Copula directional dependence of discrete time series marginals

Mohammed Alqawba, Norou Diawara & Jong-Min Kim

To cite this article: Mohammed Alqawba, Norou Diawara & Jong-Min Kim (2019): Copula directional dependence of discrete time series marginals, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2019.1630434

To link to this article: https://doi.org/10.1080/03610918.2019.1630434

Published online: 19 Jun 2019.
Copula directional dependence of discrete time series marginals

Mohammed Alqawba, Norou Diawara, and Jong-Min Kim

ABSTRACT
To understand the dynamic relationship of discrete time series processes, we adopted copula directional dependence via beta regression model applied with generalized autoregressive conditional heteroscedasticity (INGARCH) marginals. To validate the proposed method, we completed simulations of two INGARCH processes from asymmetric bivariate copula function with members such as Gaussian and Plackett copula functions. The simulations show that the proposed method is consistent for deriving directional dependent measurements regardless of the choice of the symmetric members. The proposed method is applied to the bivariate discrete time series data of the monthly counts of sandstorms and dust haze phenomena in Saudi Arabia.

ARTICLE HISTORY
Received 4 September 2018
Accepted 5 June 2019

KEYWORDS
Count time series; GLM; copula; Directional dependence; Beta regression

1. Introduction and motivation
Discrete time series data can be found in several applied disciplines such as the environmental field, biostatistics, economics, and finance (Kedem and Fokianos 2005). In some cases, the time series counts come in pairs, and researchers interested in studying the dynamic relationships of these pairs. Many of the methods applied to study such dynamic relationships are plausible only if the random processes were continuous or the relationships were symmetric. One of these methods is the Granger-causality test introduced by Granger (1969), which has become a standard technique used to study the causal relationship. Kim and Hwang (2017b) introduced a new method of investigating the casual relationships of two asymmetric time series processes by deriving measurements of the directional dependence via a Gaussian copula beta regression model with generalized autoregressive conditional heteroscedasticity (GARCH) marginals. By using directional dependence, we were able to study the joint behavior as suggested by Sungur (2005) and found which variable is more likely to lead the other variable in regression setting, and then we extended results to the integer-valued GARCH. We adopt the method via beta regression and applied it to discrete marginals, such as the INGARCH processes under copula of marginals. One of the advantages of using copula is that it allowed us to model the marginals separately and model the joint distribution, which account for the correlation structure. For the purpose of practical
application, this manuscript describes the method we developed to study the causal relationship of two discrete time series processes jointly via bivariate asymmetric copula function.

The manuscript is organized as follows. Discrete time series models under generalized linear models (GLM) are presented first with Poisson and negative binomial types of occurrence count time series. Inference and estimation technique are also provided. In Sec. 3, the copula is revised and directional dependence is developed for asymmetric bivariate copula function. In the same section, we also recall the beta regression model and derive the suggested measurements of the directional dependence of two discrete time series. Section 4 presents two simulation examples and a real life application with data recoreded in the AIQ airport station located in the Eastern Province, Saudi Arabia. We end with a conclusion in Sec. 5.

2. Discrete time series models

Analyzing and modeling discrete time series data have drawn interest because of the challenges related to the discreteness of the data. Tests were also built to check for any violation of the assumptions made on common models, such as in the autoregressive moving average models (ARMA) (see Brockwell and Davis 2013 and Shumway and Stoffer 2011).

Recently, GLMs were introduced to model time series data with added covariates and sequential history of data (Fokianos 2015). However, the most widely used GLMs make use of the normality assumption. Data limited to normality assumption do not allow us to apply to discrete data. More precisely, suppose we observe a pair of jointly distributed time series, \((X_t, Y_t)\), for \(t \in \mathbb{Z}\), in which \(\{Y_t\}\) is a discrete response time series and \(X_t = (X_{1t}, X_{2t}, ..., X_{kt})^T\) is a \(k\)-dimensional covariate vector. These variables generate a \(\sigma\)-field denoted by \(\mathcal{F}_t\) and are defined as an increasing sequence of \(\sigma\)-fields, \(\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2, \ldots\), such that

\[
\mathcal{F}_t = \sigma\{Y_t, Y_{t-1}, ..., X_{t+1}, X_t, X_{t-1}, \ldots\},
\]

and \(Y_t\) is \(\mathcal{F}_t\) measurable with the conditional density of \(Y_t\) given \(\mathcal{F}_{t-1}\) as \(f_t(y_t; \theta)\) where \(\theta\) is a fixed parameter vector. Hence, given a time series data \(\{y_1, y_2, ..., y_n\}\), the partial likelihood function is given by

\[
f(y; \theta) = \prod_{t=1}^{n} f_t(y_t; \theta|\mathcal{F}_{t-1}). \tag{2.1}
\]

However, taking into account covariates \(X_t\), we observe the pairs \((x_t, y_t)\), for \(t = 1, ..., n\), and under the conditional probability, the joint density of all \((x_{t}, x_{t})\) can be expressed as

\[
f(x, y; \theta) = f(x_1; \theta) \left[ \prod_{t=2}^{n} f_t(x_t; \theta|a_t) \right] \left[ \prod_{t=1}^{n} f_t(y_t; \theta|b_t) \right], \tag{2.2}
\]

where \(a_t = (x_1, y_1, ..., x_{t-1}, y_{t-1})\) and \(b_t = (x_1, y_1, ..., x_{t-1}, y_{t-1}, x_{t})\). Notice that the second product of the right hand of Eq. (2.2) is actually the partial likelihood in Eq. (2.1) as stated by Cox (1975), who argued that the lost information on \(\theta\) from omitting
the first part of Eq. (2.2) is rather small. Thus, the most computational gain comes from just considering the partial likelihood in Eq. (2.1) (Fokianos and Kedem 2004).

Let \( \mu_t = E[Y_t | \mathcal{F}_{t-1}] \) be the conditional expectation of the response variable \( Y_t \) given the past \( \mathcal{F}_{t-1} \). The main goal of employing GLM is to relate \( \mu_t \) to the covariates and the past values of the response. As introduced in Nelder and Wedderburn (1972) and extended in McCullagh and Nelder (1989), GLMs consist of three components: the random component, the systematic component and the link function.

1. Random component: This refers to the conditional probability distribution of the response \( \{y_t\} \) given the \( \sigma \)-field \( \mathcal{F}_{t-1} \), which belongs to an exponential family of distributions. That is, for \( t \in \mathbb{Z} \),

\[
f(y_t; \theta_t, \phi | \mathcal{F}_{t-1}) = \exp \left\{ \frac{y_t \theta_t - b(\theta_t)}{x_t(\phi)} + c(y_t; \phi) \right\},
\]

where \( \theta_t \) is the natural parameter of the distribution. The function \( x_t(\phi) \) is of the form \( \phi / \omega_t \), in which \( \phi \) is a dispersion parameter, and \( \omega_t \) is a fixed number known as weight. Most common distributions of discrete random variables, such as binomial and Poisson, belong to this family. Fokianos (2012) provided a review of modeling count time series, in which Poisson was discussed thoroughly.

1. Systematic component: A linear combination, say \( \eta_t \), of the joint process \( \{Y_t, \mu_t, X_{t+1} : t \in \mathbb{Z}\} \).

2. Link function: This specifies the link between the systematic component \( \eta_t \) and the random component by

\[
g(\mu_t) = \eta_t,
\]

where \( g : \mathbb{R}^+ \rightarrow \mathbb{R} \) is a monotone twice differentiable and known function. In linear models, the choice of \( g \) would be the identity function. However, when the data are discrete, researchers often prefer the natural log function if the data is distributed as Poisson or negative binomial, for example, and the logit function with binomial or Bernoulli data.

Under a general form of the models, we considered the link function as

\[
g(\mu_t) = \eta_t = \beta_0 + \beta' X_t + \sum_{k=1}^{p} \phi_k h(Y_{t-k}) + \sum_{l=1}^{q} \theta_l g(\mu_{t-l}),
\]

where \( h \) is a known transformation function, and \( \beta' = (\beta_1, ..., \beta_k) \) is the parameter vectors corresponding to covariate effects, and \( p \) and \( q \) are whole numbers chosen levels of AR and MA components.

The general form of Eq. (2.4) covers many different classes of discrete time series models, such as the generalized linear autoregressive moving average models (GLARMA) (Dunsmuir 2015) or the generalized autoregressive moving average models (GARMA) (Benjamin et al. 2003). Another widely used class that falls under Eq. (2.4) is the integer-valued generalized autoregressive conditional heteroskedastic (INGARCH), as considered in Ferland et al. (2006). In the next subsections, we will consider special
cases under different distributional assumptions. We will discuss the model formulation, the associated properties with it, and derive the likelihood function.

2.1. Count time series

Whenever there is a random number of events occurring over a fixed period of time, researchers commonly use either the Poisson or the negative binomial distributions to model these event counts. In this section, we will discuss the models that fall under Eq. (2.4) and model count time series regardless of the distribution as long as it belongs to the exponential family given in Eq. (2.3).

First, consider the linear model proposed by Ferland et al. (2006) with Poisson distribution and the order \((p, q)\) \(p, q \geq 0\). For \(\{Y_t : t \in \mathbb{Z}\}\), a count time series, assume that given the past \(\mathcal{F}_{t-1}\), the counts are distributed as Poisson distribution. That is, \(Y_t|\mathcal{F}_{t-1} \sim \text{Poi}(\mu_t)\) satisfies

\[
\Pr(Y_t = y_t|\mathcal{F}_{t-1}) = \frac{e^{(-\mu_t)}\mu_t^{y_t}}{y_t!}, \quad y_t = 0, 1, \ldots,
\]

with

\[
\mu_t = \beta_0 + \sum_{k=1}^{p} \phi_k Y_{t-k} + \sum_{l=1}^{q} \theta_l \mu_{t-l}, \quad t \geq \max(p, q),
\]

where \(\beta_0 > 0, \phi_k \geq 0, \ k = 1, \ldots, p, \ \theta_l \geq 0, \ l = 1, \ldots, q\) and \(0 < \sum_{k=1}^{p} \phi_k + \sum_{l=1}^{q} \theta_l < 1\) for the series to be second order stationary with \(p, q \geq 0\). Note that when \(q = 0\), the process \(\text{INGARCH}(p, 0)\) becomes \(\text{INARCH}(p)\), which is similar to the original one, \(\text{ARCH}(p)\) introduced by Engle (1982).

Now consider the process where \(p = q = 1\), i.e. \(\text{INGARCH}(1, 1)\) given by

\[
\mu_t = \beta_0 + \phi Y_{t-1} + \theta \mu_{t-1}, \quad t \geq 1.
\]

Fokianos (2012) suggested employing the following decomposition to study this process. Set

\[
\begin{align*}
Y_t &= \mu_t + Y_{t-1} - \mu_t \\
&= \beta_0 + \phi Y_{t-1} + \theta \mu_{t-1} + \epsilon_t \\
&= \beta_0 + \phi Y_{t-1} + \theta \mu_{t-1} + \epsilon_t + \theta Y_{t-1} - \theta Y_{t-1} \\
&= \beta_0 + (\phi + \theta) Y_{t-1} + \epsilon_t - \theta \epsilon_{t-1},
\end{align*}
\]

where \(\epsilon_t = Y_t - \mu_t\) are uncorrelated white noise sequence with zero mean and a constant variance. Indeed, for \(\{Y_t\}\) is a stationary process, we have:

- constant mean

\[
E[\epsilon_t] = E[Y_{t-1} - \mu_t] = E[E(Y_{t-1} - \mu_t|\mathcal{F}_{t-1})] \\
= E[E(Y_t|\mathcal{F}_{t-1}) - \mu_t] = E[\mu_t - \mu_t] = 0.
\]
• constant variance
\[ V[\varepsilon_t] = V[E(Y_t - \mu_t | F_{t-1})] + E[V(Y_t - \mu_t | F_{t-1})] = 0 + E[V(Y_t | F_{t-1})] = E[\mu_t] = E[Y_t - \varepsilon_t] = E[Y_t], \]
and \( E[Y_t] \) is constant since \( \{Y_t\} \) is assumed to be a stationary process.

• uncorrelated sequence. For \( h > 0 \),
\[ \text{Cov}(\varepsilon_t, \varepsilon_{t+h}) = E[\varepsilon_t\varepsilon_{t+h}] = E[\varepsilon_t E(\varepsilon_{t+h} | F_{t-1})] = 0. \]

After showing that \( \{\varepsilon_t\} \) is a white noise process, we can rewrite Eq. (2.8) as
\[ \left(Y_t - \frac{\beta_0}{1 - (\phi + \theta)}\right) = (\phi + \theta) \left(Y_{t-1} - \frac{\beta_0}{1 - (\phi + \theta)}\right) + \varepsilon_t + \theta \varepsilon_{t-1}. \tag{2.9} \]

This shows that \( \text{INGARCH}(1, 1) \) has the same form as the well known \( \text{ARMA}(1, 1) \), so we can easily derive the properties of the \( \text{INGARCH}(1, 1) \) process. Hence, the \( \{Y_t\} \) is a stationary process with mean
\[ E[Y_t] = \frac{\beta_0}{1 - (\phi + \theta)}, \]
and autocovariance function
\[ \text{Cov}[Y_t, Y_{t+h}] = \begin{cases} \frac{(1-(\phi + \theta)^2 + \phi^2)^E[Y_t]}{1 - (\phi + \theta)^2}, & h = 0, \\ \frac{\phi(1-\theta(\phi + \theta))(\phi + \theta)^{h-1}E[Y_t]}{1 - (\phi + \theta)^2}, & h \geq 1. \end{cases} \]

Knowing that \( \text{Corr}[Y_t, Y_{t+h}] = \frac{\text{Cov}[Y_t, Y_{t+h}]}{V[Y_t]} \), we can show that the autocorrelation (ACF) is given by
\[ \text{Corr}[Y_t, Y_{t+h}] = \frac{\phi(1-\theta(\phi + \theta))(\phi + \theta)^{h-1}}{1 - (\phi + \theta)^2 + \phi^2}, \quad h \geq 1. \]

Note that unless \( \phi = 0 \), \( \text{Var}[Y_t] > E[Y_t] \), therefore the model given in Eq. (2.7) can handle overdispersion. However, with the restriction \( 0 < \phi + \theta < 1 \), the ACF is always non-negative (i.e. \( \text{Corr}[Y_t, Y_{t+h}] \geq 0 \)), which means the process \( \text{INGARCH}(1, 1) \) can only be used when the autocorrelation is positive.

Moreover, since the link function in Eq. (2.7) is the identity link function, including covariates \( X_t \) may lead to computation challenges. Fokianos and Tjstheim (2011) suggested another special case of the model given in Eq. (2.4) that can easily include covariates and handle both positive and negative correlation by choosing the link function as the logarithmic function. For example, if \( \{Y_t : t \in \mathbb{Z}\} \) is a count time series of interest, and assuming that, given the past \( F_{t-1} \), the counts are distributed as Poisson distribution as in Eq. (2.5), that is, \( Y_t | F_{t-1} \sim \text{Poi}(\mu_t) \), they set \( \eta_t = \log(\mu_t) \) as the canonical link process, which is,
\[ \eta_t = \log(\mu_t) = \beta_0 + \phi \log(Y_{t-1} + 1) + \theta \eta_{t-1}, \quad t \geq 1, \]  
(2.10)

where \( \beta_0, \phi, \theta \in \mathbb{R} \).

The functions \( g(\mu_t) \) and \( h(Y_t) \) in Eq. (2.4) are the \( \log(\mu_t) \) and \( \log(Y_t + 1) \), respectively. This idea is found in Zeger and Qaqish (1988) and Benjamin et al. (2003).

Although the Poisson distribution is traditionally the first choice to model count data, other researchers have been considering alternative distributions. The most common second choice to model counts, after the Poisson, is the negative binomial (NB) distribution, which also allows for overdispersion in the data. Davis and Wu (2009) proposed a generalization of the model given in Eq. (2.10) with NB distribution and logit link function. Chen et al. (2016) also considered the NB distribution for modeling count data with \( p \), \( \phi \). In order to obtain the properties of the model of interest was in drawing inference on the unknown parameter vector \( \theta = (\beta_0, \phi, \theta) \) with the components as defined in Eq. (2.4). Then, the log of the partial likelihood based on a random sample as given in Eq. (2.1) is

\[ l(\theta) = \sum_{t=1}^{n} \log f(y_t; \theta_t, \phi | \mathcal{F}_{t-1}) \]

\[ = \sum_{t=1}^{n} \left\{ y_t \theta_t - b(\theta_t) \varphi(\phi) + c(y_t; \phi) \right\}. \]  
(2.13)

Now, as a special case, consider the model given in Eq. (2.7). We are interested in drawing inference on the parameter vector \( \theta = (\beta_0, \phi, \theta) \) with the components as defined in Eq. (2.4). Then, the log of the partial likelihood based on a random sample as given in Eq. (2.1) is

\[ l(\theta) = \sum_{t=1}^{n} \log f(y_t; \theta_t, \phi | \mathcal{F}_{t-1}) \]

\[ = \sum_{t=1}^{n} \left\{ y_t \theta_t - b(\theta_t) \varphi(\phi) + c(y_t; \phi) \right\}. \]  
(2.13)

Now, as a special case, consider the model given in Eq. (2.7). We are interested in drawing inference on the parameter vector \( \theta = (\beta_0, \phi, \theta) \) with the components as defined in Eq. (2.4). Then, the log of the partial likelihood based on a random sample as given in Eq. (2.1) is

\[ l(\theta) = \sum_{t=1}^{n} \log f(y_t; \theta_t, \phi | \mathcal{F}_{t-1}) \]

\[ = \sum_{t=1}^{n} \left\{ y_t \theta_t - b(\theta_t) \varphi(\phi) + c(y_t; \phi) \right\}. \]  
(2.13)

Now, as a special case, consider the model given in Eq. (2.7). We are interested in drawing inference on the parameter vector \( \theta = (\beta_0, \phi, \theta) \) with the components as defined in Eq. (2.4). Then, the log of the partial likelihood based on a random sample as given in Eq. (2.1) is

\[ l(\theta) = \sum_{t=1}^{n} \log f(y_t; \theta_t, \phi | \mathcal{F}_{t-1}) \]

\[ = \sum_{t=1}^{n} \left\{ y_t \theta_t - b(\theta_t) \varphi(\phi) + c(y_t; \phi) \right\}. \]  
(2.13)
where \( \theta_t = \log \eta_t(\theta) \), \( b(\theta_t) = \mu_t(\theta) = \eta_t(\theta) \), and \( c(y_t; \phi) = \log (y_t!) \). The dispersion functional parameter is \( \alpha_t(\phi) = \phi = 1 \) in the Poisson case. Also, recall that from Eq. (2.7)

\[
\eta_t(\theta) = \mu_t(\theta) = \beta_0 + \phi Y_{t-1} + \theta \eta_{t-1}(\theta), \quad t \geq 1.
\]

Eq. (2.14) corresponds to the full likelihood in this case. Hence, the score function is given by

\[
S_n(\theta) = \frac{\partial}{\partial \theta} l(\theta) = \sum_{t=1}^{n} \frac{\partial}{\partial \theta} l_t(\theta) = \sum_{t=1}^{n} \left\{ \frac{\partial}{\partial \theta} Y_t \log \eta_t(\theta) - \frac{\partial}{\partial \theta} \eta_t(\theta) \right\} = \sum_{t=1}^{n} \left\{ \frac{Y_t}{\eta_t(\theta)} - 1 \right\} \frac{\partial}{\partial \theta} \eta_t(\theta),
\]

where,

\[
\frac{\partial}{\partial \theta} \eta_t(\theta) = \left( \frac{\partial}{\partial \beta_0} \eta_t(\theta), \frac{\partial}{\partial \phi} \eta_t(\theta), \frac{\partial}{\partial \theta} \eta_t(\theta) \right)' = \left( 1 + \theta \frac{\partial \eta_{t-1}}{\partial \beta_0}, \ Y_{t-1} + \theta \frac{\partial \eta_{t-1}}{\partial \phi}, \ \eta_{t-1} + \theta \frac{\partial \eta_{t-1}}{\partial \theta} \right)'.
\]

Solving \( S_n(\theta) = 0 \) provides the conditional maximum likelihood estimator of \( \theta \) (Fokianos and Tjstheim 2011). The Hessian matrix is given by

\[
H_n(\theta) = -\left( \frac{\partial}{\partial \theta} \eta_t(\theta) \right) \left( \frac{\partial}{\partial \theta} \eta_t(\theta) \right)' l(\theta) = \sum_{t=1}^{n} Y_t \left( \frac{\partial}{\partial \theta} \eta_t(\theta) \right) \left( \frac{\partial}{\partial \theta} \eta_t(\theta) \right)' - \sum_{t=1}^{n} \left( \frac{Y_t}{\eta_t(\theta)} - 1 \right) \frac{\partial^2}{\partial \theta \partial \theta'} \eta_t(\theta).
\]

3. **Directional dependence by copula**

3.1. **Copula**

When the marginal distributions of random variables are not distributed as normal, using the multivariate normal distribution may not be appropriate. For many known distributions, there are no well defined multivariate extensions to univariate ones like the normal distribution. Hence, the copula is the alternative method to obtain a multivariate distribution of non-
normal marginals. The copula is a multivariate distribution with uniform univariate margins on the unit interval. The idea behind the construction of a multivariate distribution via copula is a probability integral transform. For example, for a continuous random variable, say $Y$, with distribution function $F_Y$, the transformation $F(Y)$ is uniformly distributed on the unit interval $[0,1]$. (See (Nelsen 2007) and (Joe 2014) for more details.)

**Definition 1.** A $p$-dimensional copula is a function $C : [0,1]^p \to [0,1]$ with the following properties:

1. $C(1, \ldots, a_i, \ldots, 1) = a_i \quad \forall \ i = 1,2,\ldots,p$ and $a_i \in [0,1]$.
2. $C(a_1, a_2, \ldots, a_p) = 0$ if at least one $a_i = 0$ for $i = 1,2,\ldots,p$
3. For any $a_i, a_j \in [0,1]$ with $a_i \leq a_j$, for $i = 1,2,\ldots,p$,

$$\sum_{j_1=1}^{2} \sum_{j_2=1}^{2} \cdots \sum_{j_p=1}^{2} (-1)^{j_1+j_2+\cdots+j_p} C(a_{1j_1}, a_{2j_2}, \ldots, a_{nj_p}) \geq 0.$$ (3.1)

Working with Eq. (3.1) to build a multivariate distribution is not straightforward. However, Sklar (1959) provided a theorem that helped build a multivariate distribution of known marginals through a copula, which later became one of the most fundamental theorems in this subject.

**Sklar’s Theorem.** Let $Y_1, Y_2, \ldots, Y_p$ be random variables with marginal distribution functions $F_1, F_2, \ldots, F_p$ and joint cumulative distribution function $F$, then

1. There exists a $p$-dimensional copula $C$ such that for all $y_1, y_2, \ldots, y_p \in \mathbb{R}$

$$F(y_1, y_2, \ldots, y_p) = C(F_1(y_1), F_2(y_2), \ldots, F_p(y_p)).$$ (3.2)

2. If $Y_1, Y_2, \ldots, Y_p$ are continuous, then the copula $C$ is unique. Otherwise, $C$ can be uniquely determined on $p$-dimensional rectangle $\text{Range}(F_1) \times \text{Range}(F_2) \times \cdots \times \text{Range}(F_p)$.

Based on Eq. (3.2), if the marginal distributions $F_i(\cdot)$ for $i = 1,\ldots,p$ are continuous, then the multivariate density function can be defined as

$$f(y_1, y_2, \ldots, y_p) = c(F_1(y_1), F_2(y_2), \ldots, F_p(y_p)) \prod_{i=1}^{p} f_i(y_i), \quad y_i \in \mathbb{R},$$ (3.3)

where $f_i(y_i) = \frac{\partial F_i(y_i)}{\partial y_i}$ is the marginal probability density function (pdf), with $C$ the copula function and $c$ is the copula density, defined as $c(u) = \frac{\partial C(u)}{\partial u}$.

However, when all the margins are discrete, Song (2007) stated that the multivariate probability mass function can be obtained as

$$f(y) = P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_p = y_p) = \sum_{j_1=1}^{2} \sum_{j_2=1}^{2} \cdots \sum_{j_p=1}^{2} (-1)^{j_1+j_2+\cdots+j_p} C(u_{1j_1}, u_{2j_2}, \ldots, u_{nj_p}).$$ (3.4)
where \( u_{j1} = F_j(y_j) \) and \( u_{j2} = F_j(y_j^-) \). The term \( F_j(y_j^-) \) is the left-hand limit of \( F_j \) at \( y_j \), which is equal to \( F_j(y_j) \) when the value of \( y_j \) is an integer. Other authors also considered the case where the marginals are mixed. For example, Czado et al. (2012) applied the mixed model on insurance claim data and Sen et al. (2015) on statistical pattern recognition.

There are numerous copulas available, and one of the most popular copulas in the literature is the Gaussian copula. The Gaussian copula shares many properties of the multivariate normal (Gaussian) distribution, such as the correlation structure. Therefore, we took advantage of the flexibility to manipulate the association structure by using the Gaussian copula.

### 3.2. Directional dependence via beta regression model

#### 3.2.1. Directional dependence

Consider the case where we have a bivariate distribution function, \( F(y_1, y_2) \). This function can be represented as a function of its marginal of \( Y_1 \) and \( Y_2 \), i.e. \( F_1(y_1) \) and \( F_2(y_2) \), by using Eq. (3.2). Specifically,

\[
F(y_1, y_2) = C(F_1(y_1), F_2(y_2)) = C(u_1, u_2),
\]

where \( u_1 \) and \( u_2 \) are uniformly distributed on \([0, 1]\).

Sungur (2005) defined a directional dependence by copula regression setting and provided some measurements of the directional dependence. He also distinguished between the term “directional dependence” and the “direction of the dependence.” The latter is a property of marginal distributions, while the former, which is of interest in this paper, is a property of the joint distribution, or in our case, the copula.

Let \((U_1, U_2)\) be a random pair marginally distributed as uniform on the unit interval \([0, 1]\) with the copula \( C \). Then the conditional distribution function of \( U_1 \), given \( U_2 = u_2 \) denoted by \( C_{u_2}(u_1) \), is

\[
C_{u_2}(u_1) \equiv P(U_1 \leq u_1 | U_2 = u_2) = \frac{\partial C(u_1, u_2)}{\partial u_2}, \tag{3.5}
\]

and the copula regression function of \( U_1 \) on \( U_2 \), as defined by \( r_{U_1|U_2}(u_1) \), is

\[
r_{U_1|U_2}(u_2) \equiv E[U_1 | U_2 = u_2] = 1 - \int_0^1 C_{u_2}(u_1) du_1. \tag{3.6}
\]

Note that if \( U_1 \) and \( U_2 \) are independent, i.e. \( C(u_1, u_2) = u_1 u_2 \), then \( r_{U_1|U_2}(u_2) = r_{U_2|U_1}(u_1) = 1/2 \). The functional forms of the copula regression depend on the choice of the copula function. Here, we will consider an asymmetric copula function, as explained later and use the beta regression model to derive the conditional expectation and show that it can be applied to different types of copula functions. Sungur (2005) also proposed general measurements for quantifying the directional dependence jointly. The directional dependence from \( U_1 \) to \( U_2 \) is expressed by
Table 1. Copula function.

<table>
<thead>
<tr>
<th>Copula</th>
<th>Copula function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$C(u_1, u_2) = \Phi_0(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$, $\theta \in [-1, 1]$</td>
</tr>
<tr>
<td>FGM</td>
<td>$C(u_1, u_2; \theta) = u_1 u_2 + \theta u_1 (1-u_2)(1-u_2)$, $\theta \in [-1, 1]$</td>
</tr>
<tr>
<td>Clayton</td>
<td>$C(u_1, u_2; \theta) = (u_1^{\theta} + u_2^{\theta} - 1)^{-\theta}$, $\theta \in (0, \infty)$</td>
</tr>
<tr>
<td>Frank</td>
<td>$C(u_1, u_2; \theta) = -\frac{1}{\theta} \log[1 + \frac{e^{-u_1 \theta}}{e^{-u_2 \theta}}]$, $\theta \in \mathbb{R} \setminus {0}$</td>
</tr>
<tr>
<td>Gumbel</td>
<td>$C(u_1, u_2; \theta) = \exp[-(-\log(u_1))^\theta + (-\log(u_2))^{1/\theta}]$, $\theta \geq 1$</td>
</tr>
<tr>
<td>Plackett</td>
<td>$C(u_1, u_2; \theta) = \frac{1}{\theta^2} \left[ \log(\theta (u_1 u_2) - \theta (u_1 + u_2) + 1) - \log(\theta - 1) \right]$, $\theta &gt; 0$</td>
</tr>
</tbody>
</table>

Note that $\rho_{U_1 - U_2}^2$ can be interpreted as the proportion of total variation of $U_2$, which can be explained by the copula regression of $U_2$ on $U_1$. We can also see that the variance $\text{Var}(r_{U_2|U_1}(U_1))$ is the expected square distance of the copula regression function, from the independence case, in the direction of $U_1$ to $U_2$. Thus, the predictive power of the regression is higher for the direction that corresponds to the smaller variance. For example, if $\text{Var}(r_{U_2|U_1}(U_1)) < \text{Var}(r_{U_1|U_2}(U_2))$, then the predictive power of the regression in the direction $U_1$ to $U_2$ is higher than $U_2$ to $U_1$.

Kim and Hwang (2017b) first proposed the use of Gaussian copula beta regression (Guolo and Varin 2014) to study the directional dependence of two time series processes that have asymmetric GARCH marginals. They again used the beta technique, but with stochastic volatility marginals, to investigate the directional data in Kim and Hwang (2017a). Other types of asymmetric copula were applied, such as Farlie-Gumbel-Morgenstern (FGM) copula, to test directional dependence of financial data and gene data (for examples, see Kim et al. 2008; Jung et al. 2008; Kim et al. 2009; Uhm et al. 2012).

Furthermore, a general family of copulas with asymmetric members were introduced by Durante (2009) via combining two symmetric copulas such as the ones shown in Table 1,

$$C(u_1, u_2; \phi) = C_1(u_1^\alpha, u_2^\beta) C_2(u_1^\beta, u_2^\alpha),$$  \hspace{1cm} (3.7)

where $\phi = (\theta, \alpha, \beta), \alpha, \beta \in (0, 1), \alpha \neq 1/2$ or $\beta \neq 1/2, \alpha + \beta = 1, \beta + \bar{\beta} = 1,$ and $C_1$ and $C_2$ are two symmetric copula families. $\theta$ is known as the dependence parameter and $\alpha$ and $\beta$ as the asymmetric parameters. Based on this asymmetric copula, Kim and Kim (2014) proposed a generalized multiple-step directional dependence in joint behavior based on asymmetric copula regression. However, due to the computational burden associated with the latter method, they suggested the use of Gaussian copula beta regression model to handle the asymmetry in Kim and Hwang (2017b).

### 3.2.2. Gaussian copula beta regression model

Let $\{Y_t : t = 1, \ldots, n\}$ be a time series of interest, and its components are bounded on the unit interval $(0, 1)$, and let $X_t$ be a vector of $p$ covariates. The beta regression
assumes that given $X_t$, the response variable $Y_t$ is distributed as $Beta(\mu_t, \kappa_t)$ (Paolino 2001) with probability density function

$$f(y_t; \mu_t, \kappa_t) = \frac{\Gamma(\kappa_t)}{\Gamma(\mu_t \kappa_t) \Gamma((1 - \mu_t) \kappa_t)} y_t^{\mu_t-1} (1-y_t)^{(1-\mu_t)\kappa_t-1},$$

(3.8)

where $0 < \mu_t < 1$ is the mean parameter, $\kappa_t > 0$ is the precision parameter, and $\Gamma(.)$ is the Gamma function. Also, $\text{Var}(Y_t) = \mu_t (1-\mu_t)/(1+\kappa_t)$ and the subscript $t$ emphasizes the time dependence of the beta density through $\mu_t$ and $\kappa_t$. Under the frame work of GLM, the dependence between the response $Y_t$ and the covariates $x_t$ can be captured through linking the systematic component $X_t' \beta_x$ to the mean $\mu_t$ via the logit function. That is,

$$\text{logit}(\mu_t) = X_t' \beta_x,$$

(3.9)

or $\log (\kappa_t - 1) = x_t' \beta_x$, in which $\beta_x$ is a $p$-dimensional vector of coefficients.

Guolo and Varin (2014) developed copula-based marginal extension of the beta regression for time series analysis that has advantage over observation-driven and parameter-driven models in terms of interpretation and computational difficulties. Their proposed method, beta regression, employed the probability integral transformation to relate the response variable $Y_t$ to the covariates vector $X_t$ and a standard normal error $\epsilon_t$ as stated in Masarotto and Varin (2012)

$$Y_t = F_t^{-1}\{\Phi(\epsilon_t); \beta_x\},$$

(3.10)

where $F_t^{-1}(.; \beta_x)$ is the cumulative distribution function of $Y_t$.

By applying the Gaussian copula marginal regression of Masarotto and Varin (2012), we can obtain an estimate of the parameter vector $\beta_x$ by maximizing the following closed form likelihood function

$$L(\theta) = \phi_n(\epsilon_1, \ldots, \epsilon_n; R) \prod_{t=1}^{n} f(y_t | X_t; \beta_x) / \phi(\epsilon_t),$$

(3.11)

where, $\phi_n(.; R)$ is the $n$-dimensional standard normal density function with correlation matrix $R$, and $\phi(.)$ is the standard univariate normal density function. $f(., | X_t; \beta_x)$ is beta density function given in Eq. (3.8).

In the next section, directional dependence measurements, given in Sec. 3.2.1, of two discrete time series will be derived by using the beta logit function shown in Eq. (3.9).

### 3.2.3. Directional dependence measurement from beta regression

Let $Y_{1t}$ and $Y_{2t}$ be two discrete time series, and assume that we fit models that followed the general form given in Eq. (2.4) to generate standardized residuals $\epsilon_{1t}$ and $\epsilon_{2t}$. Working with the standardized residuals helps avoid serial dependence in the components of the time series (Kim and Hwang 2017b). We then transformed the two sets of residuals, $\epsilon_{1t}$ and $\epsilon_{2t}$, to two uniformly distributed variables, $U_{1t}$ and $U_{2t}$, respectively, in $[0, 1]$ using Eq. (3.10), and perform the directional dependence via Gaussian copula marginal beta regression model.
More specifically, assume that $U_{1t}$ given $U_{2t} = u_{2t}$ follows a Beta distribution with mean parameter $\mu_{1t}$ and precision parameter $\kappa_{1t}$. Let $F(U_{1t}|\theta)$ be the distribution function of the beta random variable with mean $\mu_{1t} = E[U_{1t}|U_{2t}]$. Dependence of $U_{1t}$ on $U_{2t}$ from a random sample of size $n$ is obtained by assuming a logit model for the mean parameter, which is

$$
\text{logit}(\mu_{1t}) = \log \left( \frac{\mu_{1t}}{1 - \mu_{1t}} \right) = \beta_0 + \beta_1 u_{2t}, \quad \text{for } t = 1, \ldots, n, \quad (3.12)
$$

so that

$$
\mu_{1t} = \frac{\exp (\beta_0 + \beta_1 u_{2t})}{1 + \exp (\beta_0 + \beta_1 u_{2t})} \quad \text{and} \quad \kappa_{1t} = 1 + \exp (\beta_0 + \beta_1 u_{2t}),
$$

with correlation matrix of the errors corresponding to the white noise process.

Now, the directional dependence measurements, provided in Subsection 3.2.1 is given by

$$
\rho_{Y_{1t} \rightarrow Y_{1t}}^2 = \frac{\text{Var}(E[U_{1t}|u_{2t}])}{\text{Var}(U_{1t})} = 12\text{Var}(\mu_{1t}) = 12\sigma_{\mu_1}^2, \quad (3.13)
$$

where $\sigma_{\mu_1}^2$ is the variance that was obtained from the given data of $\mu_{1t}$ for $t = 1, \ldots, n$.

4. Data analysis

4.1. Simulated example

To verify the proposed method of measuring the directional dependence of two count time series variables, we carried out a simulation in the statistical software R (Team 2013) of two correlated INGARCH(1, 1) processes following Poisson distribution with means $\lambda_{1t}$ and $\lambda_{2t}$, respectively, for $t = 1, \ldots, n \ (n = 10000)$. To insure that the two processes were correlated and was asymmetry, we generated a pair of uniform random variables $(u_1, u_2)$ on the unit interval $[0, 1]$ from the copula family defined in Eq. (3.7) with the symmetric members chosen to be Gaussian copula ($\theta = 0.99$) × Gaussian copula ($\theta = 0.1$) with $\alpha = 0.7$ and $\beta = 0.3$ using the R package” copBasic” (Asquit, 2017). Then we generated the two INGARCH(1, 1) processes using the pairs $(u_1, u_2)$, which is

$$
Y_{1t} = F^{-1}(u_{1t}; \lambda_{1t}) \quad \text{and} \quad Y_{2t} = F^{-1}(u_{2t}; \lambda_{2t}),
$$

where $F^{-1}$ is the inverse of the Poisson distribution function, and the time dependent means are given by

$$
\lambda_{1t} = \beta_1 + \phi_1 Y_{1(t-1)} + \theta_1 \lambda_{1(t-1)} \quad \text{and} \quad \lambda_{2t} = \beta_2 + \phi_2 Y_{2(t-1)} + \theta_2 \lambda_{2(t-1)}.
$$

We chose the parameters true values to be: $\beta_1 = 1, \phi_1 = 0.6, \theta_1 = 0.3, \beta_2 = 0.8, \phi_2 = 0.2, \text{ and } \theta_2 = 0.7$. Following the steps explained in Subsection 3.2.3, this parameterization leads to having a pair with Pearson correlation $\rho = 0.38 \ (\rho = 0.15)$, and the targeted measurements of the directional dependence are $\rho_{Y_{1t} \rightarrow Y_{1t}}^2 = 0.1439$ and $\rho_{Y_{1t} \rightarrow Y_{2t}}^2 = 0.1247 \ (\rho_{\text{diff}}^2 = \rho_{Y_{1t} \rightarrow Y_{1}}^2 - \rho_{Y_{1t} \rightarrow Y_{2}}^2 = 0.0192)$.

Figure 1 shows the scatterplot of the two Poisson INGARCH(1, 1) processes $Y_{1t}$ and $Y_{2t}$. The plot shows that the relationship is not linear and asymmetry is observed. Thus,
a linear approach to investigate the directional dependence (see Wiedermann et al. (2015) for an example) is not appropriate here. Instead, a more plausible approach is the beta regression model with logit function suggested by Kim and Hwang (2017b) to derive the directional dependence measurements.

We generated 500 simulated datasets for each of the three sample sizes, $n = 100, 500, and 1000$. To examine the proposed method for measuring the directional dependence of two discrete time series processes, we calculated the mean and standard error for each one of the parameters given in this vector $\theta = (\beta_1, \phi_1, \theta_1, \beta_2, \phi_2, \theta_2, \rho_{Y_2-Y_1}, \rho_{Y_1-Y_2}, \rho_{\text{diff}})'$. We also calculated the bias and mean square error (MSE) as defined by

$$
\text{Bias}(\hat{\theta}, \theta) = E[\hat{\theta}(\theta) - \theta]
$$

and

$$
\text{MSE}(\hat{\theta}) = E[\hat{\theta}(\theta)^2].
$$

The simulation results are shown in Table 2. The parameter estimates of those from the INGARCH processes were obtained using the R package “tscount” (Liboschik et al. 2015), whereas the one corresponding to copula directional dependence via beta regression model were obtained using the R package “gcmr” (Masarotto and Varin 2017). For $n = 100$, the estimates corresponding to the marginals, i.e. INGARCH processes, resulted a large bias and mean square errors relatively, while the directional dependence estimates resulted in smaller errors even when the marginals were sometimes misspecified. However, when the sample size is larger ($n = 500, 1000$), the results show improvement and there is consistency for both the parameter estimates coming from the marginals and the one describing the joint behavior through the directional dependence. We also
obtained bootstrap 95% confidence intervals of the parameter $\hat{q}_{diff}^2$ at each sample size to test whether the directional dependence exists or not. For all of the suggested sample sizes, the confidence intervals do not include zero, which suggests that $\hat{q}_{diff}^2$ is significantly different from zero, and the directional dependence exists in our simulated example as suggested in Sungur (2005).

Finally, we repeated the same type of simulation, but with different copula functions, to construct the symmetric copula given in Eq. (3.7). In that case, we chose symmetric members to be Plackett copula ($\theta = 5000$) × Plackett copula ($\theta = 2.5$) with $z = 0.68$ and $\beta = 0.47$. The marginal parameters were chosen to be the same as before, which resulted in measurements of the directional dependence as $\rho_{Y_2 \rightarrow Y_1}^2 = 0.2820$ and $\rho_{Y_1 \rightarrow Y_2}^2 = 0.2555$ ($\rho_{diff}^2 = \rho_{Y_2 \rightarrow Y_1}^2 - \rho_{Y_1 \rightarrow Y_2}^2 = 0.0265$). Table 3 shows comparable results to the one with Gaussian members, which proves the directional dependence is well captured by the beta regression. This suggests that when we apply the beta regression model on the pair $(u_{1t}, u_{2t})$, as explained above, we do not need to make assumption on the symmetric members that construct the joint distribution function of $(u_{1t}, u_{2t})$. Further numerical comparisons were also done for the different types of copula functions shown in Table 1 and the results were all consistent.

### Table 2. Summary statistics for the 500 simulations from Gaussian copula members.

<table>
<thead>
<tr>
<th>n</th>
<th>Parameters</th>
<th>Mean</th>
<th>SE</th>
<th>Bias</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>$\beta_1$</td>
<td>1.380</td>
<td>0.8185</td>
<td>0.3804</td>
<td>0.81448</td>
</tr>
<tr>
<td></td>
<td>$\phi_1$</td>
<td>0.6344</td>
<td>0.1085</td>
<td>0.0344</td>
<td>0.0129</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.1929</td>
<td>0.1359</td>
<td>-0.1070</td>
<td>0.0299</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>1.2454</td>
<td>1.3315</td>
<td>0.4454</td>
<td>1.9715</td>
</tr>
<tr>
<td></td>
<td>$\phi_2$</td>
<td>0.2715</td>
<td>0.0981</td>
<td>0.0715</td>
<td>0.0147</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.5501</td>
<td>0.2218</td>
<td>-0.1498</td>
<td>0.0716</td>
</tr>
<tr>
<td></td>
<td>$\rho_{Y_2 \rightarrow Y_1}^2$</td>
<td>0.1538</td>
<td>0.0714</td>
<td>0.0099</td>
<td>0.0052</td>
</tr>
<tr>
<td></td>
<td>$\rho_{Y_1 \rightarrow Y_2}^2$</td>
<td>0.1357</td>
<td>0.0669</td>
<td>0.0110</td>
<td>0.0046</td>
</tr>
<tr>
<td>Bootstrap 95% CI of $\hat{q}_{diff}^2$ :</td>
<td>(0.0127,0.0261)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>$\beta_1$</td>
<td>1.0291</td>
<td>0.2433</td>
<td>0.0291</td>
<td>0.0601</td>
</tr>
<tr>
<td></td>
<td>$\phi_1$</td>
<td>0.6008</td>
<td>0.0425</td>
<td>0.0008</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.2872</td>
<td>0.0526</td>
<td>-0.0127</td>
<td>0.0029</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>0.8734</td>
<td>0.2995</td>
<td>0.0734</td>
<td>0.0951</td>
</tr>
<tr>
<td></td>
<td>$\phi_2$</td>
<td>0.2164</td>
<td>0.0372</td>
<td>0.0164</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.6689</td>
<td>0.0630</td>
<td>-0.0310</td>
<td>0.0049</td>
</tr>
<tr>
<td></td>
<td>$\rho_{Y_2 \rightarrow Y_1}^2$</td>
<td>0.1465</td>
<td>0.0333</td>
<td>0.0026</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>$\rho_{Y_1 \rightarrow Y_2}^2$</td>
<td>0.1265</td>
<td>0.0298</td>
<td>0.0018</td>
<td>0.0008</td>
</tr>
<tr>
<td>Bootstrap 95% CI of $\hat{q}_{diff}^2$ :</td>
<td>(0.0125,0.0258)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>$\beta_1$</td>
<td>1.0179</td>
<td>0.1647</td>
<td>0.0179</td>
<td>0.0274</td>
</tr>
<tr>
<td></td>
<td>$\phi_1$</td>
<td>0.5979</td>
<td>0.0300</td>
<td>-0.0020</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>0.2944</td>
<td>0.0377</td>
<td>-0.0055</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>0.8759</td>
<td>0.5774</td>
<td>0.0759</td>
<td>0.3392</td>
</tr>
<tr>
<td></td>
<td>$\phi_2$</td>
<td>0.2059</td>
<td>0.0286</td>
<td>0.0059</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>$\theta_2$</td>
<td>0.6812</td>
<td>0.0678</td>
<td>-0.0187</td>
<td>0.0049</td>
</tr>
<tr>
<td></td>
<td>$\rho_{Y_2 \rightarrow Y_1}^2$</td>
<td>0.1435</td>
<td>0.0219</td>
<td>-0.0003</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>$\rho_{Y_1 \rightarrow Y_2}^2$</td>
<td>0.1241</td>
<td>0.0195</td>
<td>-0.0006</td>
<td>0.0003</td>
</tr>
<tr>
<td>Bootstrap 95% CI of $\hat{q}_{diff}^2$ :</td>
<td>(0.0127,0.0254)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Bootstrap 95% CI of $\hat{q}_{diff}^2$ : | (0.0127,0.0261) |
| Bootstrap 95% CI of $\hat{q}_{diff}^2$ : | (0.0125,0.0258) |
| Bootstrap 95% CI of $\hat{q}_{diff}^2$ : | (0.0127,0.0254) |
4.2. Real data example

The data set used in this study consists of two time series processes of weather phenomena. The first process was the monthly count of strong sandstorms (SA) and the second...
process was the monthly count of dust haze (HZ) recorded by AQI airport station in Eastern Province, Saudi Arabia, which happens to be located in one of the major dust producing regions (Idso 1976). Both count time series processes spanned from January 1978 to December 2013. The main objective was to determine the proposed method and which time series influenced the other more through the directional dependence measurement given in Eq. (3.13) for count data. Table 4 summarizes the findings of the standardized residuals of the chosen models. Both series correspond to zero mean and one standard deviation. However, the SA count series was skewed more to the right relative to the normal distribution than the HZ series. In addition, the kurtosis of the SA series was substantially higher that the one corresponding to the HZ series, which indicated that the former had heavier tail relative to the normal distribution. Also, the high kurtosis suggests there was asymmetry in the process (Kim and Hwang 2017b). Finally, the autoregressive conditional heteroscedasticity (ARCH) tests at 4, 8, and 12 lags resulted in p-values significantly smaller than the level of significance 0.05, which suggested rejecting the null hypothesis of homoscedasticity in the data.

Table 5 reveals the parameter estimates of the beta regression models with the logit link function we used following the proposed method in Subsection 3.2. Using these estimates, we were able to calculate the measurements of directional dependence that we were interested in, i.e. $\rho^2_{SA\rightarrow HZ}$ and $\rho^2_{HZ\rightarrow SA}$.

Table 6 shows the directional dependence measurements. We can see that $\rho^2_{SA\rightarrow HZ}$ is greater than $\rho^2_{HZ\rightarrow SA}$, which suggests that directional dependence exists here. Also, it indicates that sand storms cause dust haze more than the other way around. In fact, the dust haze consisted of small sand particles suspended in the air remained from a recent sand storm (Boloorani et al. 2014), and our measurements agree with this fact.

5. Conclusion

Dependence of bivariate discrete time series data is captured through the residuals using copula directional dependance from Beta regression. The proposed model of analysis of the discrete data is developed and estimates are derived. Using simulated and real life data with different forms of copulas, directional dependence estimates are generated, and applied to the real life data where it confirms an existing conclusion on how sand-storms affect dust haze phenomena more than the other way around. Extension of this work can be obtained by considering time varying directional dependence regression measurements of the bivariate discrete time series via copula.

Acknowledgment

The authors thank the Editor and the anonymous referees for the constructive comments, which led to considerable improvement of the manuscript. The authors thank the Saudi General Authority of Meteorological & Environmental Protection for making the sandstorm data available.
**Funding**

The first author research was partially supported by Qassim University, Saudi Arabia.

**ORCID**

Norou Diawara http://orcid.org/0000-0002-8403-6793

**References**


Boolorani, D. A. N. O. S., H. Bahrami, S. Alvipanah, H. Mohammadi, and M. Nezammahalleh. 2014. *Dust storms in the west asia region*. Vice President for Science and Technology of Iran and Geoinformatics Research Institute (GRI) of the University of Tehran, Teheran.


