," *Comm. Statist. Theory Methods*, vol. 28, no. 2, pp. 315–341, 1999.

[12] M. L. Tiku, W.-K. Wong, D. C. Vaughan, and G. Bian, "Time series models in non-normal situations: symmetric innovations," *J. Time Ser. Anal.*, vol. 21, no. 5, pp. 571–596, 2000.

[13] M. L. Tiku, W.-K. Wong, and G. Bian, "Time series models with asymmetric innovations," *Comm. Statist. Theory Methods*, vol. 28, no. 6, pp. 1331–1360, 1999.

[14] A. Akkaya and M. L. Tiku, "Estimating parameters in autoregressive models in non-normal situations: asymmetric innovations," *Comm. Statist. Theory Methods*, vol. 30, no. 3, pp. 517–536, 2001.

[15] W.-K. Wong and G. Bian, "Estimating parameters in autoregressive models with asymmetric innovations," *Statist. Probab. Lett.*, vol. 71, no. 1, pp. 61–70, 2005.

[16] G. J. Janacek and A. L. Swift, "A class of models for nonnormal time series," *J. Time Ser. Anal.*, vol. 11, no. 1, pp. 19–31, 1990.

[17] G. T. Fechner, *Kollectivmasslehre*, Engleman, Leipzig, 1897.

[18] G. S. Mudholkar and A. D. Hutson, "The epsilon-skew-normal distribution for analyzing near-normal data," *J. Statist. Plann. Inference*, vol. 83, no. 2, pp. 291–309, 2000.

[19] A. D. Hutson, "Utilizing the flexibility of the epsilon-skew-normal distribution for common regression problems," *J. Appl. Stat.*, vol. 31, no. 6, pp. 673–683, 2004.

[20] R. B. Arellano-Valle, H. W. Gómez, and F. A. Quintana, "Statistical inference for a general class of asymmetric distributions," *J. Statist. Plann. Inference*, vol. 128, no. 2, pp. 427–443, 2005.

[21] A. Stuart and J. K. Ord, *Kendall's advanced theory of statistics. Vol. 1: Distribution theory*, Edward Arnold, London, sixth edition, 1994.

[22] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer Verlag, New York, 1996.