

2021

## A Class of Copula-Based Bivariate Poisson Time Series Models with Applications

Mohammed Alqawba

Dimuthu Fernando

*Old Dominion University*, [dfern002@odu.edu](mailto:dfern002@odu.edu)

Norou Diawara

*Old Dominion University*, [ndiawara@odu.edu](mailto:ndiawara@odu.edu)

Follow this and additional works at: [https://digitalcommons.odu.edu/mathstat\\_fac\\_pubs](https://digitalcommons.odu.edu/mathstat_fac_pubs)



Part of the [Mathematics Commons](#), and the [Statistical Methodology Commons](#)

---

### Original Publication Citation

Alqawba, M., Fernando, D., & Diawara, N. (2021). A class of copula-based bivariate poisson time series models with applications. *Computation*, 9(10), 1-12, Article 108. <https://doi.org/10.3390/computation9100108>

This Article is brought to you for free and open access by the Mathematics & Statistics at ODU Digital Commons. It has been accepted for inclusion in Mathematics & Statistics Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

Article

# A Class of Copula-Based Bivariate Poisson Time Series Models with Applications

Mohammed Alqawba <sup>1</sup>, Dimuthu Fernando <sup>2</sup> and Norou Diawara <sup>2,\*</sup>

<sup>1</sup> Department of Mathematics, College of Science and Arts, Qassim University, Ar Rass 51452, Saudi Arabia; m.alqawba@qu.edu.sa

<sup>2</sup> Department of Mathematics and Statistics, College of Sciences, Old Dominion University, Norfolk, VA 23529, USA; dfern002@odu.edu

\* Correspondence: ndiawara@odu.edu

**Abstract:** A class of bivariate integer-valued time series models was constructed via copula theory. Each series follows a Markov chain with the serial dependence captured using copula-based transition probabilities from the Poisson and the zero-inflated Poisson (ZIP) margins. The copula theory was also used again to capture the dependence between the two series using either the bivariate Gaussian or “t-copula” functions. Such a method provides a flexible dependence structure that allows for positive and negative correlation, as well. In addition, the use of a copula permits applying different margins with a complicated structure such as the ZIP distribution. Likelihood-based inference was used to estimate the models’ parameters with the bivariate integrals of the Gaussian or t-copula functions being evaluated using standard randomized Monte Carlo methods. To evaluate the proposed class of models, a comprehensive simulated study was conducted. Then, two sets of real-life examples were analyzed assuming the Poisson and the ZIP marginals, respectively. The results showed the superiority of the proposed class of models.



check for updates

**Citation:** Alqawba, M.; Fernando, D.; Diawara, N. A Class of Copula-Based Bivariate Poisson Time Series Models with Applications. *Computation* **2021**, *9*, 108. <https://doi.org/10.3390/computation9100108>

Academic Editor: Anna T. Lawniczak

Received: 31 August 2021

Accepted: 13 October 2021

Published: 18 October 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** count time series; copula; Markov chains; multivariate time series; zero-inflated Poisson

## 1. Introduction

Following a similar framework of building bivariate models for ordinal panel data via [1], we constructed a class of bivariate integer-valued time series models using copula theory. Applying either the bivariate Gaussian copula or the bivariate t-copula functions, we jointly modeled two copula-based Markov time series models (see [2,3]). Such a method allows for flexible dependence structures between the two time series and within each one, which allow for both positive or negative correlations, which is theoretically not presented with most existing methods.

The count time series data have multiple applications, many of which have been covered in the literature; see [4] for a comprehensive review. Sometimes, in finance, climate, public health, and crime data analysis, time series counts come as bivariate vectors that observe not only serial dependence within each time series, but also interdependence or cross-dependence between the two series. To accurately study such data, one needs to account for the two types of dependence that emerge from the observed data by applying multivariate time series models. The literature covering normally distributed multivariate time series is plentiful. However, with count time series data and due to the complexity of the computational burden of analyzing such data, the literature on bivariate or multivariate count time series is limited, and so is the zero-inflated cases of time series.

Following the concepts of the integer-valued autoregressive moving average (IN-ARMA) model, Reference [5] introduced the bivariate integer-valued moving average (BINMA) model, which allows for both positive and negative correlation between counts. Further, they presented an extension to the multivariate version starting from the BINMA

model. Reference [6] proposed a bivariate zero-inflated Poisson model to analyze occupational injuries. Reference [7] introduced a multivariate autoregressive conditional double Poisson model, which can accommodate overdispersion, serial dependence, and cross-correlation. They used a multivariate normal copula to capture the cross-correlation between time series, and parameter estimation was conducted using a two-stage estimation procedure. The work of [8] presented a bivariate integer-valued autoregressive process (BINAR(1)) in which the cross-correlation was modeled using a copula to accommodate both positive and negative correlation. They presented the use of a Frank and Gaussian copula to model dependence, and marginal time series were modeled using Poisson and negative binomial INAR(1) models. Reference [9] applied state space models for multivariate count time series and used the to analyze marketing datasets. Reference [10] proposed a new bivariate Poisson INGARCH model, which allows for positive or negative cross-correlation between time series. Reference [11] proposed a class of flexible bivariate Poisson INGARCH(1,1) models whose dependence was established by a special multiplicative factor. The parameter estimation was conducted using maximization by parts the algorithm and its modified version to reduce the computation time. Reference [12] proposed a model for longitudinal data where the univariate margins were selected from the class of zero-inflated distributions, and the dependence structure was modeled using a D-vine copula. Reference [13] presented a comprehensive review on a copula and its applications in different fields. They presented the use of a copula in time series under both univariate and multivariate setups. Reference [14] also wrote a comprehensive review, but on multivariate time series for count time series.

The rest of the paper is organized as follows. In Section 2, we first provide a brief background on the Poisson and ZIP regression model and the copula theory. Then, Section 3 presents the proposed class of copula-based bivariate models to analyze two dependent time series, each being modeled via copula-based Markov chains and then jointly handled using either the bivariate Gaussian or t-copula functions. In Section 5, we give some simulation results. Section 6 provides two real data examples. The conclusion is given in Section 7.

## 2. Background

### 2.1. The Poisson and ZIP Distributions

The probability mass function (pmf) of the well-known Poisson distribution is defined as:

$$f(y_t) = \frac{e^{-\lambda} \lambda^{y_t}}{y_t!},$$

where  $\lambda > 0$  is the intensity (mean) parameter. To account for a large number of zeros (zero-inflation), a modified version of the Poisson distribution is given by the zero-inflated Poisson (ZIP), which handles the presence of excess zeros. It is described by two parameters denoted as the extra probability thrust  $\omega$  and the intensity parameter  $\lambda$ . This is defined by:

$$f_t(y_t) = \omega I_{\{y_t=0\}} + (1 - \omega) \frac{e^{-\lambda} \lambda^{y_t}}{y_t!}, \quad (1)$$

where  $I_{\{y_t=0\}}$  is the indicator function,  $\omega \in [0, 1]$  is the zero-inflation parameter, and  $\lambda > 0$  is the intensity parameter, or the mean, of the baseline Poisson distribution. Note that the pmf (1) is a mixture of the degenerate distribution with a point mass at zero and a Poisson distribution. There is a large literature associated with the ZIP, starting from [15] to [16].

### 2.2. Copula

As a multivariate cumulative distribution function (cdf), the copula is a joint function that captures the dependence structure between variables. With uniform margins  $U(0, 1)$  as in [17], a  $p$ -dimensional copula is a function  $C : [0, 1]^p \rightarrow [0, 1]$  with the following properties:

1.  $C(1, \dots, a_i, \dots, 1) = a_i, \forall i = 1, 2, \dots, p$  and  $a_i \in [0, 1]$ ;
2.  $C(a_1, a_2, \dots, a_p) = 0$  if at least one  $a_i = 0$  for  $i = 1, 2, \dots, p$ ;
3. For any  $a_{i_1}, a_{i_2} \in [0, 1]$  with  $a_{i_1} \leq a_{i_2}$ , for  $i = 1, 2, \dots, p$ ,

$$\sum_{j_1=1}^2 \sum_{j_2=1}^2 \dots \sum_{j_p=1}^2 (-1)^{j_1+j_2+\dots+j_p} C(a_{1j_1}, a_{2j_2}, \dots, a_{pj_p}) \geq 0.$$

The copula is unique if the marginal distributions are continuous, but not when some of the components are discrete. In fact, under Sklar’s theorem, if  $X_1, X_2, \dots, X_p$  are random variables with marginal distribution functions  $F_1, F_2, \dots, F_p$  and joint cumulative distribution function  $F$ , then the following hold:

1. There exists a  $p$ -dimensional copula  $C$  such that for all  $x_1, x_2, \dots, x_p \in \mathbb{R}$ :

$$F(x_1, x_2, \dots, x_p) = C(F_1(x_1), F_2(x_2), \dots, F_p(x_p));$$

2. If  $X_1, X_2, \dots, X_p$  are continuous, then the copula  $C$  is unique. Otherwise,  $C$  can be uniquely determined on  $p$ - dimensional rectangle  $Range(F_1) \times Range(F_2) \times \dots \times Range(F_p)$ .

There are a number of copula functions, i.e.,  $C$ , from which one can choose. Table 1 shows some of the popular functions of the copula families. For more details on these families, see [18].

**Table 1.** Bivariate copula functions.

Copula	Copula Function
Gaussian	$C(u_1, u_2; \delta) = \Phi_\delta(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \delta \in [-1, 1]$
Frank	$C(u_1, u_2; \delta) = -\frac{1}{\delta} \log \left[ 1 + \frac{(e^{-\delta u_1} - 1)(e^{-\delta u_2} - 1)}{e^{-\delta} - 1} \right], \delta \in \mathbb{R} \setminus \{0\}$
Gumbel	$C(u_1, u_2; \delta) = \exp \left[ - \left( (-\log(u_1))^\delta + (-\log(u_2))^\delta \right)^{1/\delta} \right], \delta \geq 1$
Clayton	$C(u_1, u_2; \delta) = (u_1^{-\delta} + u_2^{-\delta} - 1)^{-1/\delta}, \delta > 0$
Plackett	$C(u_1, u_2; \delta) = \frac{[1+(\delta-1)(u_1+u_2)] - \sqrt{[1+(\delta-1)(u_1+u_2)]^2 - 4u_1u_2\delta(\delta-1)}}{2(\delta-1)}, \delta \geq 0$
BVT	$C(u_1, u_2; \delta) = \tau_\delta(\tau^{-1}(u_1), \tau^{-1}(u_2)), \delta \in [-1, 1]$

### 3. Constructing the Bivariate Models

Assume we observe the following series of a two-dimensional vector,  $\{Y_t\}_{t=1}^n$ , where  $Y_t = (Y_{1t}, Y_{2t})'$  for  $t = 1, 2, \dots, n$ . Using copula theory, we can separately study  $\{Y_{1t}\}_{t=1}^n$  and  $\{Y_{2t}\}_{t=1}^n$  and their joint behavior, which would describe the cross-dependence among the bivariate series and with the assumption that each series  $\{Y_{1t}\}_{t=1}^n$  and  $\{Y_{2t}\}_{t=1}^n$  follows a copula-based Markov process (see [2,3,19] for examples).

We were interested in studying the mean vector of the bivariate series, i.e.,  $\mu_t$ , and more importantly, the correlation matrix, say  $\Gamma(t, t - 1)$ , where:

$$\begin{aligned} \mu_t &= E(Y_t) \\ &= \begin{bmatrix} E(Y_{1t}) \\ E(Y_{2t}) \end{bmatrix}, \end{aligned}$$

and:

$$\begin{aligned} \Gamma(t, t - 1) &= \text{COV}(\mathbf{Y}_t, \mathbf{Y}_{t-1}) \\ &= \begin{bmatrix} \text{COV}(Y_{1t}, Y_{1,t-1}) & \text{COV}(Y_{1t}, Y_{2,t-1}) \\ \text{COV}(Y_{2t}, Y_{1,t-1}) & \text{COV}(Y_{2t}, Y_{2,t-1}) \end{bmatrix}, \end{aligned} \tag{2}$$

where the diagonal elements of the matrix in (2) describe the autocovariance among each of the two series, while the off-diagonal elements describe the cross-correlation between  $Y_{1t}$  and  $Y_{2t}$ .

Hence, observing both types of correlations, the joint probability distribution of  $Y_{1t}$  and  $Y_{2t}$  given  $Y_{1,t-1}$  and  $Y_{2,t-1}$ , respectively, for  $t = 1, \dots, n$ , is given by:

$$f(Y_{1t}, Y_{2t} | Y_{1,t-1}, Y_{2,t-1}) = \int_{V^{-1}(F_{1,t}^-)}^{V^{-1}(F_{1,t}^+)} \int_{V^{-1}(F_{2,t}^-)}^{V^{-1}(F_{2,t}^+)} V_2(z_1, z_2, R) dz_2 dz_1, \tag{3}$$

where  $V^{-1}$  is either the inverse cdf of the normal distribution or the t-distribution with  $V_2(\cdot, R)$  being the bivariate normal or t-distribution, respectively. Here,  $R$  is the correlation matrix associated with the bivariate distribution capturing the cross-sectional dependence, given by:

$$R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

where  $\rho$  is a dependence parameter of either the Gaussian or the t-copula function that explains the cross-dependence between the two series. Furthermore,  $F_{i,t}^+ = F(y_{it} | y_{i,t-1})$  and  $F_{i,t}^- = F(y_{it} - 1 | y_{i,t-1})$ , for  $i = 1, 2$ , where:

$$F(y_{it} | y_{i,t-1}) = \frac{F_{12}(y_{it}, y_{i,t-1}) - F_{12}(y_{it}, y_{i,t-1} - 1)}{f_{t-1}(y_{i,t-1}; \theta)}$$

is the transition cdf of  $Y_{it}$  given  $Y_{i,t-1}$ , for  $i = 1, 2$ , and:

$$F_{12}(y_{it}, y_{i,t-1}) = C(F_t(y_{it}), F_{t-1}(y_{i,t-1}); \delta),$$

where  $C(\cdot; \delta)$  is a bivariate copula function with dependence parameter  $\delta$ , describing the serial dependence in a single series, and  $\theta = (\omega, \lambda)$  is a vector of the marginal parameters of the ZIP distribution and reduces to a scalar with the Poisson distribution, i.e.,  $\theta = \lambda$ .

#### 4. Estimation Method

To draw a likelihood-based inference, the log-likelihood function is built. Such a function has no closed form, so its maximization is presented next.

With the likelihood function construction for  $t = 1$ , the joint (bivariate) distribution of  $Y_{11}$  and  $Y_{21}$  is given by:

$$f(Y_{11}, Y_{21}) = \int_{V^{-1}(F_{1,1}^-)}^{V^{-1}(F_{1,1}^+)} \int_{V^{-1}(F_{2,1}^-)}^{V^{-1}(F_{2,1}^+)} V_2(z_1, z_2, R) dz_2 dz_1, \tag{4}$$

and for  $t = 2, \dots, n$ , the conditional bivariate distribution of  $Y_{1t}$  and  $Y_{2t}$  given  $Y_{1,t-1}$  and  $Y_{2,t-1}$  is given by:

$$f(Y_{1t}, Y_{2t} | Y_{1,t-1}, Y_{2,t-1}) = \int_{V^{-1}(F_{1,t}^-)}^{V^{-1}(F_{1,t}^+)} \int_{V^{-1}(F_{2,t}^-)}^{V^{-1}(F_{2,t}^+)} V_2(z_1, z_2, R) dz_2 dz_1. \tag{5}$$

Hence, combining the functions in (4) and (5), the likelihood function is given by:

$$L(\boldsymbol{\theta}; \mathbf{y}) = f(Y_{11}, Y_{21}) \cdot \prod_{t=2}^n f(Y_{1t}, Y_{2t} \mid Y_{1,t-1}, Y_{2,t-1}), \tag{6}$$

where  $\boldsymbol{\theta} = (\boldsymbol{\theta}', \delta_1, \delta_2, \rho)'$ , where  $\boldsymbol{\theta}$  is the marginal parameter vector and  $\delta_1$  and  $\delta_2$  are the copula parameters to deal with the first and second time series, respectively. The bivariate dependence parameter is captured by  $\rho$ . Therefore, taking the log of the function in (6), we obtain the log-likelihood function as follows:

$$\log L(\boldsymbol{\theta}; \mathbf{y}) = l(\boldsymbol{\theta}; \mathbf{y}) = \log f(Y_{11}, Y_{21}) + \sum_{t=2}^n \log f(Y_{1t}, Y_{2t} \mid Y_{1,t-1}, Y_{2,t-1}). \tag{7}$$

Maximizing the log-likelihood function in (7) provides ML estimates for the proposed class of models. However, within the log-likelihood function exists a bivariate normal or  $t$ -integral function that does not have a closed function as shown in (3). Hence, we evaluated the bivariate integral function using the standard randomized importance sampling method introduced by [20], which has been proven to be effective with dimensions less than ten. Then, the parameter estimates, i.e.,  $\hat{\boldsymbol{\theta}}$ , can be obtained by:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}; \mathbf{y}).$$

This maximizing technique produces a numerically calculated Hessian matrix that provides the Fisher information matrix (FIM). Taking the inverse of the FIM yields the standard errors of the ML estimates of  $\boldsymbol{\theta}$ . In the next section, we evaluate the proposed class of bivariate models via comprehensive simulation studies to show the effectiveness of the estimation method.

### 5. Simulation Studies

A comprehensive simulation study was conducted to evaluate the proposed method, and the asymptotic properties of the parameter estimates were validated. For each univariate time series, we considered a copula-based Markov model, where a copula family was used for the joint distribution of subsequent observations, and then, coupled these two time series using another copula at each time point. Here,  $\lambda_1, \lambda_2$  denote the means of the marginal distributions.  $\delta_1$  and  $\delta_2$  measure the serial dependence within each time series, and  $\rho$  denotes the cross-correlation between two time series. The Gaussian copula and Student's  $t$ -copula were selected as candidate copula families with true parameters ( $\lambda_1 = 3, \lambda_2 = 5, \delta_1 = 0.6, \delta_2 = 0.4, \rho = 0.5$ ). The simulation was performed using sample sizes of 100, 300, and 500 while replicating it 300 and 500 times. For each of these five parameter estimates, the standard error was calculated, and the results are displayed in Tables 2 and 3.

Tables 2 and 3 illustrate that the parameter estimates converge to true values as the number of replicates increases and the standard error decreases as the sample size increases. The results showed that the estimates become more and more robust as the sample size increases.

**Table 2.** Parameter estimates using the Gaussian copula for a univariate and joint distribution for 300 and 500 replicates.

Sample Size	Parameters	300 Replicates		500 Replicates	
		Estimate	SE	Estimate	SE
100	$\lambda_1$	3.185	0.433	3.166	0.412
	$\lambda_2$	5.119	0.395	5.104	0.382
	$\delta_1$	0.501	0.064	0.501	0.062
	$\delta_2$	0.348	0.079	0.343	0.078
	$\rho$	0.450	0.072	0.450	0.071
300	$\lambda_1$	3.170	0.224	3.161	0.235
	$\lambda_2$	5.091	0.215	5.090	0.216
	$\delta_1$	0.510	0.035	0.510	0.035
	$\delta_2$	0.349	0.041	0.348	0.041
	$\rho$	0.439	0.041	0.442	0.040
500	$\lambda_1$	3.173	0.194	3.164	0.191
	$\lambda_2$	5.092	0.166	5.085	0.167
	$\delta_1$	0.512	0.026	0.512	0.027
	$\delta_2$	0.350	0.034	0.353	0.034
	$\rho$	0.442	0.033	0.443	0.033

**Table 3.** Parameter estimates using the Gaussian copula for univariate distributions and the t-copula for the joint distribution for 300 and 500 replicates.

Sample Size	Parameters	300 Replicates		500 Replicates	
		Estimate	SE	Estimate	SE
100	$\lambda_1$	3.200	0.434	3.181	0.416
	$\lambda_2$	5.124	0.399	5.108	0.388
	$\delta_1$	0.495	0.065	0.495	0.064
	$\delta_2$	0.346	0.080	0.342	0.080
	$\rho$	0.426	0.074	0.429	0.071
300	$\lambda_1$	3.190	0.232	3.184	0.241
	$\lambda_2$	5.094	0.214	5.096	0.217
	$\delta_1$	0.504	0.037	0.504	0.036
	$\delta_2$	0.347	0.042	0.347	0.042
	$\rho$	0.418	0.042	0.420	0.041
500	$\lambda_1$	3.193	0.196	3.183	0.192
	$\lambda_2$	5.101	0.168	5.094	0.170
	$\delta_1$	0.505	0.026	0.505	0.028
	$\delta_2$	0.349	0.035	0.352	0.035
	$\rho$	0.419	0.033	0.421	0.033

Expanding the results, we performed the simulations for the ZIP choosing the Gaussian copula for the univariate and joint distributions with true parameters ( $\lambda_1 = 3, \lambda_2 = 5, \omega_1 = 0.3, \omega_2 = 0.4, \delta_1 = 0.6, \delta_2 = 0.4, \rho = -0.5$ ) where the marginals follow the Poisson or ZIP marginals. here for the zero-inflated Poisson distribution where  $\omega_1$  and  $\omega_2$  represent the zero-inflation parameters for the ZI marginal densities. The results in Table 4 indicate the robustness of the parameter estimates even under negative cross-correlation.

**Table 4.** Parameter estimates using the Gaussian copula for the univariate and joint distributions for 300 replicates with the Poisson and ZIP marginals.

Sample Size	Parameters	Poisson		ZIP	
		Estimate	SE	Estimate	SE
100	$\lambda_1$	3.168	0.432	3.402	0.387
	$\omega_1$			0.332	0.084
	$\lambda_2$	4.993	0.378	5.205	0.386
	$\omega_2$			0.405	0.067
	$\delta_1$	0.503	0.065	0.539	0.0841
	$\delta_2$	0.341	0.082	0.354	0.099
	$\rho$	-0.453	0.075	-0.491	0.095
300	$\lambda_1$	3.146	0.225	3.402	0.204
	$\omega_1$			0.337	0.044
	$\lambda_2$	4.978	0.199	5.201	0.191
	$\omega_2$			0.409	0.038
	$\delta_1$	0.510	0.035	0.549	0.044
	$\delta_2$	0.345	0.040	0.362	0.052
	$\rho$	-0.444	0.039	-0.479	0.052
500	$\lambda_1$	3.147	0.191	3.401	0.173
	$\omega_1$			0.338	0.035
	$\lambda_2$	4.970	0.157	5.199	0.148
	$\omega_2$			0.409	0.03
	$\delta_1$	0.510	0.026	0.55	0.033
	$\delta_2$	0.350	0.031	0.367	0.041
	$\rho$	-0.441	0.032	-0.476	0.041

## 6. Applications

The proposed models were applied to two real-life examples, and the results are presented in Sections 6.1 and 6.2.

### 6.1. Application to Forgery and Fraud Data

We considered monthly counts of forgery and fraud in the 61st police car beat in Pittsburgh, PA. For a selected county, two count time series were selected to fit the proposed bivariate Poisson class of models. The empirical means for the two count time series were 2.632 and 8.243, respectively. The cross-correlation between the two time series was 0.307, which provided sufficient evidence that the two series are dependent. Monthly counts of forgery and fraud are displayed in Figure 1. The fraud count series has higher counts and larger fluctuations when compared to the fraud time series over time. Figure 2 shows the distributions of the count via bar plots and sample autocorrelation functions (ACFs) for the forgery and Fraud count series, respectively. The bar plots suggest that both counts follow Poisson distributions, whereas the ACFs indicate that the two series are serially dependent. Hence, we can now investigate both types of dependences properly using the proposed method of bivariate Poisson time series.



### Monthly counts of forgery and fraud

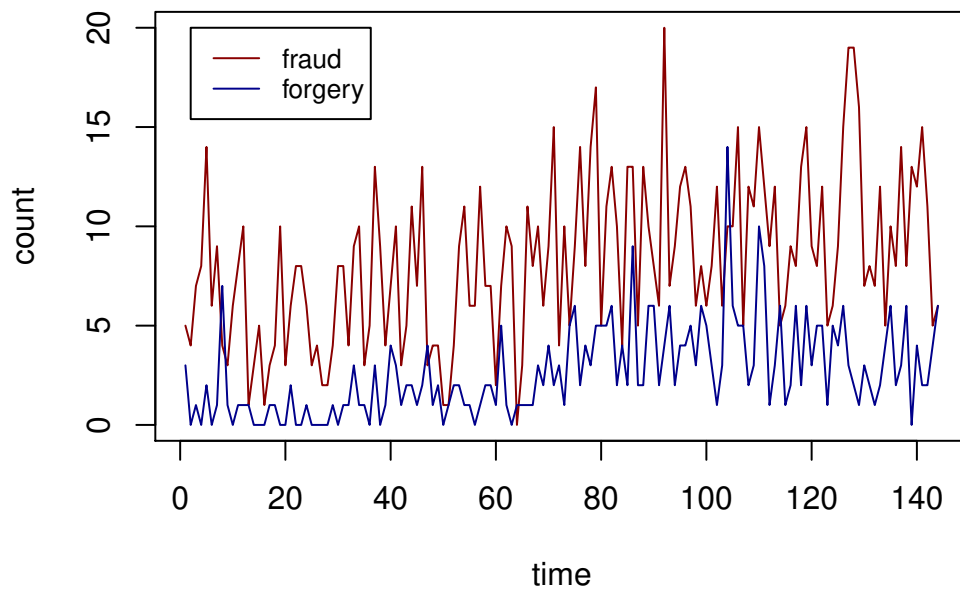


Figure 1. Forgery and fraud count series.

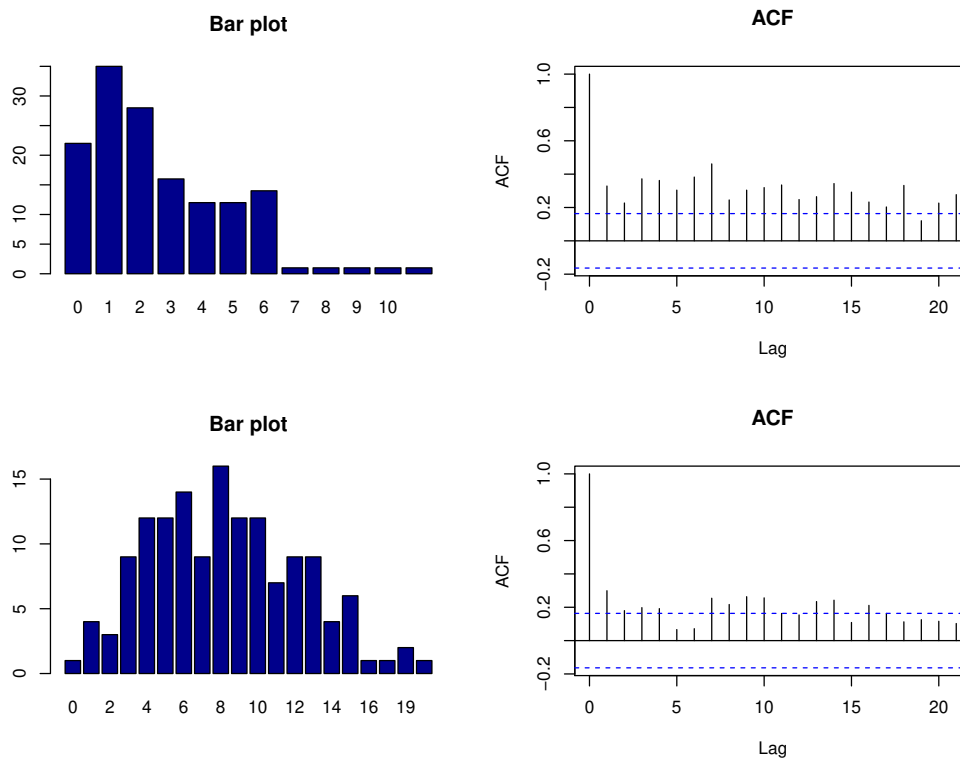


Figure 2. Bar plot and ACFs for counts of forgery (top) and fraud (bottom).

Different copula families were fit for the univariate and joint distributions, and their corresponding AIC values are reported in Table 5. Parameter estimates and the corresponding standard deviations are recorded in Table 6.

**Table 5.** AIC values for the bivariate Poisson model with different copula families.

Univariate Copula Family	Joint Copula Family	
	Gaussian	t
Gaussian	1448.63	1462.68
Frank	1461.7	1475.61
Clayton	1433.78	1447.85
Gumbel	1458.4	2315.55

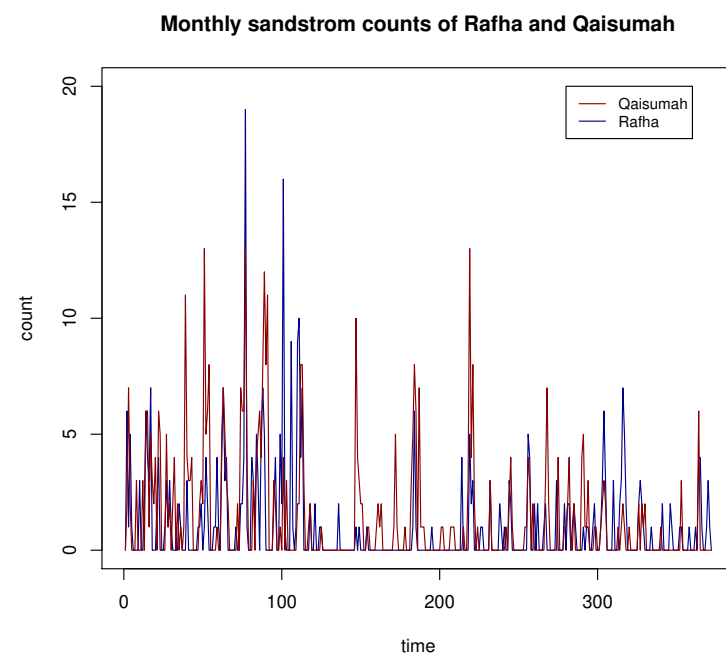
**Table 6.** Parameter estimates for the bivariate Poisson model.

Parameter	Estimate	SE
$\lambda_1$	2.695	0.0657
$\lambda_2$	8.332	0.0350
$\delta_1$	0.261	0.0527
$\delta_2$	0.192	0.0478
$\rho$	0.178	0.0532

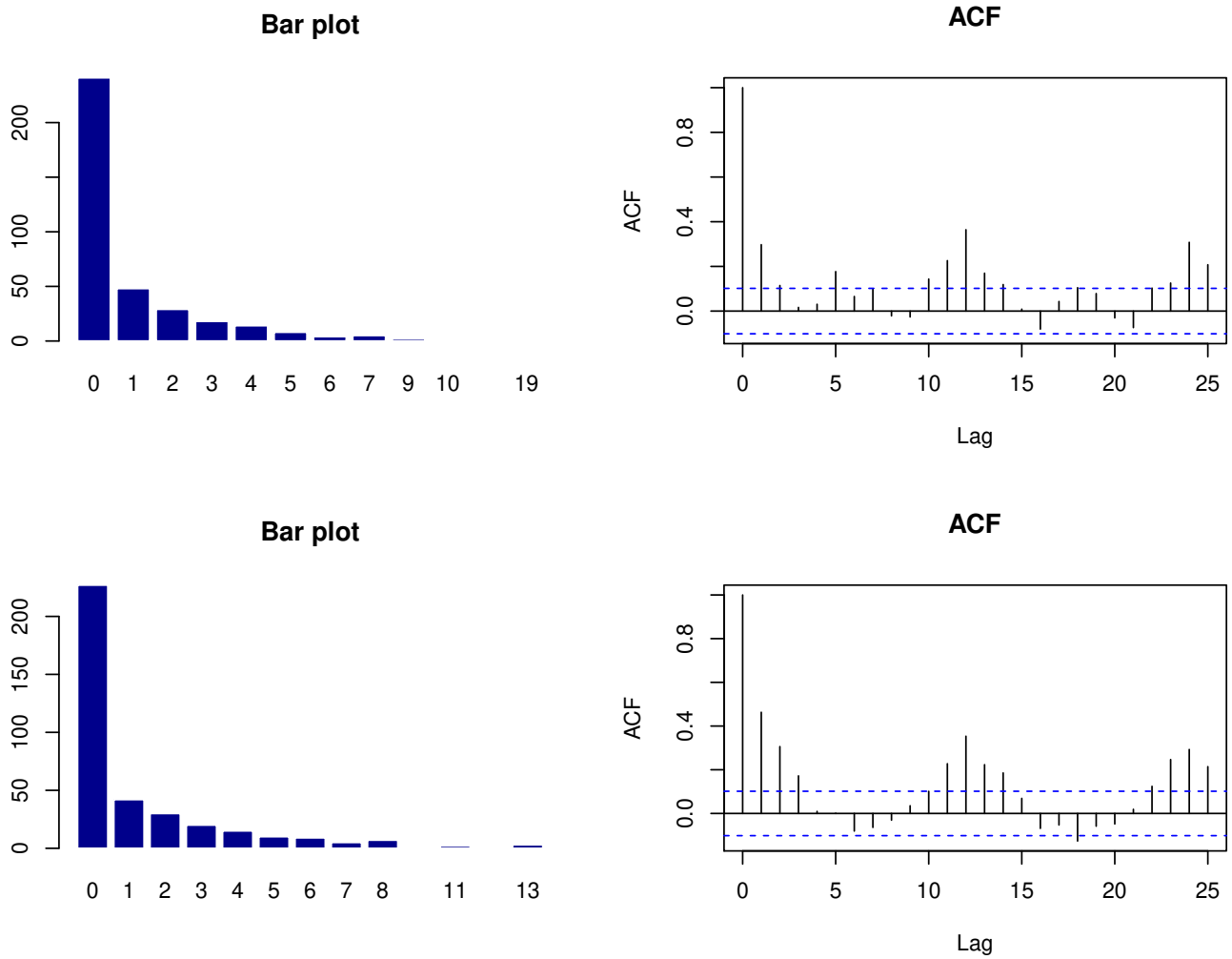
6.2. Application to Sandstorm Data

The dataset used for this application consisted of the monthly count of strong sandstorms recorded by the AQI airport station in Eastern Province, Saudi Arabia. The zero-inflated proportions for the two time series were 65% and 61%, respectively. The cross-correlation between the two time series was 0.465, and this provided evidence that the two counts are cross-correlated. For the proposed BZIP model, the parameter estimates and the corresponding standard deviations are recorded in Table 8.

The two sandstorm count series are shown in Figure 3. The plot shows that there exist the frequent occurrence of zeros for both time series. The bar plots and ACFs of sandstorm counts for Rafha and Qaisumah are shown in Figure 4. The plots show that there exists the frequent occurrence of zeros and low-ordered autocorrelation. Hence, we applied the bivariate ZIP distribution to model the sandstorm counts.



**Figure 3.** Sandstorm counts of Rafha and Qaisumah.



**Figure 4.** Bar plot and ACFs for sandstorm counts of Rafha (**top**) and Qaisumah (**bottom**).

Different copula families were fit for the univariate and joint distributions, and their corresponding AIC values are reported in Table 7. Table 8 represents the parameter estimates and corresponding standard deviations for the fitted bivariate zero-inflated model using the Gaussian copula for the joint and marginal distributions.

The standard errors are very small under both the Poisson and ZIP marginals, which emphasizes the robustness of our proposed method.

**Table 7.** AIC values for the BZIP model with different copula families.

Univariate Copula Family	Joint Copula Family	
	Gaussian	t
Gaussian	2232.2	2272.09
Frank	2122.9	2153.99
Clayton	2089.8	2268.88
Gumbel	4574.24	7646.1

**Table 8.** Parameter estimates for the bivariate zero-inflated Poisson model.

Parameter	Estimate	SE
$\lambda_1$	1.384	0.0003
$\omega_1$	0.502	0.0003
$\lambda_2$	2.505	0.0604
$\omega_2$	0.690	0.1501
$\delta_1$	0.327	0.0024
$\delta_2$	0.300	0.0454
$\rho$	0.210	0.0469

## 7. Summary

In this paper, we proposed a class of bivariate integer-valued time series models that was constructed via copula theory. The use of Markov chains to capture the serial dependence via copula-based transition probabilities with the Poisson and the zero-inflated Poisson (ZIP) margins was described. The copula theory was also used again to capture the dependence between the two series using either the bivariate Gaussian or t-copula functions. Simulated examples were given to evaluate the likelihood-based estimation method with importance sampling to evaluate the bivariate normal or t-integrals. Two sets of count data were analyzed using the introduced class of models with the Poisson and ZIP margins, respectively. The simulation studies and real-life examples proved the effectiveness of the proposed method.

A future extension will be constructed for a multivariate class with a larger dimension using the vine copula.

**Author Contributions:** Conceptualization, M.A., D.F. and N.D.; methodology, M.A., D.F. and N.D.; software, M.A., D.F. and N.D.; validation, M.A., D.F. and N.D.; formal analysis, M.A., D.F. and N.D.; investigation, M.A., D.F. and N.D.; data curation, M.A., D.F. and N.D.; writing-original draft preparation, M.A., D.F. and N.D.; writing-review and editing, M.A., D.F. and N.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** This research was considered “exempt” by the Institutional Review Board of Old Dominion University.

**Data Availability Statement:** Data is attached to the document.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Nikoloulopoulos, A.K.; Moffatt, P.G. Coupling couples with copulas: Analysis of assortative matching on risk attitude. *Econ. Inq.* **2019**, *57*, 654–666. [[CrossRef](#)]
2. Joe, H. Markov models for count time series. In *Handbook of Discrete-Valued Time Series*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2016; pp. 49–70.
3. Alqawba, M.; Diawara, N. Copula-based Markov zero-inflated count time series models with application. *J. Appl. Stat.* **2021**, *48*, 786–803. [[CrossRef](#)]
4. Davis, R.A.; Fokianos, K.; Holan, S.H.; Joe, H.; Livsey, J.; Lund, R.; Pipiras, V.; Ravishanker, N. Count time series: A methodological review. *J. Am. Stat. Assoc.* **2021**, 1–50. [[CrossRef](#)]
5. Quoreshi, A.S. Bivariate time series modeling of financial count data. *Commun. Stat.-Theory Methods* **2006**, *35*, 1343–1358. [[CrossRef](#)]
6. Wang, K.; Lee, A.H.; Yau, K.K.; Carrivick, P.J. A bivariate zero-inflated Poisson regression model to analyze occupational injuries. *Accid. Anal. Prev.* **2003**, *35*, 625–629. [[CrossRef](#)]
7. Heinen, A.; Rengifo, E. Multivariate autoregressive modeling of time series count data using copulas. *J. Empir. Financ.* **2007**, *14*, 564–583. [[CrossRef](#)]
8. Karlis, D.; Pedeli, X. Flexible bivariate INAR (1) processes using copulas. *Commun. Stat.-Theory Methods* **2013**, *42*, 723–740. [[CrossRef](#)]

9. Ravishanker, N.; Venkatesan, R.; Hu, S. Dynamic models for time series of counts with a marketing application. In *Handbook of Discrete-Valued Time Series*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2016.
10. Cui, Y.; Zhu, F. A new bivariate integer-valued GARCH model allowing for negative cross-correlation. *Test* **2018**, *27*, 428–452. [[CrossRef](#)]
11. Cui, Y.; Li, Q.; Zhu, F. Flexible bivariate Poisson integer-valued GARCH model. *Ann. Inst. Stat. Math.* **2020**, *72*, 1449–1477. [[CrossRef](#)]
12. Sefidi, S.; Ganjali, M.; Baghfalaki, T. Pair copula construction for longitudinal data with zero-inflated power series marginal distributions. *J. Biopharm. Stat.* **2021**, *31*, 233–249. [[CrossRef](#)] [[PubMed](#)]
13. Größer, J.; Okhrin, O. Copulae: An overview and recent developments. In *Wiley Interdisciplinary Reviews: Computational Statistics*; Wiley: Hoboken, NJ, USA, 2021; p. e1557.
14. Fokianos, K. Multivariate Count Time Series Modelling. *arXiv* **2021**, arXiv:2103.08028.
15. Lambert, D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **1992**, *34*, 1–14. [[CrossRef](#)]
16. Young, D.S.; Roemmele, E.S.; Shi, X. Zero-inflated modeling part II: Zero-inflated models for complex data structures. In *Wiley Interdisciplinary Reviews: Computational Statistics*; Wiley: Hoboken, NJ, USA, 2020; p. e1540.
17. Nelsen, R.B. *An Introduction to Copulas*; Springer: Cham, Switzerland, 2007.
18. Joe, H. *Dependence Modeling with Copulas*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2014.
19. Sun, L.H.; Huang, X.W.; Alqawba, M.S.; Kim, J.M.; Emura, T. *Copula-Based Markov Models for Time Series: Parametric Inference and Process Control*; Springer Nature: Singapore, 2020.
20. Genz, A.; Bretz, F. *Computation of Multivariate Normal and t Probabilities*; Springer Science & Business Media: Cham, Switzerland, 2009; Volume 195.