

Research Article

A Semiparametric Approach for Modeling Partially Linear Autoregressive Model with Skew Normal Innovations

Leila Sakhabakhsh,¹ Rahman Farnoosh ,² Afshin Fallah,³
and Mohammadhassan Behzadi¹

¹Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran

²School of Mathematics, Iran University of Science and Technology, Tehran, Iran

³Faculty of Basic Sciences, Imam Khomeini International University, Qazvin, Iran

Correspondence should be addressed to Rahman Farnoosh; rfarnoosh@iust.ac.ir

Received 23 November 2021; Accepted 24 January 2022; Published 25 February 2022

Academic Editor: Nikos E. Mastorakis

Copyright © 2022 Leila Sakhabakhsh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The nonlinear autoregressive models under normal innovations are commonly used for nonlinear time series analysis in various fields. However, using this class of models for modeling skewed data leads to unreliable results due to the disability of these models for modeling skewness. In this setting, replacing the normality assumption with a more flexible distribution that can accommodate skewness will provide effective results. In this article, we propose a partially linear autoregressive model by considering the skew normal distribution for independent and dependent innovations. A semiparametric approach for estimating the nonlinear part of the regression function is proposed based on the conditional least squares approach and the nonparametric kernel method. Then, the conditional maximum-likelihood approach is used to estimate the unknown parameters through the expectation-maximization (EM) algorithm. Some asymptotic properties for the semiparametric method are established. Finally, the performance of the proposed model is verified through simulation studies and analysis of a real dataset.

1. Introduction

One of the most widely used classes for time series analysis is the class of autoregressive models. The normality of innovations is a common assumption for autoregressive models. However, such an assumption may be unrealistic in many empirical situations. In recent years, more attention has focused on nonnormal innovations rather than normal innovations. Tarami and Porahmadi [1] considered multivariate autoregressive processes with the t distribution for modeling volatile time series data. Jacobs and Lewis [2] analyzed an autoregressive model with nonnormal innovations. Ghasemi et al. [3] considered autoregressive models with generalized hyperbolic innovations. The most important limitation of the normal distribution is that it cannot model skewness. In this article, we consider the skew normal (SN) distribution (Azzalini [4]) for modeling the uncertainty of the innovations in the time series analysis. This skew-

symmetric distribution is a generalization of the normal distribution that enables it to model asymmetric observations as well as the symmetric data. A non-Gaussian autoregressive model with epsilon SN innovations is considered by Bondon [5]. Sharafi and Nematollahi [6] introduced an autoregressive model of order one with SN innovations and proposed some methods for parameters estimation.

Many researchers have recently studied nonlinear autoregressive models in various fields of science. For example, Tsay [7]; Farnoosh and Mortazavi [8]; Hajrajabi and Mortazavi [9]; Farnoosh et al. [10] and Ortega Contreras et al. [11] considered nonlinear time series models and analyzed various datasets. Farnoosh and Mortazavi [8] considered the Gaussian first-order nonlinear autoregressive model with dependent innovations to estimate the yearly amount of deposit in Iran's Tejarat-Bank. Hajrajabi and Mortazavi [9] proposed a nonlinear autoregressive model

with SN innovations and presented asymptotic behaviors of the estimators. Tong [12] and Haggren and Ozaki [13] investigated nonlinear models for modeling sound vibrations. A class of nonlinear additive autoregressive models with exogenous variables to analyze the nonlinear time series is proposed by Chen and Tsay [14]. In addition, the estimation of autoregression function in nonlinear autoregressive models through semiparametric methodology is interested in literature. Fan and Yao [15] discussed modern parametric and nonparametric approaches for estimating nonlinear models. Zhuoxi et al. [16] proposed a semiparametric approach for a nonlinear autoregressive model considering the innovations are normal distribution with mean zero and fixed variance. Hajrajabi and Fallah [17] followed Zhuoxi et al. [16] by assuming the SN innovations. Farnoosh et al. [10] studied a partially linear autoregressive model by considering independent innovations with mean zero and fixed variance. To estimate the regression function, similarly to Zhuoxi et al. [16], they used a semiparametric approach.

This article aims at developing a partially linear autoregressive model with SN innovations for both independent and dependent innovations. This model is an extension of the proposed model by Farnoosh et al. [10]. The estimation of our proposed model consists of two parts. In the first part, we use a semiparametric approach consisting of parametric estimation and nonparametric adjustment introduced by Zhuoxi et al. [16]. For parametric estimation, the conditional least squares approach is used to estimate the model's nonlinear function. Also, the smooth kernel approach is used to estimate the nonparametric adjustment. The second part is to compute the conditional maximum-likelihood (CML) estimators of parameters using the EM algorithm. We also derived the closed iterative forms for the CML estimators of parameters.

The plan of the article proceeds as follows: Section 2 covers the brief properties of the SN distribution. In Section 3, the SN partially linear autoregressive models with independent and dependent innovations are introduced. This section also shows how to estimate the nonlinear part of the models considering the semiparametric approach. In Section

4, the CML estimation of the model parameters via the EM algorithm is discussed. The performance of suggested methods is investigated by simulation in Section 5. A real dataset is also considered in this section to explain the applicability of the proposed models. Finally, conclusions are provided in Section 6. Some asymptotic behaviors of the estimators are given in the Appendix.

2. A Brief Introduction about the SN Distribution

Let Z be a random variable with univariate SN distribution, denoted by $Z \sim SN(\mu, \sigma^2, \lambda)$, where μ , σ^2 , and λ indicate the location, scale, and skewness parameters, respectively. Then, the density function of Z is

$$f_{SN}(Z) = \frac{2}{\sigma} \varphi\left(\frac{z-\mu}{\sigma}\right) \phi\left(\lambda \frac{z-\mu}{\sigma}\right), \quad z \in \mathbb{R}, \quad (1)$$

where $\varphi(\cdot)$ is the density function of the standard normal distribution and $\phi(\cdot)$ is its cumulative distribution function, $\mu \in \mathbb{R}$, $\sigma > 0$, and $\lambda \in \mathbb{R}$.

Lemma 1. *If $Z \sim SN(\mu, \sigma^2, \lambda)$, then*

- (a) $E(Z) = \mu + \sigma\sqrt{2/\pi}\delta(\lambda)$
- (b) $Var(Z) = \sigma^2\{1 - 2/\pi\delta^2(\lambda)\}$
- (c) $SK(Z) = \sqrt{2}(4 - \pi)\lambda^3/\{\pi + (\pi - 2)\lambda^2\}^{1.5}$
- (d) $KU(Z) = 3 + 8(\pi - 2)\lambda^4/\{\pi + (\pi - 2)\lambda^2\}^2$

where $\delta(\lambda) = \lambda/\sqrt{1 + \lambda^2}$, and $SK(Z)$ and $KU(Z)$ are the coefficients of skewness and kurtosis, respectively.

Lemma 2. *From [18] (Theorem 1) and [19] (p. 201), if $U \sim TN(0, \sigma^2)I_{\{u>0\}}$, $V \sim N(0, \sigma^2)$ and U and V be independent, then*

$$Z = \mu + \delta(\lambda)U + V\sqrt{1 - \delta^2(\lambda)}, \quad (2)$$

distributed as $SN(\mu, \sigma^2, \lambda)$. Also, the joint density of Z and U is

$$f_{Z,U}(z, u) = \frac{\sqrt{1 + \lambda^2}}{\pi\sigma^2} \exp\left\{\frac{-1}{2\sigma^2} \left[(z - \mu)^2 + (1 + \lambda)^2 \left(u - \frac{\lambda}{\sqrt{1 + \lambda^2}} (z - \mu) \right)^2 \right]\right\}. \quad (3)$$

Lemma 3. *Suppose $U \sim TN(0, \sigma^2)I_{\{u>0\}}$, $V \sim N(0, \sigma^2)$, and Z is defined as $Z = \mu + \delta(\lambda)U + V\sqrt{1 - \delta^2(\lambda)}$, then*

$$U|Z = z \sim TN(\mu_U, \sigma_U^2)I_{a_1=0 < u < a_2=\infty}, \quad (4)$$

and also

$$E(U|Z = z) = \mu_U + \frac{\varphi(\mu_U/\sigma_U)}{\phi(\mu_U/\sigma_U)}\sigma_U, \quad (5)$$

$$E(U^2|Z = z) = \mu_U^2 + \sigma_U^2 + \frac{\varphi(\mu_U/\sigma_U)}{\phi(\mu_U/\sigma_U)}\sigma_U\mu_U.$$

3. Semiparametric Approach in the Proposed Model

In this section, we consider partially linear autoregressive models of the following forms.

3.1. *Model with Independent Innovations (Model I).* Consider the following model:

$$z_t = \gamma z_{t-1} + f(z_{t-2}) + \nu_t, \quad t = 2, 3, \dots, n, |\gamma| < 1, \quad (6)$$

where $\nu_t \sim SN(\mu, \sigma^2, \lambda)$, $f(\cdot)$ is a nonlinear autoregressive function, and γ is an unknown parameter. Also, ν_t and y_t are independent for each t .

At first, we estimate $f(x)$ by an initial guess $r(x, \beta)$ as a known function of x and β . The parameter β can be estimated by using the conditional nonlinear least squares errors (CNLSE) method based on data z_0, z_1, \dots, z_n as follows:

$$\begin{aligned} \hat{\beta} &= \arg \min_{\beta \in \Theta} \sum_{t=2}^n (z_t - E(z_t | z_{t-1}, z_{t-2}))^2, \\ &= \arg \min_{\beta \in \Theta} \sum_{t=2}^n (z_t - \gamma z_{t-1} - r(z_{t-2}, \beta) - \mu_{\nu_t})^2, \end{aligned} \quad (7)$$

where $\mu_{\nu_t} = \mu + \sigma\sqrt{2/\pi}\delta(\lambda)$ (see Lemma 1).

We use the semiparametric form of $r(x, \beta)\xi(x)$ to adjust the initial approximation, where $\xi(x)$ shows a nonparametric adjustment.

For estimating $\xi(x)$, we use the following local L2-fitting criterion:

$$q_n(x, \xi) = \frac{1}{h_n} \sum_{t=2}^n k\left(\frac{z_{t-2} - x}{h_n}\right) (f(z_{t-2}) - r(z_{t-2}, \hat{\beta})\xi)^2, \quad (8)$$

where $k(\cdot)$ and h_n are the kernel function and bandwidth, respectively. We get the estimator $\hat{\xi}(x)$ of $\xi(x)$ by minimizing (2) with respect to $\xi(x)$ as follows:

$$\hat{\xi}(x) = \frac{\sum_{t=2}^n [K(z_{t-2} - x/h_n)r(z_{t-2}, \hat{\beta})f(z_{t-2})]}{\sum_{t=2}^n K(z_{t-2} - x/h_n)r^2(z_{t-2}, \hat{\beta})}, \quad (9)$$

and the estimator of autoregression function of the model is

$$\hat{f}(x) = r(x, \hat{\beta})\hat{\xi}(x). \quad (10)$$

However, the function $f(x)$ in the formula $\hat{\xi}(x)$ in equation (9) is unknown. Therefore, by using

$$\begin{aligned} \tilde{\nu}_t &= \nu_t - \mu_{\nu_t}, \\ &= z_t - \gamma z_{t-1} - f(z_{t-2}) - \mu_{\nu_t}, \end{aligned} \quad (11)$$

and regarding the fact that $\tilde{\nu}_t$ are small values, one can obtain

$$\tilde{\xi}(x) = \frac{\sum_{t=2}^n [K(z_{t-2} - x/h_n)r(z_{t-2}, \hat{\beta})(z_t - \hat{\gamma}z_{t-1} - \hat{\mu}_{\nu_t})]}{\sum_{t=2}^n [K(z_{t-2} - x/h_n)r^2(z_{t-2}, \hat{\beta})]}. \quad (12)$$

Finally, the estimator of $f(x)$ is

$$\tilde{f}(x) = r(x, \hat{\beta})\tilde{\xi}(x). \quad (13)$$

3.2. Model with Dependent Innovations (Model II). By considering the partially linear autoregressive model in equation (6) with dependent innovations as first-order autoregressive AR(1), we have

$$\begin{aligned} z_t &= \gamma z_{t-1} + f(z_{t-2}) + \varepsilon_t, \\ \varepsilon_t &= \rho \varepsilon_{t-1} + \nu_t, \quad |\rho| < 1, \end{aligned} \quad (14)$$

where $\nu_t \sim SN(\mu, \sigma^2, \lambda)$ and $f(\cdot)$ is a nonlinear autoregression function similar to Model (I). Also, ν_t and y_t are independent for each t .

From (14), we can write

$$\varepsilon_t = z_t - \gamma z_{t-1} - f(z_{t-2}) \longrightarrow \varepsilon_{t-1} = z_{t-1} - \gamma z_{t-2} - f(z_{t-3}). \quad (15)$$

And, therefore,

$$z_t = \gamma z_{t-1} + f(z_{t-2}) + \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) + \nu_t. \quad (16)$$

We want to estimate the unknown regression functions $f(z_{t-2})$ and $f(z_{t-3})$ that can be formed as $r(z_{t-2}, \beta)$ and $r(z_{t-3}, \beta)$, respectively. As it is shown in Section 3.1, to estimate β , we can apply the CNLSE approach as follows:

$$\hat{\beta} = \arg \min_{\beta \in \Theta} \sum_{t=2}^n (z_t - \gamma z_{t-1} - r(z_{t-2}, \beta) - \rho(z_{t-1} - \gamma z_{t-2} - r(z_{t-3}, \beta) - \mu_{\nu_t}))^2, \quad (17)$$

where $\mu_{\nu_t} = \mu + \sigma\sqrt{2/\pi}\delta(\lambda)$ (see Lemma 1).

By applying the same idea as in Model (I), the local L2-fitting criterion is

$$q_n(x, \xi) = \frac{1}{h_n} \sum_{t=3}^n k\left(\frac{z_{t-2} - x}{h_n}\right) (f(z_{t-2}) - r(z_{t-2}, \hat{\beta})\xi)^2 + \frac{1}{h_n} \sum_{t=3}^n k\left(\frac{z_{t-3} - x}{h_n}\right) (f(z_{t-3}) - r(z_{t-3}, \hat{\beta})\xi)^2. \quad (18)$$

By minimizing (18) with respect to $\xi(x)$, one can write

$$\widehat{\xi}(x) = \frac{\sum_{t=3}^n [K(z_{t-2} - x/h_n)r(z_{t-2}, \widehat{\beta})f(z_{t-2}) + K(z_{t-3} - x/h_n)r(z_{t-3}, \widehat{\beta})f(z_{t-3})]}{\sum_{t=3}^n [K(z_{t-2} - x/h_n)r^2(z_{t-2}, \widehat{\beta}) + K(z_{t-3} - x/h_n)r^2(z_{t-3}, \widehat{\beta})]}. \quad (19)$$

Unfortunately, equation (19) includes the unknown function $f(\cdot)$. Therefore, by using

$$\begin{aligned} \widetilde{\varepsilon}_t &= \varepsilon_t - \mu_{\varepsilon_t} = z_t - \gamma z_{t-1} - f(z_{t-2}) - \mu_{\varepsilon_t}, \\ \widetilde{\varepsilon}_{t-1} &= \varepsilon_{t-1} - \mu_{\varepsilon_{t-1}} = z_{t-1} - \gamma z_{t-2} - f(z_{t-3}) - \mu_{\varepsilon_{t-1}}, \end{aligned} \quad (20)$$

and regarding the fact that $\widetilde{\varepsilon}_t$ are small values, we have

$$\widetilde{\xi}(x) = \frac{\sum_{t=3}^n [K(z_{t-2} - x/h_n)r(z_{t-2}, \widehat{\beta})(z_t - \widehat{\gamma}z_{t-1} - \widehat{\mu}_{\varepsilon_t}) + K(z_{t-3} - x/h_n)r(z_{t-3}, \widehat{\beta})(z_{t-1} - \widehat{\gamma}z_{t-2} - \widehat{\mu}_{\varepsilon_{t-1}})]}{\sum_{t=3}^n [K(z_{t-2} - x/h_n)r^2(z_{t-2}, \widehat{\beta}) + K(z_{t-3} - x/h_n)r^2(z_{t-3}, \widehat{\beta})]}. \quad (21)$$

Finally, the estimator of $f(x)$ is obtained as

$$\widetilde{f}(x) = r(x, \widehat{\beta})\widetilde{\xi}(x). \quad (22)$$

4. Conditional Maximum-Likelihood Estimation

There are several methods for estimating the parameters in time series models. This article implements the CML estimation method. The conditional likelihood function of the Model (II) given the observed data are defined by

$$\begin{aligned} L(\boldsymbol{\theta}|\mathbf{z}) &= \prod_{t=3}^n f_{SN}(z_t | z_{t-2}, z_{t-3}, \boldsymbol{\theta}), \\ &= \prod_{t=3}^n \frac{2}{\sigma} \phi\left(\frac{z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu}{\sigma}\right) \\ &\quad \times \Phi\left(\lambda\left(\frac{z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu}{\sigma}\right)\right) \\ &= \left(\frac{2}{\sigma}\right)^{n-2} \exp\left\{-\frac{1}{2} \sum_{t=3}^n (z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu)^2\right\} \\ &\quad \times \prod_{t=3}^n \Phi\left(\lambda\left(\frac{z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu}{\sigma}\right)\right). \end{aligned} \quad (23)$$

where $f(z_{t-2}) = r(z_{t-2}, \beta)\xi(z_{t-2})$, $f(z_{t-3}) = r(z_{t-3}, \beta)\xi(z_{t-3})$ and $\theta = (\beta, \mu, \sigma^2, \lambda, \gamma, \rho)$ is the unknown parameters vector. Since the likelihood function in equation (23) is complicated, we need a computational approach to maximize it.

Therefore, an EM algorithm is developed to calculate the ML estimate of the parameters. To do this, we consider the missing data problem. By defining the variables $U_t \sim TN(0, \sigma^2)I_{\{u_t > 0\}}$, $V_t \sim N(0, \sigma^2)$, and

$$z_t = \gamma z_{t-1} + f(z_{t-2}) + \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) + \mu + \delta(\lambda)U_t + \sqrt{1 - \delta^2}(\lambda)V_t, \quad (24)$$

and using Lemma 2, the conditional distribution of observation in the Model (II) is given by

$$z_t | z_{t-1}, z_{t-2}, z_{t-3} \sim SN\left((\gamma z_{t-1} + f(z_{t-2}) + \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) + \mu), \sigma^2, \lambda\right), \quad t = 3, \dots, n. \quad (25)$$

Let \mathbf{z} and $\{u_t; t = 3, \dots, n\}$ be the incomplete and missing data, respectively, and using Lemma 2, the joint density function of the complete data can be written as

$$\begin{aligned} f_{Z_t, U_t}(z_t, u_t) &= \frac{\sqrt{1+\lambda^2}}{\pi\sigma^2} \exp\left\{\frac{-1}{2\sigma^2} \left[(z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu)^2 \right. \right. \\ &\quad \left. \left. + (1+\lambda)^2 \left(u_t - \frac{\lambda}{\sqrt{1+\lambda^2}} (z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu) \right)^2 \right]\right\} \\ &= \frac{1}{\pi\sigma^2 \sqrt{1-\delta^2(\lambda)}} \exp\left\{\frac{-1}{2\sigma^2(1-\delta^2(\lambda))} \times \left[(z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu)^2 \right. \right. \\ &\quad \left. \left. - 2u_t\delta(\lambda)(z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu) + u_t^2 \right]\right\}. \end{aligned} \quad (26)$$

Therefore, the complete data likelihood and log likelihood functions are, respectively,

$$\begin{aligned} L_C(\theta | \mathbf{z}, \mathbf{u}) &= \prod_{t=3}^n f_{Z_t, U_t}(z_t, u_t) = (\pi\sigma^2)^{-(n-2)} (1-\delta^2(\lambda))^{-n-2/2} \times \exp\left\{\frac{-1}{2\sigma^2(1-\delta^2(\lambda))} \right. \\ &\quad \left. \cdot \sum_{t=3}^n \left[\begin{aligned} &(z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho z_{t-1} - \gamma z_{t-2} - f(z_{t-3}) - \mu)^2 \\ &- 2u_t\delta(\lambda)(z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu) + u_t^2 \end{aligned} \right]\right\}, \\ l_C(\theta | \mathbf{z}, \mathbf{u}) &= -(n-2)\log(\sigma^2) - \frac{n-2}{2}\log(1-\delta^2(\lambda)) - \frac{1}{2\sigma^2(1-\delta^2(\lambda))} \\ &\quad \times \left[\sum_{t=3}^n (z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu)^2 \right. \\ &\quad \left. - 2\delta(\lambda) \sum_{t=3}^n u_t (z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu) + \sum_{t=3}^n u_t^2 \right]. \end{aligned} \quad (27)$$

The EM algorithm iterates between *E* and *M* steps. The *E* step obtains the conditional expectation of $l_C(\theta | \mathbf{z}, \mathbf{u})$ given the observed data and current parameters. Based on Lemma 3, we have

$$U_t | Z_t = z_t \sim TN(\mu_{U_t}, \sigma_U^2) I_{u_t > 0}, \quad (28)$$

where

$$\begin{aligned} \mu_{U_t} &= \delta(\lambda)(z_t - \gamma z_{t-1} - f(z_{t-2}) - \rho(z_{t-1} - \gamma z_{t-2} - f(z_{t-3})) - \mu), \\ \sigma_U &= \sigma \sqrt{1-\delta^2(\lambda)}. \end{aligned} \quad (29)$$

Calculating the conditional expectation of (27) yields

$$\begin{aligned}
E[l_c(\hat{\theta}|\mathbf{z}, \mathbf{u})|z] &= -(n-2)\log(\hat{\sigma}^2) - \frac{(n-2)}{2}\log(1 - \delta^2(\hat{\lambda})) \\
&\quad - \frac{1}{2\hat{\sigma}^2(1 - \delta^2(\hat{\lambda}))} \left[\sum_{t=3}^n (z_t - \hat{\gamma}z_{t-1} - \tilde{f}(z_{t-2}) - \hat{\rho}z_{t-1} - \hat{\gamma}z_{t-2} - \tilde{f}(z_{t-3}) - \hat{\mu})^2 \right. \\
&\quad \left. - 2\delta(\hat{\lambda}) \sum_{t=3}^n a_{1t}(z_t - \hat{\gamma}z_{t-1} - \tilde{f}(z_{t-2}) - \hat{\rho}(z_{t-1} - \hat{\gamma}z_{t-2} - \tilde{f}(z_{t-3})) - \hat{\mu}) + \sum_{t=3}^n a_{2t} \right], \tag{30}
\end{aligned}$$

where

$$\begin{aligned}
a_{1t} &= E_{\hat{\theta}}[U_t|Z_{t-2}, Z_{t-3} = z_{t-2}, z_{t-3}] \\
&= \hat{\mu}_{U_t} + \frac{\phi(\hat{\mu}_{U_t}/\hat{\sigma}_U)}{\Phi(\hat{\mu}_{U_t}/\hat{\sigma}_U)}\hat{\sigma}_U, \\
a_{2t} &= E_{\hat{\theta}}[U_t^2|Z_{t-2}, Z_{t-3} = z_{t-2}, z_{t-3}] \\
&= \hat{\mu}_{U_t}^2 + \hat{\sigma}_U^2 + \frac{\phi(\hat{\mu}_{U_t}/\hat{\sigma}_U)}{\Phi(\hat{\mu}_{U_t}/\hat{\sigma}_U)}\hat{\sigma}_U\hat{\mu}_{U_t}, \tag{31}
\end{aligned}$$

with

$$\begin{aligned}
\hat{\mu}_{U_t} &= \delta(\hat{\lambda})(z_t - \hat{\gamma}z_{t-1} - \tilde{f}(z_{t-2}) \\
&\quad - \hat{\rho}(z_{t-1} - \hat{\gamma}z_{t-2} - \tilde{f}(z_{t-3})) - \hat{\mu}), \tag{32} \\
\hat{\sigma}_U &= \hat{\sigma}\sqrt{(1 - \delta^2(\hat{\lambda}))},
\end{aligned}$$

where $\tilde{f}(z_{t-2}) = r(z_{t-2}, \hat{\beta})\xi(z_{t-2})$ and $\tilde{f}(z_{t-3}) = r(z_{t-3}, \hat{\beta})\xi(z_{t-3})$ are given in (22).

The M step of the algorithm maximizes the expectation computed in (30).

Given the values of the parameters in k th iteration, equating the first-order derivatives of (30) to zero and solving the resulted system of equations, the maximum-likelihood estimates of model parameters in $k+1$ th iteration of the algorithm are obtained to be

$$\begin{aligned}
\hat{\mu}^{(k+1)} &= \frac{1}{n-2} \left[\sum_{t=3}^n (z_t - \hat{\gamma}^{(k)}z_{t-1} - \tilde{f}^{(k)}(z_{t-2}) - \hat{\rho}^{(k)}(z_{t-1} - \hat{\gamma}^{(k)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3})) \right] - \delta(\hat{\lambda}^{(k)}) \sum_{t=3}^n a_{1t}^{(k)}, \\
\hat{\sigma}^{2(k+1)} &= \frac{1}{2(n-2)(1 - \delta^2(\hat{\lambda}^{(k)}))} \left[\sum_{t=3}^n a_{2t}^{(k)} - 2\delta(\hat{\lambda}^{(k)}) \right. \\
&\quad \cdot \sum_{t=3}^n a_{1t}^{(k)} (z_t - \hat{\gamma}^{(k)}z_{t-1} - \tilde{f}^{(k)}(z_{t-2}) - \hat{\rho}^{(k)}(z_{t-1} - \hat{\gamma}^{(k)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3})) - \hat{\mu}^{(k+1)}) \\
&\quad \left. + \sum_{t=3}^n (z_t - \hat{\gamma}^{(k)}z_{t-1} - \tilde{f}^{(k)}(z_{t-2}) - \hat{\rho}^{(k)}(z_{t-1} - \hat{\gamma}^{(k)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3})) - \hat{\mu}^{(k+1)})^2 \right], \\
\hat{\gamma}^{(k+1)} &= \frac{\sum_{t=3}^n (z_{t-1} - \hat{\rho}^{(k)}z_{t-2})(z_t - \tilde{f}^{(k)}(z_{t-2}) - \hat{\rho}^{(k)}(z_{t-1} - \tilde{f}^{(k)}(z_{t-3})) - \hat{\mu}^{(k+1)})}{\sum_{t=3}^n (z_{t-1} - \hat{\rho}^{(k)}z_{t-2})^2} - \frac{\delta(\hat{\lambda}^{(k)}) \sum_{t=3}^n a_{1t}^{(k)} (z_{t-1} - \hat{\rho}^{(k)}z_{t-2})}{\sum_{t=3}^n (z_{t-1} - \hat{\rho}^{(k)}z_{t-2})^2}, \\
\hat{\rho}^{(k+1)} &= \frac{\sum_{t=3}^n (z_{t-1} - \hat{\gamma}^{(k+1)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3}))(z_t - \hat{\gamma}^{(k+1)}z_{t-1} - \tilde{f}^{(k)}(z_{t-2}) - \hat{\mu}^{(k+1)})}{\sum_{t=3}^n (z_{t-1} - \hat{\gamma}^{(k+1)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3}))^2} - \frac{\delta(\hat{\lambda}^{(k)}) \sum_{t=3}^n a_{1t}^{(k)} (z_{t-1} - \hat{\gamma}^{(k+1)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3}))}{\sum_{t=3}^n (z_{t-1} - \hat{\gamma}^{(k+1)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3}))^2}, \\
\hat{\lambda}^{(k+1)} &= \arg \max_{\lambda} \sum_{t=3}^n \Phi \left(\lambda \left(\frac{z_t - \hat{\gamma}^{(k+1)}z_{t-1} - \tilde{f}^{(k)}(z_{t-2}) - \hat{\rho}^{(k+1)}(z_{t-1} - \hat{\gamma}^{(k+1)}z_{t-2} - \tilde{f}^{(k)}(z_{t-3})) - \hat{\mu}^{(k+1)}}{\hat{\sigma}^{(k+1)}} \right) \right). \tag{33}
\end{aligned}$$

Then, updating the parameter β from equation (17) yields

$$\widehat{\beta}^{(k+1)} = \arg \min_{\beta \in \Theta} \sum_{t=3}^n \left(z_t - \widehat{\gamma}^{(k+1)} z_{t-1} - r(z_{t-2}, \beta) - \widehat{\rho}^{(k+1)}(z_{t-1} - \widehat{\gamma}^{(k+1)} z_{t-2} - r(z_{t-3}, \beta)) - \widehat{\mu}_{\nu_t}^{(k+1)} \right)^2, \quad (34)$$

where

$$\widehat{\mu}_{\nu_t}^{(k+1)} = \widehat{\mu}^{(k+1)} + \widehat{\sigma}^{(k+1)} \sqrt{\frac{2}{\pi}} \delta \left(\widehat{\lambda}^{(k+1)} \right), \quad (35)$$

Therefore, the nonparametric estimator of $\xi(x)$ is updated as follows:

$$\begin{aligned} \widetilde{\xi}^{(k+1)}(x) = & \frac{\sum_{t=3}^n K(z_{t-2} - x/h_n) r(z_{t-2}, \widehat{\beta}^{(k+1)}) (z_t - \widehat{\gamma}^{(k+1)} z_{t-1} - \widehat{\mu}_{\varepsilon_t}^{(k+1)})}{\sum_{t=3}^n \left[K(z_{t-2} - x/h_n) r^2(z_{t-2}, \widehat{\beta}^{(k+1)}) + K(z_{t-3} - x/h_n) r^2(z_{t-3}, \widehat{\beta}^{(k+1)}) \right]} \\ & + \frac{\sum_{t=3}^n K(z_{t-3} - x/h_n) r(z_{t-3}, \widehat{\beta}^{(k+1)}) (z_{t-1} - \widehat{\gamma}^{(k+1)} z_{t-2} - \widehat{\mu}_{\varepsilon_{t-1}}^{(k+1)})}{\sum_{t=3}^n \left[K(z_{t-2} - x/h_n) r^2(z_{t-2}, \widehat{\beta}^{(k+1)}) + K(z_{t-3} - x/h_n) r^2(z_{t-3}, \widehat{\beta}^{(k+1)}) \right]}. \end{aligned} \quad (36)$$

Finally, the semiparametric estimation of autoregression function in the $k + 1$ th iteration is

$$\begin{aligned} \widetilde{f}^{(k+1)}(z_{t-2}) &= r(z_{t-2}, \widehat{\beta}^{(k+1)}) \widetilde{\xi}^{(k+1)}(z_{t-2}), \\ \widetilde{f}^{(k+1)}(z_{t-3}) &= r(z_{t-3}, \widehat{\beta}^{(k+1)}) \widetilde{\xi}^{(k+1)}(z_{t-3}). \end{aligned} \quad (37)$$

As it can be seen, there is no closed iterative form for the ML estimator of parameter λ . Therefore, an iterative procedure similar to Newton–Raphson should be employed to calculate the corresponding value for this parameter.

5. Simulation Study

5.1. Simulation Study 1. We present a simulation study to examine the performance of the suggested methods by using the R programming environment. Two partially linear autoregressive models are considered with the different situations of independent and dependent innovations, respectively, as

$$\begin{aligned} \text{Model (I): } z_t &= \gamma z_{t-1} + f(z_{t-2}) + \nu_t, \\ \text{Model (II): } z_t &= \gamma z_{t-1} + f(z_{t-2}) + \varepsilon_t, \quad \varepsilon_t = \rho \varepsilon_{t-1} + \nu_t, \end{aligned} \quad (38)$$

where $\nu_t \sim SN(0.1, 0.25, 0.85)$. The data were simulated from Model (II) in (14) with sample sizes $n = 100, 200, 400$, and 500 iterations considering the nonlinear function $f(x) = 5e^{-x^2}$. We assume $r(x, \beta) = \beta e^{-x^2}$. Also, for choosing the bandwidth h_n , we perform an opening window method, that is, considering several bandwidths [20].

Table 1 presents some descriptive statistics for the simulated data from Model (II). The Kolmogorov–Smirnov

(K–S) test is also provided. The test statistics and p -value reject the normality of the datasets.

The values of $f(z_t)$ and their semiparametric estimates from Model (II) are displayed in Figure 1(a) with selected bandwidth. This figure presents that the semiparametric estimator of the autoregression function with AR(1) innovation performs well.

As can be seen from Figure 1(b), the estimated values of data z_t are close to the exact values of data. Figure 1(c) presents the autocorrelation function (ACF) of the residuals of the model with AR(1) innovations. This figure illustrates that the residuals of the model are uncorrelated.

Finally, we calculate the root of mean squared error (RMSE) for comparing the efficiency and accuracy of the suggested partially linear autoregressive models as

$$\text{RMSE} = \sqrt{\frac{1}{n-2} (f(z_t) - \widetilde{f}(z_t))^2}. \quad (39)$$

Table 2 reports the RMSE of Model (I) and Model (II) with three values of $h_n \in \{0.02, 0.04, 0.06\}$ and different sample sizes $n \in \{100, 200, 400\}$. As we can see from Table 2, the simulation results show that the model with AR(1) innovations has more performance in comparison to the model with independent innovations. Also, the results show that the RMSE decreases when the sample size increases for all sample sizes n and all values of h_n .

Also, the values of $f(z_t)$ and their semiparametric estimates under two above models with selected bandwidth are displayed in Figure 2(a). The solid line is the regression function $f(z_t)$, and the blue line and red line are the semiparametric fitted function of Models I and II, respectively.

TABLE 1: Descriptive statistics for the simulated data from Model (II).

Data set	n	Min	Max	Mean	Standard deviation	Skewness	K-S (statistics)	K-S (p-value)
1	100	0.4143	5.5	2.0569	1.2471	1.1899	0.7159	<2.2e-16
2	200	0.1489	5.5541	1.9708	1.1528	1.1981	0.7187	<2.2e-16
3	400	0.2380	5.9470	1.9462	1.1483	1.1851	0.7041	<2.2e-16

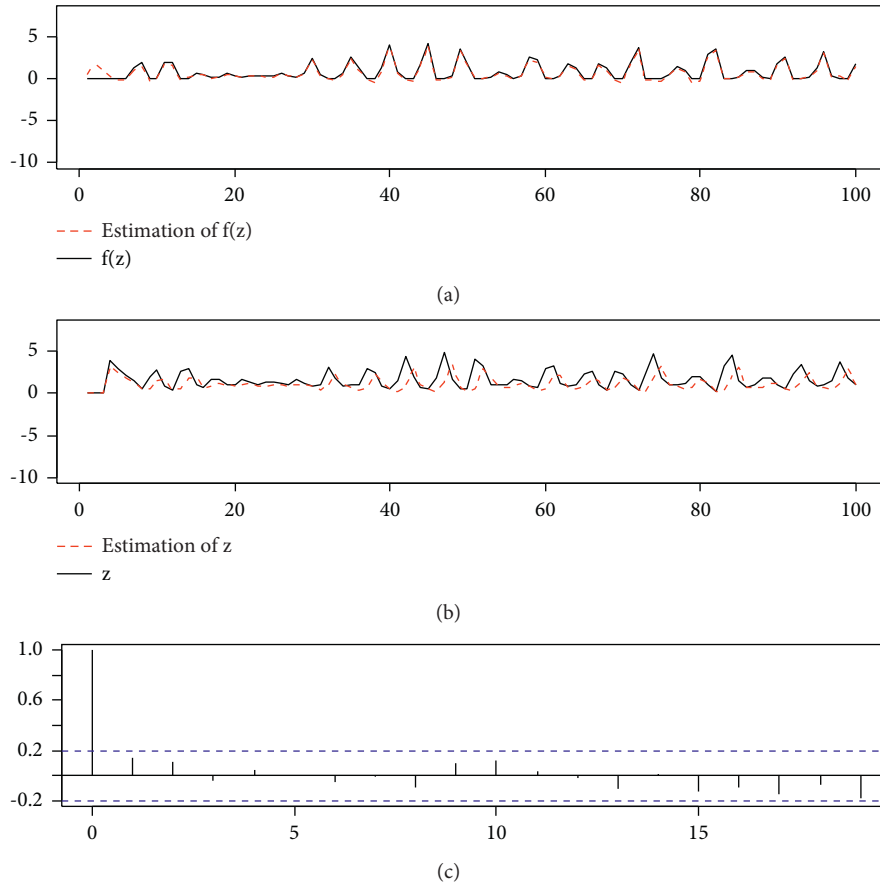
FIGURE 1: (a) Exact and estimated values of $f(z_t)$. (b) Exact and estimated values of Z . (c) ACF of residuals, $h_n = 0.06, n = 100$.

TABLE 2: Values of the RMSE for Model I and Model II, with AR(1) innovations and 500 iterations.

n	h_n	RMSE (model II)	RMSE (model I)
100	0.02	0.5198	1.4865
	0.04	0.6305	1.7414
	0.06	0.6824	2.0679
200	0.02	0.4723	1.4131
	0.04	0.5563	1.7014
	0.06	0.6674	1.9159
400	0.02	0.3916	1.3632
	0.04	0.5463	1.6686
	0.06	0.5593	1.8909

Figure 2(b) presents the exact and estimated values of data under Models I and II. The solid line corresponds to exact values of z_t , and the blue line and red line correspond to semiparametric estimates of z_t from Models I and II, respectively. We can see that the estimation results of Model II are very close.

5.2. Simulation Study 2. This section presents a simulation study to compare the performance of the proposed model with normal model in modeling time series data with SN innovations. We consider a partially linear autoregressive model of the following form:

$$z_t = \gamma z_{t-1} + f(z_{t-2}) + v_t, \quad t = 3, \dots, n, \quad (40)$$

where $v_t \sim SN(1, 1, \lambda)$ and chose two different functions $f_1(x) = 5e^{-x^2}$ by assuming $r(x, \beta_1) = \beta_1 e^{-x^2}$, $f_2(x) = 0.7 \sin x + 0.1x$ by assuming $r(x, \beta_2) = \beta_2 \sin x$.

The datasets are generated with sample sizes $n = 50, 100, 200$ and different values of $\lambda = \{-2, -1, 0, 1, 2\}$. Also, the bandwidth h_n is chosen by an opening the window technique [20].

The values of RMSE of the model in equation (40) under normal and SN innovations are presented in Tables 3 and 4.

The RMSE results show the SN model has a better performance than the normal model for positive and negative skewness parameters. Also, there are no significant

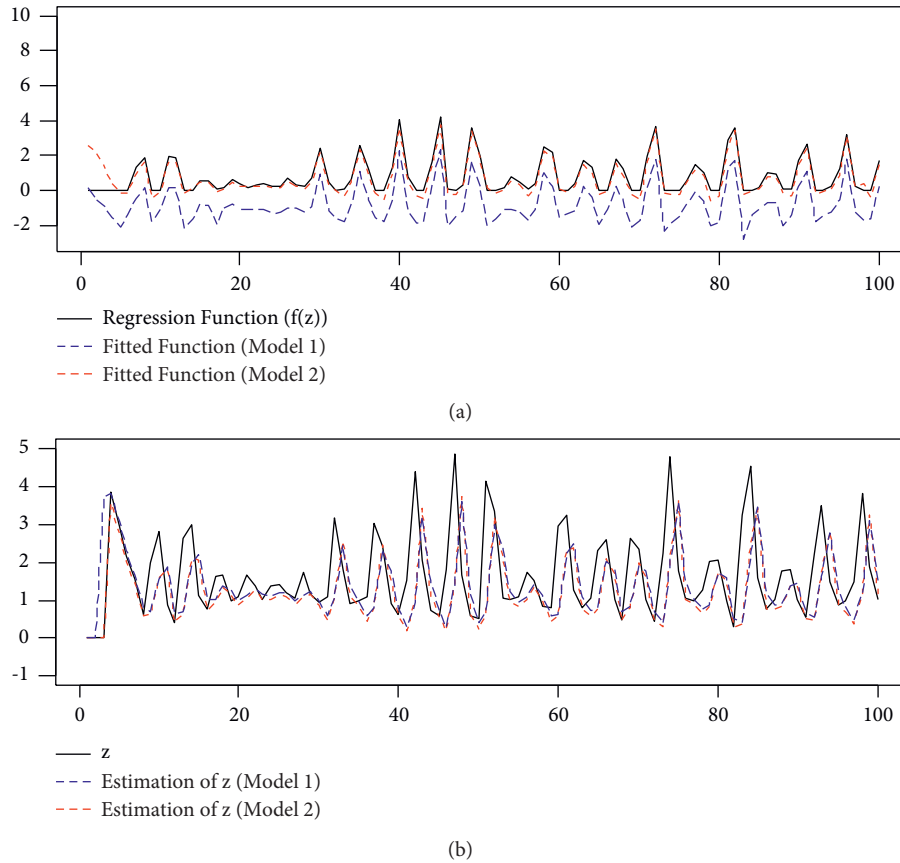


FIGURE 2: (a) Exact and estimated values of $f(z_t)$ from Model I and Model II. (b) Exact and estimated values of data from Model I and Model II, $h_n = 0.04$, $n = 100$.

TABLE 3: RMSE for different values of λ under the normal and SN innovations, $f_1(x)$.

n	λ	Model	
		Skew normal	Normal
50	-2	2.1273	2.2415
	-1	2.2927	2.4291
	0	2.3441	2.3605
	1	1.8298	2.21151
	2	1.2074	1.8185
100	-2	1.9364	2.0192
	-1	2.0985	2.2491
	0	2.3092	2.3405
	1	1.3833	1.9486
	2	1.1510	1.8178
200	-2	1.9129	2.0187
	-1	2.0211	2.1326
	0	2.1266	2.1683
	1	1.3339	1.8965
	2	1.1431	1.7354

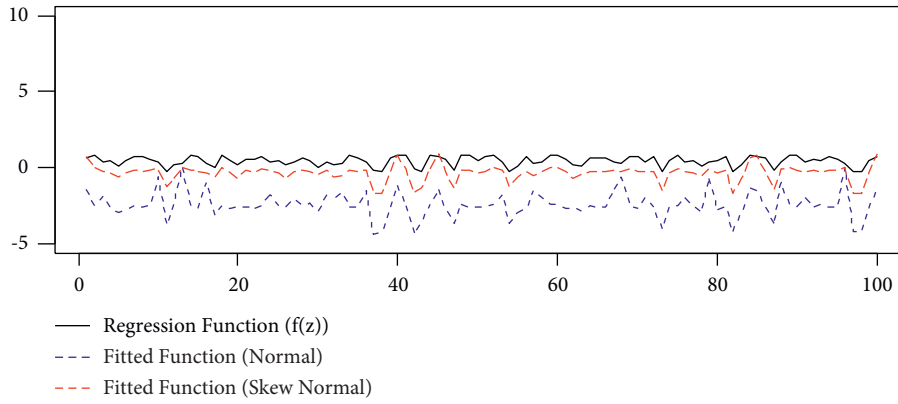
TABLE 4: RMSE for different values of λ under the normal and SN innovations, $f_2(x)$.

n	λ	Model	
		Skew normal	Normal
50	-2	0.9397	1.0150
	-1	0.9716	1.0148
	0	1.5302	1.5306
	1	1.2725	1.7291
	2	1.2526	1.7133
100	-2	0.9239	0.9821
	-1	0.9441	1.009
	0	1.5051	1.5177
	1	1.2598	1.7216
	2	1.1280	1.7044
200	-2	0.8601	0.9152
	-1	0.9092	0.9500
	0	1.4758	1.4871
	1	1.1902	1.6908
	2	1.0654	1.6690

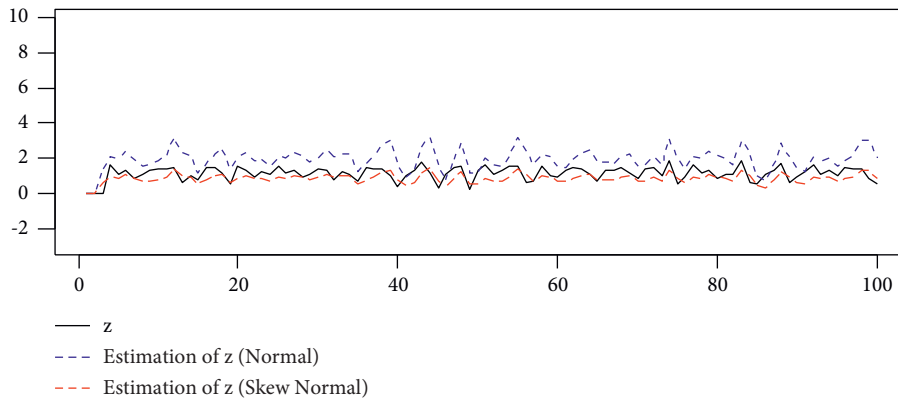
differences between the RMSE values of normal and SN models when $\lambda = 0$.

Figures 3(a) and 4(a) show the values of $f(z_t)$ and its semiparametric estimator under the normal and SN model with

selected bandwidth $h_n = 0.12$, $\lambda = 2$ and sample size 100 for two functions $f_1(x)$ and $f_2(x)$. Also, the exacted and estimated values of data are shown in Figures 3(b) and 4(b). The figures show that the SN model performs better than the normal model.

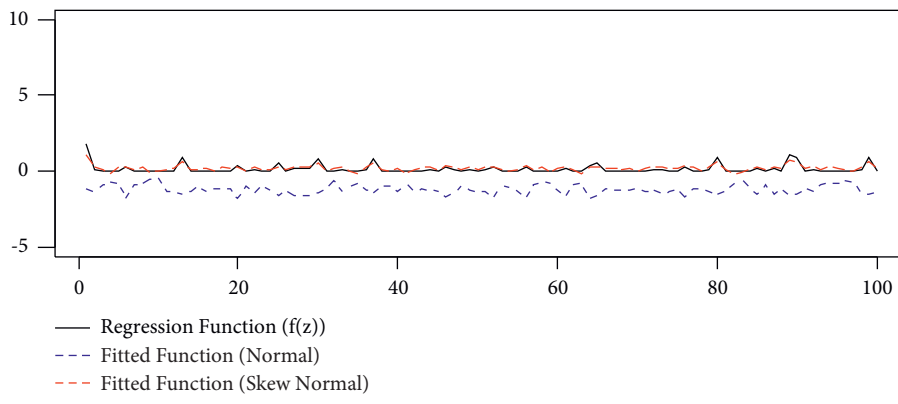


(a)



(b)

FIGURE 3: (a) Exact and estimated values of $f(z_t)$ from normal and SN models. (b) Exact and estimated values of data from normal and SN models, $h_n = 0.12, n = 100, \lambda = 2, f_1(x) = 5e^{-x^2}$.



(a)

FIGURE 4: Continued.

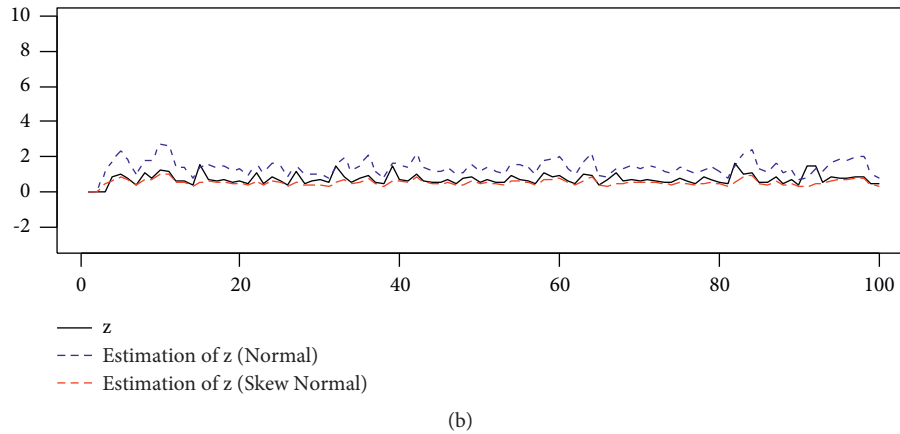


FIGURE 4: (a) Exact and estimated values of $f(z_t)$ from normal and SN models. (b) Exact and estimated values of data from normal and SN models, $h_n = 0.12$, $n = 100$, $\lambda = 2$, $f_2(x) = 0.7 \sin x + 0.1x$.

TABLE 5: The descriptive statistics of the values for ARW data, 1974–2008.

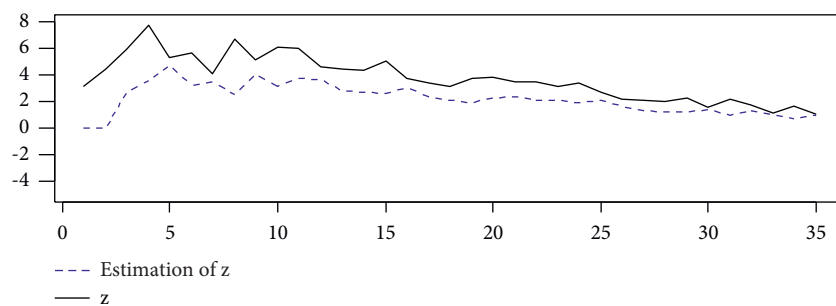
n	Mean	Min	Max	Variance	Skewness	Sum
35	3.74	1.08	7.75	2.77	0.37	131.12

TABLE 6: The ML estimates and the RMSE value for ARW data with independent innovations.

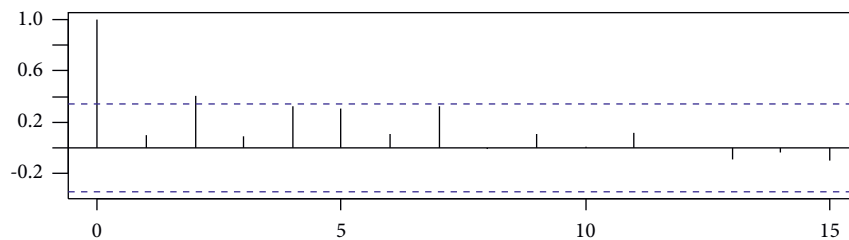
n	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\lambda}$	$\hat{\gamma}$	RMSE
35	0.7596	0.4989	2.3339	0.6293	1.7314

TABLE 7: The ML estimates and the RMSE value for ARW data with AR(1) innovations.

n	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\rho}$	RMSE
35	-0.6299	2.7006	2.7227	-0.33	0.6998	1.5411



(a)



(b)

FIGURE 5: (a) Exact and estimated values of data based on Model (I). (b) ACF of residuals, $h_n = 0.04$.

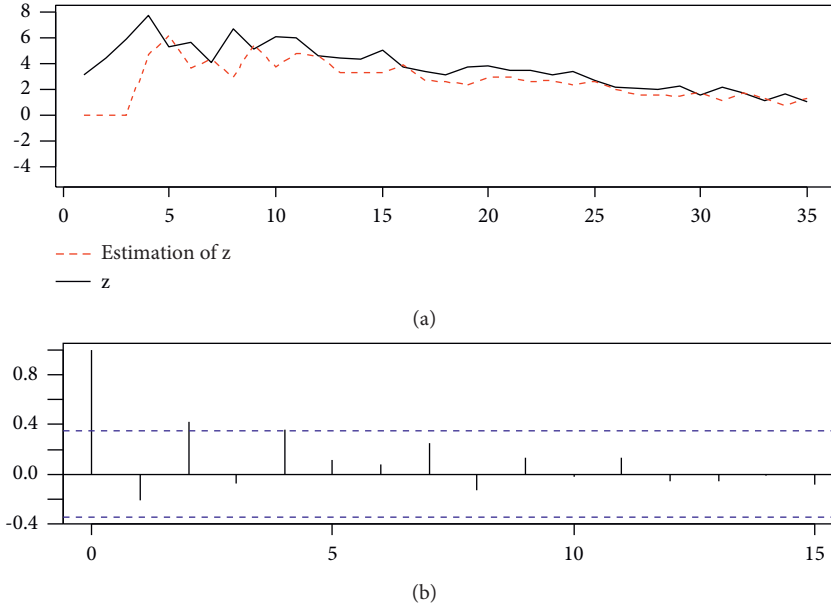


FIGURE 6: (a) Exact and estimated values of data based on Model (II). (b) ACF of residuals, $h_n = 0.04$.

6. Empirical Application

This section analyzes an application of the partially linear autoregressive model with SN innovations to a real-world data set. This dataset consists of three *Pinus eldarica* trees randomly selected from a plantation at Garagpas-Kelardasht site, located in the western part of the Mazandaran province in the north of Iran. Farnoosh et al. [10] and Hajrajabi and Mortazavi [9] studied this dataset by applying different autoregressive models. Considering the proposed semiparametric approach and by assuming $r(x, \beta) = \beta e^{-x^2}$, we estimate autoregression function in two models (i.e., partially linear autoregressive models with AR(1) and independent innovations) for the ARW data from 1974 to 2008 using the SN innovations.

Table 5 shows descriptive statistics for the ARW data. We can see that the data have skewness.

Tables 6 and 7 report the ML estimates and RMSE values for the models with independent and AR(1) innovations, respectively.

Figures 5(a) and 6(a) show the exact values of the ARW data and their estimates in the partially linear autoregressive models with independent and AR(1) innovations, respectively. For diagnostic checking of the fitted models, the residuals of models are analyzed by using ACF plot. Figures 5(b) and 6(b) show the ACF of residuals in the partially linear autoregressive models with independent and AR(1) innovations, respectively. The residuals of the models are almost uncorrelated. As we can see from these figures, the model with AR(1) innovations fits better than the model with independent innovations in describing the ARW data. Also, the comparison of RMSE for the two discussed models shows that the suggested semiparametric approach for a partially linear autoregressive model with AR(1) innovations is more efficient.

7. Conclusion

This article suggested the partially linear autoregressive model with SN innovations from a semiparametric point of view. Both independent and dependent innovations are considered. The CNLSE approach and the local L2-fitting criterion are used to estimate the regression function. For parameter estimation, we applied the CML method using the EM algorithm. The findings of the simulation studies indicated the proposed model is quite flexible for modeling skewed data. Furthermore, the proposed semiparametric method is used to show that the partially linear autoregressive model under SN innovations is an efficient model for modeling the ARW data of Kelardasht site. The results of the study verified the effectiveness of the proposed model.

Appendix

A. The Asymptotic Behaviors of Estimators

To investigate asymptotic behaviors of the estimators, we consider the following assumptions A1-A12 of Farnoosh et al. [21]:

- (A1) $f(\cdot)$ is Lipschitz continuous and all moments of innovations are finite with Lipschitz density function. Also, the sequence $\{Z_t\}$ is a stationary ergodic sequence of integrable random variables.
- (A2) $\partial r / \partial \beta_i, \partial^2 r / \partial \beta_i \partial \beta_j, \partial^3 r / \partial \beta_i \partial \beta_j \partial \beta_k$ exist and are continuous for all $\beta \in \Theta$, where $i, j, k = 1, \dots, m$.
- (A3) $E(Z_t | Z_{t-1}, \dots, Z_0) = E(Z_t | Z_{t-1}, \dots, Z_{t-k})$, a.s. $t \geq k \geq 1$, where k is constant.
- (A4) $E(U_t^2(\mu_0, \sigma_0, \lambda_0, \beta_0, \gamma_0) | \partial r(\beta_0, z_{t-1}) / \partial \beta_i, \partial r(\beta_0, z_{t-2}) / \partial \beta_j) < +\infty$, $E(U_t^2(\mu_0, \sigma_0, \lambda_0, \beta_0, \gamma_0, \rho_0) | \partial r(\beta_0, z_{t-1}) / \partial \beta_i, \partial r(\beta_0, z_{t-1}) / \partial \beta_j) < +\infty$, $l = 2, 3$

where U_t in the independent innovation case is given by

$$\begin{aligned} U_t(\mu_0, \sigma_0, \lambda_0, \beta_0, \gamma_0) &= Z_t - E(Z_t | Z_{t-1}, Z_{t-2}) \\ &= Z_t - \gamma_0 Z_{t-1} - f(Z_{t-2}) - \mu_0 \\ &\quad - \sqrt{\frac{2}{\pi}} \sigma_0 \frac{\lambda_0}{\sqrt{1 + \lambda_0^2}} \end{aligned} \tag{A.1}$$

and in the dependent innovation case is as

$$\begin{aligned} U_t(\mu_0, \sigma_0, \lambda_0, \beta_0, \gamma_0, \rho_0) &= Z_t - \gamma_0 Z_{t-1} - f(Z_{t-2}) \\ &\quad - \rho(Z_{t-1} - \gamma_0 Z_{t-2} - f(Z_{t-3})) \\ &\quad - \mu_0 - \sqrt{\frac{2}{\pi}} \sigma_0 \frac{\lambda_0}{\sqrt{1 + \lambda_0^2}} \end{aligned} \tag{A.2}$$

and consider the following matrices:

$$\begin{aligned} B_l &= E\left(\frac{\partial r(\beta_0, z_{t-l})}{\partial \beta_i}, \frac{\partial r(\beta_0, z_{t-l})}{\partial \beta_j}\right), \quad i, j = 1, \dots, m, \quad l = 2, 3, \\ D &= E\left(U_t^2(\mu_0, \sigma_0, \lambda_0, \beta_0, \gamma_0) \left(\frac{\partial r(\beta_0, z_{t-2})}{\partial \beta_i}, \frac{\partial r(\beta_0, z_{t-2})}{\partial \beta_j}\right)\right), \quad i, j = 1, \dots, m, \\ D_l &= E\left(U_t^2(\mu_0, \sigma_0, \lambda_0, \beta_0, \gamma_0, \rho_0) \left(\frac{\partial r(\beta_0, z_{t-l})}{\partial \beta_i}, \frac{\partial r(\beta_0, z_{t-l})}{\partial \beta_j}\right)\right), \quad i, j = 1, \dots, m, \quad l = 2, 3. \end{aligned} \tag{A.3}$$

We will assume throughout that $B_l, D,$ and D_l ($l = 1, 2$) are positive definite matrices.

(A5) For $i, j = 1, \dots, m,$ the expectations $E[(z_t - r)\partial r/\partial \beta_i], E[(z_t - r)\partial^2 r/\partial \beta_i \partial \beta_j]$ and $E|\partial r/\partial \beta_i, \partial r/\partial \beta_j|$ are finite, where r and its partial derivatives are evaluated at β_0 and $Z_{t-2}, Z_{t-3}.$

(A6) For $i, j, k = 1, \dots, m,$ there exist functions:

$$\begin{aligned} G^{(0)}(Z_t - l), G_i^{(1)}(Z_t - l), G_{i,j}^{(2)}(Z_t - l), G_{i,j,k}^{(3)}(Z_t - l), \\ l = 2, 3, \end{aligned} \tag{A.4}$$

such that

$$\begin{aligned} |r| &\leq G^{(0)}, \\ \left|\frac{\partial r}{\partial \beta_i}\right| &\leq G_i^{(1)}, \\ \left|\frac{\partial^2 r}{\partial \beta_i \partial \beta_j}\right| &\leq G_{i,j}^{(2)}, \\ \left|\frac{\partial^3 r}{\partial \beta_i \beta_j \beta_k}\right| &\leq G_{i,j,k}^{(3)}, \end{aligned} \tag{A.5}$$

for all $\beta \in \Theta,$ and we assume

$$\begin{aligned} E(G_{i,j,k}^{(3)}(Z_t - l), Z_t), \\ E\{G_{i,j,k}^{(3)}(Z_t - l), G^{(0)}(Z_t - l)\}, \\ E\{G_{i,j}^{(2)}(Z_t - l), G_i^{(1)}(Z_t - l)\}, \end{aligned} \tag{A.6}$$

are finite for $l = 2, 3.$

(A7) The sequence $(Z_t)_{t \in \mathbb{Z}}$ is α -mixing (see [22]).

(A8) The random variables Z_0 and Z_1 have the same distribution $\pi(\cdot)$ such that the density $m(\cdot)$ of $\pi(\cdot)$ exists, bounded, continuous, and strictly positive in a neighborhood of the point $x.$

(A9) Functions $f(x)$ and $r(x, \beta)$ are bounded and continuous with respect to $x,$ in a neighborhood of the point x away from 0. Set $r(x, \beta_0) = r_{\beta_0}(x).$

(A10) Function $r(x, \beta)$ has a continuous derivative with respect to β and its derivative at the point β_0 is uniformly bounded with respect $x.$

(A11) The kernel function $K: R^1 \rightarrow R^+$ is a compactly symmetric bounded function, such that $K(x) > 0$ for x in a set of positive Lebesgue measures.

(A12) $h_n = \tau n^{-1/5},$ where $\tau > 0.$

Considering the assumptions A1–A12, we have the following theorems and lemmas:

Lemma 4. Under the assumptions of A1–A12, as $n \rightarrow \infty,$ we have

$$\begin{aligned} (a) \quad n^{-4/5} \sum_{t=2}^n K(z_{t-2} - x/h_n) f(z_{t-2}) r(z_{t-2}, \hat{\beta}_n) &\xrightarrow{P} \tau m(x) f(x) r_{\beta_0}(x) \\ (b) \quad n^{-4/5} \sum_{t=3}^n K(z_{t-3} - x/h_n) f(z_{t-3}) r(z_{t-3}, \hat{\beta}_n) &\xrightarrow{P} \tau m(x) f(x) r_{\beta_0}(x) \\ (c) \quad n^{-4/5} \sum_{t=2}^n K(z_{t-2} - x/h_n) r^2(z_{t-2}, \hat{\beta}_n) &\xrightarrow{P} \tau m(x) r_{\beta_0}^2(x) \\ (d) \quad n^{-4/5} \sum_{t=3}^n K(z_{t-3} - x/h_n) r^2(z_{t-3}, \hat{\beta}_n) &\xrightarrow{P} \tau m(x) r_{\beta_0}^2(x) \end{aligned}$$

where $m(x)$ and $r_{\beta_0}^2(x)$ are defined in (A8) and (A9), respectively.

Theorem 1. Consider the estimator $\hat{f}(x)$ in equation (10). Then, $\hat{f}(x) \xrightarrow{P} f(x)$, as $n \rightarrow \infty$.

Proof. Using Lemma 4 and the strong consistency of $\hat{\beta}_n, \hat{\gamma}_n$ and $\hat{\mu}_{\gamma_N} = \hat{\mu}_N + \hat{\sigma}_N \sqrt{2/\pi \hat{\lambda}_N} / \sqrt{1 + \hat{\lambda}_N^2}$, we can prove Theorem 1. \square

Theorem 2. Let $\tilde{f}(x)$ be the autoregression function estimator given in equation (13) for the model with independent innovations. Then, $|\tilde{f}(x) - \hat{f}(x)| \xrightarrow{P} 0$ as $n \rightarrow \infty$.

Proof. We can obtain the following equality for independent innovations:

$$\tilde{f}(x) - \hat{f}(x) = r(x, \hat{\beta}_n) \frac{\sum_{t=2}^n K(z_{t-2} - x/h_n) \tilde{\gamma}_t r(z_{t-2}, \hat{\beta}_n)}{\sum_{t=2}^n K(z_{t-2} - x/h_n) r^2(z_{t-2}, \hat{\beta}_n)}. \tag{A.7}$$

One can write

$$\begin{aligned} n^{-4/5} \sum_{t=2}^n K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t r(z_{t-2}, \hat{\beta}_n) &= n^{-4/5} \sum_{t=2}^n K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t (r(z_{t-2}, \hat{\beta}_n) - r(z_{t-2}, \beta_0)) + n^{-4/5} \sum_{t=2}^n K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t r(z_{t-2}, \beta_0), \\ &= C_n + D_n. \end{aligned} \tag{A.8}$$

It is known that $\max_{1 \leq t \leq n} |\tilde{\gamma}_t| = O((\log n)^{1/2})$ a.s. as $n \rightarrow \infty$ (see [23]). Then,

$$\begin{aligned} |C_n| &\leq n^{-4/5} \sum_{t=2}^n k_0 |\tilde{\gamma}_t| |r(z_{t-2}, \hat{\beta}_n) - r(z_{t-2}, \beta_0)|, \\ &\leq n^{-4/5} n O((\log n)^{1/2}) \cdot O\left(\left(\frac{\log_2 n}{n}\right)^{1/2}\right) \\ &= \frac{O((\log_2 n \log n)^{1/2})}{n^{3/10}}, \end{aligned} \tag{A.9}$$

where k_0 is an upper bound of the kernel function. Hence, $C_n \rightarrow 0$ a.s. as $n \rightarrow \infty$. Since

$$\begin{aligned} E(D_n) &= n^{-4/5} E \sum_{t=2}^n K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t r(z_{t-2}, \beta_0) = 0, \\ E \left\{ n^{-4/5} \sum_{t=2}^n K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t r(z_{t-2}, \beta_0) \right\}^2 \\ &= n^{-8/5} \left\{ E \left(\sum_{t=2}^n K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t^2 r^2(z_{t-2}, \beta_0) \right) \right. \\ &\quad \left. + 2 \sum_{1 \leq t \leq t' \leq n} E \left[K\left(\frac{z_{t-2} - x}{h_n}\right) \tilde{\gamma}_t r(z_{t-2}, \beta_0) \times K\left(\frac{z_{t'-2} - x}{h_n}\right) \tilde{\gamma}_{t'} r(z_{t'-2}, \beta_0) \right] \right\} \leq n^{-8/5} \cdot n \cdot k^* \cdot \sigma^2 = O\left(\frac{1}{n^{3/5}}\right). \end{aligned} \tag{A.10}$$

where $k^* > 0$ is a constant, and $D_n \xrightarrow{P} 0$ as $n \rightarrow \infty$. Hence, by lemma A.1 and the strong consistency of $\hat{\beta}_n$ and $\hat{\mu}_{\gamma_n}$, we have

$$|\tilde{f}(x) - \hat{f}(x)| \xrightarrow{P} 0. \tag{A.11} \quad \square$$

Theorem 3. Let $\tilde{f}(x)$ be the defined estimator in equation (22) for the model with dependent innovations. Then $|\tilde{f}(x) - \hat{f}(x)| \xrightarrow{P} 0$, as $n \rightarrow \infty$.

Proof. For the dependent innovations, we have

$$\begin{aligned} \tilde{f}(x) - \hat{f}(x) &= r(x, \hat{\beta}_n) \frac{\sum_{t=3}^n K(z_{t-2} - x/h_n) \tilde{\varepsilon}_t r(z_{t-2}, \hat{\beta}_n) + \sum_{t=3}^n K(z_{t-3} - x/h_n) \tilde{\varepsilon}_{t-1} r(z_{t-3}, \hat{\beta}_n)}{\sum_{t=3}^n [K(z_{t-2} - x/h_n) r^2(z_{t-2}, \hat{\beta}_n) + K(z_{t-3} - x/h_n) r^2(z_{t-3}, \hat{\beta}_n)]}, \\ &= r(x, \hat{\beta}_n) \frac{E_n + F_n}{\sum_{t=3}^n [K(z_{t-2} - x/h_n) r^2(z_{t-2}, \hat{\beta}_n) + K(z_{t-3} - x/h_n) r^2(z_{t-3}, \hat{\beta}_n)]}, \end{aligned} \tag{A.12}$$

where ε_t is a stationary AR(1) process given by $\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t$, and $\varepsilon_0 = \nu_0 \approx 0$. To finish the proof, it is enough to prove $E_n \rightarrow^P 0$, and $F_n \rightarrow^P 0$ as $n \rightarrow \infty$. We have

$$\begin{aligned} E_n &= n^{-8/5} \sum_{t=3}^n \tilde{\varepsilon}_t K\left(\frac{z_{t-2} - x}{h_n}\right) r(z_{t-2}, \hat{\beta}_n), \\ &\leq n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t K\left(\frac{z_{t-2} - x}{h_n}\right) \hat{\rho}^{t-j} \nu_j r(z_{t-2}, \hat{\beta}_n) \\ &\leq n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t K\left(\frac{z_{t-2} - x}{h_n}\right) \nu_j r(z_{t-2}, \hat{\beta}_n). \end{aligned} \tag{A.13}$$

The last term in (A.13) can be written as

$$\begin{aligned} &n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t K\left(\frac{z_{t-2} - x}{h_n}\right) \nu_j r(z_{t-2}, \hat{\beta}_n), \\ &= n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t K\left(\frac{z_{t-2} - x}{h_n}\right) \nu_j (r(z_{t-2}, \hat{\beta}_n) - r(z_{t-2}, \beta_0)) + n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t K\left(\frac{z_{t-2} - x}{h_n}\right) \nu_j r(z_{t-2}, \beta_0) \\ &= G_n + H_n. \end{aligned} \tag{A.14}$$

Notice that $\max_{1 \leq j \leq n} |\nu_j| = O((\log n)^{1/2})$ a.s. as $n \rightarrow \infty$ (see [8]). Then,

$$\begin{aligned} |G_n| &\leq n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t \left| K\left(\frac{z_{t-2} - x}{h_n}\right) \right| |\nu_j| |r(z_{t-2}, \hat{\beta}_n) - r(z_{t-2}, \beta_0)|, \\ &\leq \sum_{t=3}^n \sum_{j=0}^t k_0 \cdot O((\log n)^{1/2}) \cdot O\left(\left(\frac{\log_2 n}{n}\right)^{1/2}\right) \\ &= \frac{O((\log_2 n \log n)^{1/2})}{n^{1/10}}, \end{aligned} \tag{A.15}$$

where k_0 is an upper bound of the kernel. Therefore, $G_n \rightarrow 0$ a.s. as $n \rightarrow \infty$.

Since

$$\begin{aligned}
 E(H_n) &= n^{-8/5} E \left\{ \sum_{t=3}^n \sum_{j=0}^t K \left(\frac{z_{t-2} - x}{h_n} \right) \nu_j r(z_{t-2}, \beta_0) \right\} = 0 \\
 E \left\{ n^{-8/5} \sum_{t=3}^n \sum_{j=0}^t K \left(\frac{z_{t-2} - x}{h_n} \right) \nu_j r(z_{t-2}, \beta_0) \right\}^2 &= n^{-16/5} \left\{ \sum_{t=3}^n \sum_{j=0}^t K^2 \left(\frac{z_{t-2} - x}{h_n} \right) E(\nu_j)^2 r^2(z_{t-2}, \beta_0) \right\} \\
 &\quad + 2n^{-16/5} E \left\{ \sum_{t', t=3}^n \sum_{j', j=0}^{t', t} K \left(\frac{z_{t-2} - x}{h_n} \right) \nu_j r(z_{t-2}, \beta_0) \nu_{j'} K \left(\frac{z_{t'-2} - x}{h_n} \right) r(z_{t'-2}, \beta_0) \right\}, \\
 &\leq n^{-16/5} \cdot n^2 \cdot k^* \cdot \sigma^2 \\
 &= O \left(\frac{1}{n^{6/5}} \right).
 \end{aligned} \tag{A.16}$$

where $k^* > 0$ is a constant. Therefore, $H_n \xrightarrow{P} 0$, as $n \rightarrow \infty$. Using the equality $E_n = G_n + H_n$, it is found that $E_n \xrightarrow{P} 0$ as $n \rightarrow \infty$. Similarly, we can prove $F_n \xrightarrow{P} 0$ as $n \rightarrow \infty$. Thus,

$$|\tilde{f}(x) - \hat{f}(x)| \xrightarrow{P} 0. \quad (\text{A.18}) \quad \square$$

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest in the publication of this article.

Authors' Contributions

All authors contributed equally. All authors read and approved the final manuscript.

References

- [1] B. Tarami and M. Pourahmadi, "Multi-variate t autoregressions: innovations, prediction variances and exact likelihood equations," *Journal of Time Series Analysis*, vol. 24, no. 6, pp. 739–754, 2003.
- [2] P. A. Jacobs and P. A. W. Lewis, "A mixed autoregressive-moving average exponential sequence and point process," *J Time Ser Anal*, vol. 24, pp. 739–754, 2003.
- [3] S. Ghasemi, Z. Khodadadi, and M. Maleki, "Autoregressive processes with generalized hyperbolic innovations," *Communications in Statistics - Simulation and Computation*, vol. 1-13, 2019.
- [4] A. Azzalini, "A class of distribution which includes the normal ones," *Scandinavian Journal of Statistics*, vol. 12, pp. 171–178, 1985.
- [5] P. Bondon, "Estimation of autoregressive models with epsilon-skew-normal innovations," *Journal of Multivariate Analysis*, vol. 100, no. 8, pp. 1761–1776, 2009.
- [6] M. Sharafi and A. R. Nematollahi, "AR(1) model with skew-normal innovations," *Metrika*, vol. 79, no. 8, pp. 1011–1029, 2016.
- [7] R. S. Tsay, *An Introduction to Analysis of Financial Data with R*, John Wiley & Sons, New Jersey, USA, 2013.
- [8] R. Farnoosh and S. J. Mortazavi, "A semiparametric method for estimating nonlinear autoregressive model with dependent errors," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 17, pp. 6358–6370, 2011.
- [9] A. Hajrajabi and S. J. Mortazavi, "The first-order nonlinear autoregressive Model with skew normal innovations: a semiparametric approach," *Iranian Journal of Science and Technology Transaction A-Science*, vol. 43, no. 2, pp. 579–587, 2019.
- [10] R. Farnoosh, M. Hajebi, and S. J. Mortazavi, "A semi-parametric estimation for regression functions in the partially linear autoregressive time series models," *Appl Appl Math*, vol. 9, no. 2, pp. 573–591, 2014.
- [11] J. A. Ortega-Contreras, Y. S. Shmaliy, and J. A. Andrade-Lucio, "Three-Wheeled omnidirectional robot localization in RFID-Tag environments using UFIR filtering," *Wseas Transactions on Signal Processing*, vol. 17, pp. 16–21, 2021.
- [12] H. Tong, *Nonlinear Time Series*, The Clarendon Press Oxford University Press, New York, NY, USA, 1990.
- [13] V. Haggan and T. Ozaki, "Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model," *Biometrika*, vol. 68, no. 1, pp. 189–196, 1981.
- [14] R. Chen and R. S. Tsay, "Functional-coefficient autoregressive models," *Journal of the American Statistical Association*, vol. 88, no. 421, pp. 298–308, 1993.
- [15] J. Fan and Q. Yao, *Nonlinear Time Series: Nonparametric and Parametric Methods*, Springer, New York, USA, 2003.
- [16] Y. Zhuoxi, W. Dehui, and S. Ningzhong, "Semiparametric estimation of regression function in autoregressive models," *J Stat Prob Lett*, vol. 79, no. 2, pp. 165–172, 2009.
- [17] A. Hajrajabi and A. Fallah, "Nonlinear semiparametric AR(1) model with skew-symmetric innovations," *Communications*

- in Statistics - Simulation and Computation*, vol. 47, no. 5, pp. 1453–1462, 2018.
- [18] N. Henz, “A probabilistic representation of the skew-normal,” *Scand Stat*, vol. 13, pp. 171–178, 1986.
- [19] A. Azzalini, “Further results on a class of distributions which includes the normal ones,” *Scand Stat*, vol. 12, pp. 171–178, 1986.
- [20] R. Farnoosh and A. Nademi, “Mixture of autoregressive-autoregressive conditionally heteroscedastic models: semi-parametric approach,” *Journal of Applied Statistics*, vol. 41, no. 2, pp. 271–275, 2014.
- [21] R. Farnoosh, M. Hajebi, and S. Y. Samadi, “A semiparametric estimation for the first-order nonlinear autoregressive time series model with independent and dependent errors,” *Iranian Journal of Science and Technology Transaction A-Science*, vol. 43, no. 3, pp. 905–917, 2019.
- [22] Y. A. Davydov, “Mixing conditions for Markov chains,” *Theory Probab Appl*, vol. 18, pp. 312–328, 1973.
- [23] Z. Yu, D. Wang, and N. Shi, “Semiparametric estimation of regression functions in autoregressive models,” *Statistics & Probability Letters*, vol. 79, no. 2, pp. 165–172, 2009.