Austrian Journal of Statistics
2026, Volume 55, Issue 3, 36–47.
http://www.ajs.or.at/

Modelling of Regression with Non-Gaussian AR Errors

Anna Thomas
Department of Statistics
and Data Science
CHRIST (Deemed to be University)
Bengaluru, India

Nimitha John
Department of Statistics
and Data Science
CHRIST (Deemed to be University)
Bengaluru, India

Abstract

This paper presents a new forecasting methodology aimed at addressing challenges in regression models where the residuals exhibit both autocorrelated errors and non-Gaussian innovations. Traditional regression models often assume independent and normally distributed errors; however, in many real-world applications, these assumptions are violated, leading to biased and inefficient parameter estimates. We employ the Optimal Estimating Function (EF) method and the Cochran-Orcutt procedure to effectively handle autocorrelation and non-Gaussian innovations in the error structure. Through comprehensive simulation studies, we evaluate the robustness and efficiency of this combined approach in estimating the model parameters under various scenarios. The methodology is further validated by applying it to both simulated and real-world datasets, showcasing its versatility and practical relevance. For the real data analysis, we use the Rubber dataset, which consists of observations on rubber production and the area of cultivation over a specified period. This application demonstrates the practical utility of the proposed method in dealing with non-Gaussian autocorrelated errors, providing more accurate forecasts and parameter estimates than traditional methods.

Keywords: regression, autocorrelation, logistic errors, estimating function, Cochran-Orchutt.

1. Introduction

In statistical analysis, regression models are a fundamental tool that is frequently used to investigate the relationship between independent and dependent variables. One of the basic assumptions of regression analysis is that the errors are uncorrelated and normally distributed. However, this assumption is often broken, particularly when dealing with time series data or other situations where observations could show non-normality or autocorrelation. In cases where the errors are autocorrelated and non-gaussian, standard OLS estimates are ineffective and biased. Some techniques, such the Cochran-Orcutt process, generalized least squares, or the Estimating Function approach, are used to solve these issues by accounting for the autocorrelation and non-Gaussian nature of the errors. These methods yield more robust and reliable estimates, thereby enhancing the overall performance of the regression model in the presence of autocorrelated non-Gaussian errors through improved estimator efficiency, reduced bias under model misspecification, increased robustness to heavy-tailed error distributions, and superior finite-sample behavior compared to classical OLS-based methods.

The literature contains several regression models with autocorrelated errors and non-gaussian time series models. Cochrane and Orcutt (1949) point out the importance of modification of standard methods for estimating parameters when the errors are autocorrelated. They proposed a tentative procedure to estimate the parameters in the regression model when the errors are autocorrelated. Lee and Lund (2004) examines the characteristics of both ordinary and generalized least squares estimators in a simple linear regression with autocorrelated stationary errors. Beach and MacKinnon (1978) proposes a maximum likelihood procedure to estimate the parameters of regression with autocorrelated errors. Alpuim and El-Shaarawi (2008) examine the statistical characteristics of the least squares and maximum likelihood estimators in a linear regression model with autoregressive stationary processes for the error terms. Lark (2000) discuss the problem of spatial dependency in regression errors and show how the maximum likelihood (ML) approach may be applied to prevent bias when estimating the standard deviation of regression errors in spatial data analysis.

To analyze time series data, Balakrishna and Others (2021) shows several non-gaussian autoregressive type models. Using quantile functions, Cai (2009) provides a Bayesian method for non-Gaussian autoregressive time series models. In order to facilitate effective Bayesian analysis, Aktekin, Polson, and Soyer (2020) suggest a class of multivariate non-Gaussian time series models that take into consideration both time and series correlations using a common random environment. Ranganath and Balakrishna (2018) develop some non-gaussian time series models to analyze and estimate financial data. Sengupta and Kay (1989) proposes a maximum likelihood estimation for estimating non-gaussian time series models. The problem of parameter estimation in statistical modeling refers to calculating utilizing common methods such as the method of moments, maximum likelihood, and least squares. Nevertheless, a closed-form expression is not obtained when time series models with nongaussian errors are estimated using these conventional methods.

Estimating function (EF) methodology provides an attractive alternative in such situations. Unlike likelihood-based methods, EF does not require full specification of the joint distribution and can yield consistent and asymptotically efficient estimators under weaker assumptions. Moreover, EF methods have been shown to perform well in the presence of model misspecification and non-Gaussian innovations. An estimating function approach for product autoregressive models was proposed by Balakrishna and Muhammed (2017), highlighting its effectiveness when traditional estimation techniques are analytically intractable.

Several non-Gaussian time series models have been developed in the literature. Among non-Gaussian distributions, the logistic distribution has received attention due to its heavier tails compared to the normal distribution and its analytical tractability. Logistic innovations are particularly suitable for modeling data with moderate tail heaviness and symmetric but non-normal behavior, which arise frequently in real-world applications. Therefore in this paper, we have used AR(1) with Logistic errors and combined with regression. A unit root process and a first-order cointegration framework of Autoregressive processes of order 1 with logistic innovations were presented by John and Balakrishna (2018). The parameters were obtained using maximum likelihood estimation (MLE). Nevertheless, the conventional MLE method does not yield an obvious analytical solution in this situation. Since closed-form expressions lack non-normal time series models, traditional estimate techniques are frequently insufficient. Thus, we investigate estimating the parameters of the AR(1) model using the Godambe estimating function.

Despite these developments, there is limited work that simultaneously addresses regression modeling, autocorrelated errors, and non-Gaussian innovations within a unified framework. In particular, regression models with AR(1) errors driven by logistic innovations have not been adequately explored. This motivates the present study, which combines the Cochran–Orcutt procedure for handling autocorrelation with the estimating function approach to accommodate non-Gaussian logistic errors. The proposed combination leverages the strengths of both methods: Cochran–Orcutt effectively removes serial dependence, while the estimating function framework provides robustness and analytical feasibility in the absence of a closed-form likelihood.

To the best of our knowledge, regression with autocorrelated non gaussian innovations has not yet been studied in literature. Therefore, it is our interest to focus on regression with autocorrelated errors with Logistic innovations.

The remainder of the paper is structured as follows: Section 2 discusses the Godambe estimating function methodology and Cochran Orcutt procedure. Section 3 deals with the proposed methodology for estimating the parameters of regression with autocorrelated errors having Logistic Innovations. Section 4 deals with the Algorithm used for the proposed study. Section 5 demonstrates the simulation study and the real data analysis for the proposed model. Finally, the conclusions of the study are presented in Section 6.

2. Methodology

2.1. Regression with autocorrelated errors

Consider a simple linear regression model,

$$Y_t = \beta_0 + \beta_1 Z_t + \delta_t \qquad \text{for } t = 1, 2, \ldots, n. \tag{1}$$

In this model, we typically assume that the error terms are independent, as required by the classical linear regression assumptions. However, when the data is time-dependent, this assumption of independent errors may no longer hold, leading to the presence of autocorrelation among the errors. If the errors, δ_t, exhibit autocorrelation in Equation (1), we can assume that the structure of δ_t follows an AR(1) process, represented as:

$$\delta_t = \phi_1 \delta_{t-1} + a_t \tag{2}$$

where a_t is an independent and identically distributed (i.i.d.) error term, following $\mathcal{N}(0, \sigma_a^2)$. Equation (2) can be rewritten using the back-shift operator B as follows:

$$\phi(B)\delta_t = a_t \tag{3}$$

where $\phi(B) = 1 - \phi_1 B$, and B is the back-shift operator. Assuming that the inverse of $\phi(B)$ exists, we can express δ_t as:

$$\delta_t = \frac{a_t}{\phi(B)}. \tag{4}$$

Substituting Equation (4) into the original regression Equation (1), we get:

$$Y_t = \beta_0 + \beta_1 Z_t + \frac{a_t}{\phi(B)}. \tag{5}$$

To correct for the autocorrelation in the error terms, we transform Equation (1) into a regression model with uncorrelated errors a_t. Multiplying both sides of Equation (1) by $\phi(B)$, we obtain:

$$\phi(B)Y_t = \phi(B)\beta_0 + \beta_1 \phi(B) Z_t + a_t. \tag{6}$$

This transformation results in the following regression model with uncorrelated errors:

$$Y_t^* = \beta_0^* + \beta_1 Z_t^* + a_t \tag{7}$$

where $Y_t^* = \phi(B)Y_t$, $\beta_0^* = (1-\phi_1)\beta_0$, and $Z_t^* = \phi(B)Z_t$.

This transformation process, known as the Cochran-Orcutt procedure, is commonly used to correct for first-order serial correlation in the residuals of a time series regression model. After the transformation, the parameters β_0^* and β_1 can be estimated using appropriate estimation methods such as ordinary least squares (OLS).

2.2. Godambe estimating function

The Generalized Method of Moments (GMM) and the Conditional Least Squares (CLS) method can be thought of as hybrid forms of the Estimating Function (EF) methodology. The EF theory of stochastic processes (Godambe 1985) facilitates parameter estimation in various non-normal time series models. This technique is particularly useful when likelihood-based inference becomes challenging, as is often true with non-Gaussian time series.

Consider a discrete-time stochastic process $\{Z_t, t = 1, 2, \dots\}$, where the finite - dimensional distribution is indexed by a parameter vector θ, belonging to an open subset $\mathcal{A}$ of the q-dimensional Euclidean space. This process is realized by the sequence $\{z_t, t = 1, 2, \dots\}$. Let the underlying probability space be denoted by $(\Omega, \mathcal{F}, P_\theta)$, and the σ-field generated by $\{Z_1, Z_2, \dots, Z_t, t \geq 1\}$ be denoted by $\mathcal{F}_t^x$. We define an r-dimensional vector of martingales $h_t(\theta) = h_t(Z_1, Z_2, \dots, Z_t, \theta)$ for $t = 1, 2, \dots, n$. We now consider square-integrable d-dimensional martingale estimating functions of the following form, and a class $\mathcal{E}$ with zero mean:

$$\mathcal{E} = \left\{ f_n(\theta) : f_n(\theta) = \sum_{t=1}^{n} b_{t-1}(\theta) k_t(\theta) \right\},$$

where $b_{t-1}(\theta)$ are $d \times r$ matrices that depend on $Z_1, Z_2, \dots, Z_{t-1}$ and the parameter vector θ, and are referred to as weight matrices.

We examine the simple estimating functions $n_t(\theta)$ and $r_t(\theta)$, which are linear and quadratic forms, respectively. These functions capture the deviation of the process from its conditional mean and variance:

$$n_t(\theta) = Z_t - \alpha_{t-1}(\theta) \quad \text{and} \quad r_t(\theta) = n_t^2(\theta) - \gamma_{t-1}^2(\theta), \tag{8}$$

where Z_t is the observed data, $\alpha_{t-1}(\theta)$ and $\gamma_{t-1}^2(\theta)$ are model-dependent functions of the parameter θ.

The conditional mean $\alpha_{t-1}(\theta)$ and conditional variance $\gamma_{t-1}^2(\theta)$ are defined as:

$$\alpha_{t-1}(\theta) = E(Z_t \mid \mathcal{F}_{t-1}^x), \quad \gamma_{t-1}^2(\theta) = \mathrm{Var}(Z_t \mid \mathcal{F}_{t-1}^x), \tag{9}$$

where $\mathcal{F}_{t-1}^x$ represents the σ-field generated by the past observations up to time $t-1$.

By setting $k_t(\theta) = (n_t(\theta), r_t(\theta))$, the conditional covariance matrix $V_{t-1}(\theta)$ of $k_t(\theta)$ is a key component for constructing an optimal estimating function. This covariance matrix takes the form:

$$V_{t-1}(\theta) = \begin{pmatrix} V_{t-1}^{11} & V_{t-1}^{12} \\ V_{t-1}^{21} & V_{t-1}^{22} \end{pmatrix}, \tag{10}$$

where each element of the matrix is defined as follows:

$$V_{t-1}^{11} = \gamma_{t-1}^2(\theta), \quad V_{t-1}^{12} = V_{t-1}^{21} = \alpha_{3t}(\theta) = E\left[(Z_t - \alpha_{t-1}(\theta))^3 \mid \mathcal{F}_{t-1}^x\right],$$

and

$$V_{t-1}^{22} = \alpha_{4t}(\theta) = E\left[(Z_t - \alpha_{t-1}(\theta))^4 \mid \mathcal{F}_{t-1}^x\right].$$

Thus, $V_{t-1}(\theta)$ captures the second, third, and fourth moments of the process, linking the distributional properties of Z_t to the parameter vector θ.

The next crucial element in the estimation procedure is the weight matrix, which is the matrix of conditional expectations of the derivatives of $k_t(\theta)$ with respect to θ. This matrix plays a significant role in the optimality of the estimating function. It is given by:

$$E\left[\frac{\partial k_t(\theta)}{\partial \theta} \mid \mathcal{F}_{t-1}^x\right] = \begin{pmatrix} -\frac{\partial \alpha_{t-1}(\theta)}{\partial \theta} & -\frac{\partial \gamma_{t-1}^2(\theta)}{\partial \theta} \end{pmatrix}. \tag{11}$$

Finally, an optimal estimating equation, denoted as $f_n^*(\theta) = 0$, can be derived by combining the weight matrix and the covariance matrix $V_{t-1}(\theta)$. As suggested by Godambe (1985), this optimal equation takes the form:

$$\begin{aligned} f_n^*(\theta) =& \sum_{t=1}^{n} \left(1 - \frac{V_{12}^{t-1} V_{21}^{t-1}}{V_{11}^{t-1} V_{22}^{t-1}}\right)^{-1} \left[\frac{1}{V_{11}^{t-1}} \frac{\partial \alpha_{t-1}(\theta)}{\partial \theta} - \frac{V_{21}^{t-1}}{V_{11}^{t-1} V_{22}^{t-1}} \frac{\partial \gamma_{t-1}^2(\theta)}{\partial \theta}\right] n_t(\theta) \\ &+ \left[-\frac{V_{12}^{t-1}}{V_{11}^{t-1} V_{22}^{t-1}} \frac{\partial \alpha_{t-1}(\theta)}{\partial \theta} + \frac{1}{V_{22}^{t-1}} \frac{\partial \gamma_{t-1}^2(\theta)}{\partial \theta}\right] r_t(\theta). \end{aligned} \tag{12}$$

The parameter estimates $\hat{\theta}$ are then obtained by solving the optimal estimating equation $f_n^*(\theta) = 0$, which balances the contributions from both the linear and quadratic deviations in the process.

3. Proposed methodology

Consider a simple linear regression model:

$$Y_t = \beta_0 + \beta_1 Z_t + \delta_t, \tag{13}$$

where Y_t is the dependent variable, Z_t is the independent variable, β_0 and β_1 are the regression coefficients, and δ_t is the error term at time t. Assume that δ_t exhibits autocorrelation and follows an AR(1) process with logistic innovations. The structure of δ_t is given by the AR(1) model:

$$\delta_t = \phi_1 \delta_{t-1} + a_t, \tag{14}$$

where ϕ_1 is the autoregressive parameter and a_t are the logistic innovations. The random errors a_t follow a logistic distribution with scale parameter s and location parameter λ, and its probability density function (pdf) is given by:

$$f(a_t; \lambda, s) = \frac{1}{s} \frac{e^{-(a_t - \lambda)/s}}{\left(1 + e^{-(a_t - \lambda)/s}\right)^2}, \quad -\infty < a_t < \infty. \tag{15}$$

In this setting, the main parameter of interest is ϕ_1, which governs the autocorrelation in the error process. Using the stationary mean adjustment α, the AR(1) model can be rewritten as:

$$\delta_t = \alpha + \phi_1(\delta_{t-1} - \alpha) + a_t, \tag{16}$$

where $\eta_t = a_t - \lambda$. Since standard estimation methods (such as OLS) may not accurately estimate the parameters of the model due to the logistic distribution of the errors, we develop an optimal estimating function based on a combination of linear and quadratic estimating functions, denoted by $n_t(\theta)$ and $r_t(\theta)$, respectively.

The elementary linear estimating function is:

$$n_t(\theta) = \delta_t - \alpha - \phi_1(\delta_{t-1} - \alpha), \tag{17}$$

and the quadratic estimating function is:

$$r_t(\theta) = n_t^2(\theta) - \frac{s^2\pi^2}{3} = \eta_t^2 - \frac{s^2\pi^2}{3}. \tag{18}$$

Here, $\frac{s^2\pi^2}{3}$ is the variance of the logistic distribution. Next, we need the partial derivatives of the conditional mean $\alpha_{t-1}(\theta)$ and the conditional variance $\gamma_{t-1}^2(\theta)$ with respect to the parameter vector $\theta = (\alpha, \phi_1, s)$. These are given by:

$$\frac{\partial \alpha_{t-1}(\theta)}{\partial \theta} = (\delta_{t-1} - \alpha, 1 - \phi_1, 0), \tag{19}$$

and

$$\frac{\partial \gamma_{t-1}^2(\theta)}{\partial \theta} = \left(0, 0, \frac{2s\pi^2}{3}\right). \tag{20}$$

The conditional dispersion matrix $V_{t-1}(\theta)$ of the joint process $(n_t(\theta), r_t(\theta))$ is:

$$V_{t-1}(\theta) = \begin{pmatrix} \frac{s^2\pi^2}{3} & 0 \\ 0 & 2\left(\frac{s^2\pi^2}{3}\right)^2 \end{pmatrix}. \tag{21}$$

Using this, we can construct the optimal estimating function by substituting these quantities into the general form of the optimal estimating equation:

$$\begin{aligned} f_n^*(\theta) = \sum_{t=1}^{n} \Bigg\{ & \frac{1}{\frac{s^2\pi^2}{3}} (\delta_{t-1} - \alpha, 1 - \phi_1, 0) \\ & \times (\delta_t - \alpha - \phi_1(\delta_{t-1} - \alpha)) \\ & + \frac{1}{2\left(\frac{s^2\pi^2}{3}\right)^2} (0, 0, \frac{2s\pi^2}{3}) \times \left(\eta_t^2 - \frac{s^2\pi^2}{3}\right) \Bigg\}. \end{aligned} \tag{22}$$

Solving this optimal estimating function gives us the estimating equations for the AR(1) process with logistic errors. These equations are:

$$\sum_{t=1}^{n} (\delta_{t-1} - \alpha)(\delta_t - \alpha - \phi_1(\delta_{t-1} - \alpha)) = 0. \tag{23}$$

The solution to this equation provides an estimate for ϕ_1, the autoregressive parameter of the AR(1) process:

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{n} (\delta_t - \alpha)(\delta_{t-1} - \alpha)}{\sum_{t=1}^{n} (\delta_{t-1} - \alpha)^2}. \tag{24}$$

Now the new model is

$$Y_t = \beta_0 + \beta_1 Z_t + \phi_1 \delta_{t-1} + a_t, \tag{25}$$

to fix autocorrelation we will use the Cochran Orcutt procedure. Equation (28) can be written as

$$\phi(B)\delta_t = a_t \tag{26}$$

$$\delta_t = \frac{a_t}{\phi(B)} \tag{27}$$

where B is the back-shift operator $\phi(B) = 1 - \phi_1 B$, assuming that the inverse exists. Therefore Equation (5) becomes

$$Y_t = \beta_0 + \beta_1 Z_t + \frac{a_t}{\phi(B)}. \tag{28}$$

The Equation (5) can be transformed into a regression model with uncorrelated errors a_t as

$$\phi(B)Y_t = \phi(B)\beta_0 + \beta_1\phi(B)Z_t + a_t \tag{29}$$

$$Y_t^* = \beta_0^* + \beta_1 Z_t^* + a_t \tag{30}$$

where $Y_t^* = \phi(B)Y_t, \beta_0^* = (1-\phi)\beta_0,\ Z_t^* = \phi(B)X_t$.

4. Algorithm

The following is a step-by-step summary of the algorithm used to implement a simple linear regression model with autocorrelated logistic errors in the random term:

Step 1: Model specification with autocorrelated Logistic error in the random term

$$Y_t = \beta_0 + \beta_1 Z_t + \delta_t.$$

Step 2: Give initial values for β_0, β_1 and the AR(1) autoregressive parameter ϕ. These values can be set based on prior knowledge or random initialization.
Step 3: Generate AR(1) with Logistic errors.
Step 4: Fit simple linear regression model using OLS.
Step 5: Confirming the autoregressive model of order 1 in the error term of the regression model using estimates obtained from Step 4.
Step 6: Estimate the parameters of AR(1) with Logistic errors using Estimating function.
Step 7: Using Cochran Orcutt procedure to the model obtained in Step 3 to obtain the final estimates of the model by capturing dependencies of error.
Step 8: Check for the assumption of uncorrelated errors for the residuals.
Step 9: Finally, compute the mean squared error (MSE) of the estimated parameters to assess the accuracy of the model.

5. Numerical analysis

5.1. Simulation study

A simulation study was conducted to evaluate the performance of regression models with autocorrelated errors using the model described below:

$$Y_t = \beta_0 + \beta_1 Z_t + \delta_t \tag{31}$$

In this model, the error term δ_t follows an autoregressive process of order 1 with Logistic errors. Specifically, we simulated AR(1) errors where the logistic distribution has a scale parameter of 1 and a location parameter $\lambda = 1$. The model parameters used for the simulation were $\beta_0 = \{2, 3, 1\}$ and $\beta_1 = \{3, 4, 5\}$. These values were chosen randomly within reasonable ranges to represent different intercept and slope levels and to assess the performance of the estimators under varying parameter settings. Furthermore, the AR(1) parameter ϕ was set to 0.5 for all simulations.

The parameters of the AR(1) model with logistic errors were estimated using the estimation function method. Simulations were conducted across varying sample sizes n = (50,100, 200,500,1000), with each configuration repeated over 100 iterations to obtain robust results. All simulations and model estimations were implemented using R software. The table below presents the mean estimates of β_0 and β_1, along with their corresponding mean square errors (MSE). This approach allowed us to assess the accuracy and precision of the parameter estimates for different model configurations under the influence of autocorrelated Logistic errors. The averaged results provide insights into how well the Estimating Function method performs in the presence of such errors across different sample sizes.

Table 1 presents the average estimates and corresponding Mean Squared Errors (MSE) of the regression model with Autoregressive Logistic errors. The simulation results demonstrate that, as the sample size increases, the MSE of the parameter estimates systematically decreases, indicating an improvement in the precision of the estimates with larger samples. One significant observation is that the estimation of ordinary least squares (OLS) fails to account for the autocorrelation present in the error structure introduced by the AR(1) logistic errors. This inability to accommodate the autoregressive nature of the errors results in biased and inefficient estimates when using OLS, as it assumes that the errors are independent and identically distributed (i.i.d.).

To address this limitation, the Estimating Function (EF) method, followed by the Cochran-Orcutt procedure, is employed to model the regression with AR(1) Logistic errors. The EF method captures the non gaussian structure in the errors effectively, producing more accurate and efficient estimates. The results indicate that this approach leads to better performance in terms of both bias and MSE. As seen in Table 1, the EF method's estimates show consistently lower MSE values, particularly as sample sizes grow, making it a superior choice for modeling regression with autocorrelated Logistic errors.

5.2. Data analysis

The proposed methodology is applied to the Rubber dataset, which is sourced from the FAOSTAT database. This dataset spans a time period from 1961 to 2022, containing 62 yearly observations. The variables under study are the area under rubber cultivation and rubber production. In the analysis, rubber production is considered the dependent variable, while the area under rubber cultivation serves as the independent variable. Intially we have estimated the parameters by using OLS estimation and it is found that the residuals are autocorrelated. The ACF and PACF plot of residuals are given below in Figures 1 and 2.

From ACF and PACF plot it is found that residuals are autocorrelated. AR(1) model is fitted to the residuals and to assess the adequacy of the existing AR(1) with normal error model. First, we performed an AR model with normal errors and examined the residuals. It is found from the Shapiro wilks test that the residuals are not normal. The QQ plot of normal residual error is given below in Figure 3.

From the Figure 3, the assumption of normality is rejected by the residual series that is derived from the AR(1) series using normal errors. Hence we fitted the proposed model to check whether the data can be modelled using AR with Logistic errors. The parameter $\hat{\phi}$ estimated using Estimating function is 0.2517887. The QQ plot of residuals from logistic errors is given in Figure 4.

This visualization indicate that the residuals are better modeled by a logistic distribution rather than a normal distribution. This is evident from the patterns observed in the QQ plot, which align more closely with the characteristics of the logistic distribution. We used

Table 1: The average estimates with corresponding MSE and standard error

Sample Size	Estimating Function		OLS	
	$\hat{\beta}_0$(MSE, SE)	$\hat{\beta}_1$(MSE, SE)	$\hat{\beta}_0$(MSE, SE)	$\hat{\beta}_1$(MSE, SE)
$\beta_0 = 2,\ \beta_1 = 3$				
50	1.982(0.373, 0.613)	3.026(0.087, 0.295)	4.045(4.478, 0.548)	3.026(0.087, 0.295)
100	2.129(0.278, 0.514)	3.017(0.050, 0.223)	4.021(4.248, 0.406)	3.017(0.050, 0.223)
200	2.043(0.090, 0.298)	2.997(0.021, 0.146)	4.005(4.076, 0.237)	2.997(0.021, 0.146)
500	2.087(0.047, 0.199)	2.997(0.008, 0.088)	3.999(4.019, 0.156)	2.997(0.008, 0.088)
1000	2.118(0.041, 0.164)	2.986(0.004, 0.062)	4.016(4.076, 0.109)	2.986(0.004, 0.062)
$\beta_0 = 3,\ \beta_1 = 4$				
50	2.473(0.807, 0.731)	4.026(0.087, 0.295)	5.045(4.478, 0.548)	4.026(0.087, 0.295)
100	2.656(0.482, 0.606)	4.017(0.050, 0.223)	5.021(4.248, 0.406)	4.017(0.050, 0.223)
200	2.553(0.328, 0.359)	3.997(0.021, 0.146)	5.005(4.076, 0.237)	3.997(0.021, 0.146)
500	2.608(0.210, 0.240)	3.997(0.008, 0.088)	4.999(4.019, 0.156)	3.997(0.008, 0.088)
1000	2.645(0.165, 0.198)	3.986(0.004, 0.062)	5.016(4.076, 0.109)	3.986(0.004, 0.062)
$\beta_0 = 1,\ \beta_1 = 5$				
50	1.492(0.491, 0.502)	5.026(0.087, 0.295)	3.045(4.478, 0.548)	5.026(0.087, 0.295)
100	1.602(0.541, 0.425)	5.017(0.050, 0.223)	3.021(4.248, 0.406)	5.017(0.050, 0.223)
200	1.533(0.341, 0.240)	4.997(0.021, 0.146)	3.005(4.076, 0.237)	4.997(0.021, 0.146)
500	1.565(0.344, 0.160)	4.997(0.008, 0.088)	2.999(4.019, 0.156)	4.997(0.008, 0.088)
1000	1.591(0.366, 0.132)	4.986(0.004, 0.062)	3.016(4.076, 0.109)	4.986(0.004, 0.062)

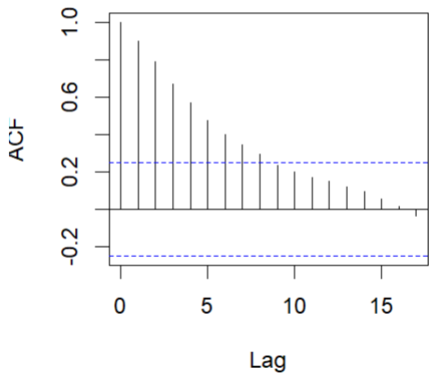

Figure 1: ACF plot

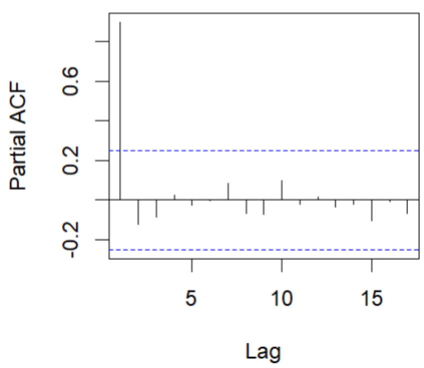

Figure 2: PACF plot

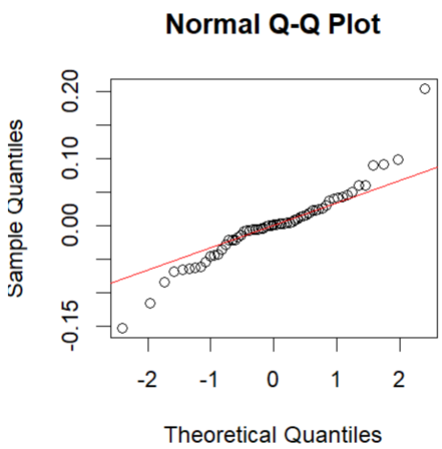

Figure 3: QQ plot of the residuals

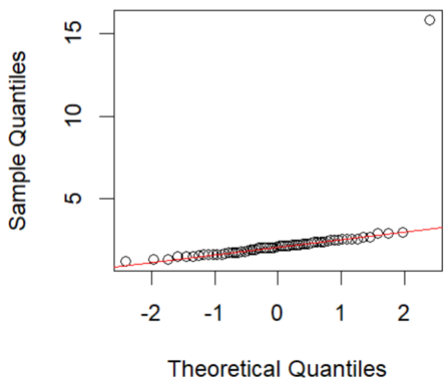

Figure 4: QQ plot of the residuals

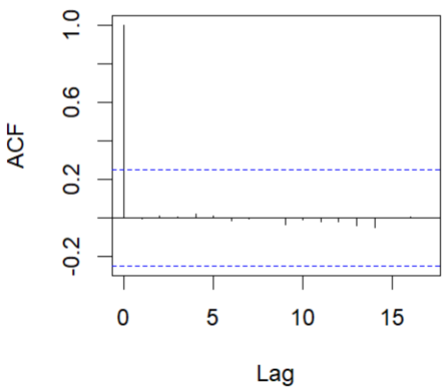

Figure 5: ACF plot

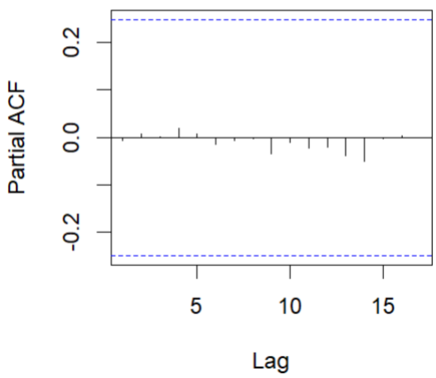

Figure 6: PACF plot

the Kolmogorov-Smirnov test to determine if the residuals follow a logistic distribution using the parameter estimates of the estimating function. The p-value obtained is 0.9464, which indicates that logistic distribution is suitable for these residuals. Thus, we can confirm that the residuals follow the logistic distribution. To get an accurate response, we thus use the Estimating Function Method, which allows for logistic errors. Then, we used the Cochran-Orcutt procedure to fix autocorrelation in the regression model. The final transformed model estimates of the regression model after fixing autocorrelation are β_0 = -6.123115 and β_1 = 1.670254. Then we have checked whether the residuals of transformed model is uncorrelated or not.

After fixing autocorrelation, the above ACF and PACF plot of residuals of the final transformed regression model which is given in Figures 5 and 6 shows that the residuals are uncorrelated. We have compared the accuracy of the proposed model with existing regression with AR(1) normal errors. The accuracy metrics of both models are given in Table 2.

From the above accuracy table given in Table 2, it is observed that the regression with AR(1) Logistic errors gives better accuracy than with Regression with AR(1) Normal errors when the regression data exhibits both autocorrelation and the non-Gaussian innovations for the residuals of Autocorrelated structure.

Table 2: The accuracy measures of different models

Model	RMSE	MAE	MAPE
Reg with AR(1) Logistic Errors	2.923375	2.321316	18.83136
Reg with AR(1) Normal Errors	4.3098	4.54219	34.71602

6. Conclusion

The present study concludes by presenting a new forecasting approach intended to overcome the drawbacks of conventional regression models in the presence of autocorrelated errors and non-Gaussian innovations in the residuals. By combining the Cochran-Orcutt technique with the Optimal Estimating Function (EF) method, we offer a reliable method that increases the accuracy and efficiency of parameter estimation even when faced with such difficult error structures. Comparing this methodology to usual methodologies, the results of extensive simulation experiments and the real-world application using the Rubber dataset show how effective it is in delivering more reliable forecasts and parameter estimations. This combined approach offers a practical solution for a wide range of regression modeling scenarios where the usual assumptions of independent and Gaussian errors are violated. In the real data analysis, we also compared the proposed methodology with a regression model assuming AR(1) normal errors. The results demonstrate that when the residuals of the regression model are autocorrelated and exhibit non-Gaussian behavior, the proposed methodology provides more accurate results than traditional approaches.

References

Aktekin T, Polson NG, Soyer R (2020). "A Family of Multivariate Non-Gaussian Time Series Models." *Journal of Time Series Analysis*, **41**(5), 691–721. doi:10.1111/jtsa.12513.

Alpuim T, El-Shaarawi A (2008). "On the Efficiency of Regression Analysis with AR(p) Errors." *Journal of Applied Statistics*, **35**(7), 717–737. doi:10.1080/02664760801920492.

Balakrishna N, Muhammed PA (2017). "Estimating Function Method for Product Autoregressive Models." *Communications in Statistics - Simulation and Computation*, **46**(5), 3962–3979. doi:10.1080/03610918.2015.1122049.

Balakrishna N, Others (2021). *Non-Gaussian Autoregressive-Type Time Series*. Springer. doi:10.1007/978-981-16-8084-5.

Beach CM, MacKinnon JG (1978). "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors." *Econometrica*, **46**(1), 51–58. doi:10.2307/1913644.

Cai Y (2009). "Autoregression with Non-Gaussian Innovations." *Journal of Time Series Econometrics*, **1**(2). doi:10.2202/1941-1928.1001.

Cochrane D, Orcutt GH (1949). "Application of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms." *Journal of the American Statistical Association*, **44**(245), 32–61. doi:10.1080/01621459.1949.10483290.

Godambe VP (1985). "The Foundations of Finite Sample Estimation in Stochastic Processes." *Biometrika*, **72**(2), 419–428. doi:10.1093/biomet/72.2.419.

John N, Balakrishna N (2018). "Unit Root and Cointegration with Logistic Errors." *Journal of the Indian Society of Agricultural Statistics*, **72**(1), 39–48.

Lark RM (2000). "Regression Analysis with Spatially Autocorrelated Error: Simulation Studies and Application to Mapping of Soil Organic Matter." *International Journal of Geographical Information Science*, **14**(3), 247–264. doi:10.1080/136588100240903.

Lee J, Lund R (2004). "Revisiting Simple Linear Regression with Autocorrelated Errors." *Biometrika*, **91**(1), 240–245. doi:10.1093/biomet/91.1.240.

Ranganath CG, Balakrishna N (2018). *Modelling and Analysis of Financial Time Series Using Some Non-Gaussian Models*. Ph.D. thesis, Cochin University of Science and Technology.

Sengupta D, Kay SM (1989). "Efficient Estimation of Parameters for Non-Gaussian Autoregressive Processes." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **37**(6), 785–794. doi:10.1109/29.32276.

Affiliation:

Anna Thomas
Department of Statistics and Data Science
CHRIST (Deemed to be University)
Bengaluru
India
E-mail: anna.thomas@res.christuniversity.in

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 55, Issue 3 *Submitted:* 2025-05-31
doi:10.17713/ajs.v55.i3.2232 *Accepted:* 2026-04-21