

# Circular Functional Relationship Model with Wrapped Cauchy Errors

Ali H. Abuzaid

Department of Mathematics, Faculty of Science, Al Azhar University – Gaza, Palestine  
alizaid33@yahoo.com

Walaa A. Abu El-laban

Department of Applied Statistics, Faculty of Economic and Administrative Sciences  
Al Azhar University – Gaza, Palestine  
alaqsa-w@hotmail.com

Abdul Ghapor Hussin

Department of Defence Science, Faculty of Defence Science and Technology  
National Defence University of Malaysia, Malaysia  
ghapor@upnm.edu.my

## Abstract

This paper extends the simple linear regression model with wrapped Cauchy error to the functional case when both variables are subjected to wrapped Cauchy errors. Assuming the ratio between the two error variances is known and the slope parameter equals one the maximum likelihood estimates are obtained. The closed-form expression for the maximum likelihood estimators are not available and the estimates are obtained iteratively by choosing a suitable initial values. The quality of estimates and the accuracy of the model are illustrated via simulations and the results revealed an acceptable performance of the estimators where they are unbiased, consistent and robust. The sampling variances of the model parameters are obtained via bootstrapping methods and consequently the confidence intervals were constructed. The proposed model is illustrated with an application on the analysis of wind directions data at two cities in the Gaza strip, Palestine.

**Keywords:** Bootstrap, EIVM, robustness, Wind direction.

## 1. Introduction

The functional relationship model is a part of the general class of Error-In-Variables Models (EIVM) which also known as the measurement error or random regression models. The study of EIVM had been firstly explored by (Adcock 1877, 1878) and then (Kendall 1951, 1952) formally made a distinction between functional and structural relationship between the two variables. In the ordinary regression we supposed that the explanatory variables measured without error and all the errors are in the response variable, but the EIVM assumed that the errors are in both response and explanatory variables. Hence, there is no distinction between the response and explanatory variables.

The extension of EIVM to the case when both variables are circular has not received much attention. The un-replicated linear functional relationship for two circular variables was first introduced by (Hussin 1997) assuming that errors of both variables are independently distributed with von Mises distribution with equal variances.

Later, (Hussin 2003) improved the model by estimating the concentration parameters of error for any ratio  $\lambda$  using the asymptotic properties of the Bessel function. Then,

(Hussin 2005) extended the model for replicated observations and considered the Fisher information matrix of model parameters.

(Caires and Wyatt 2003) proposed the simple linear circular functional relationship model in which the value of the slope parameter  $\beta$  is fixed to be one, and the errors are assumed to follow von Mises distribution for any ratio  $\lambda$ . (Hussin et al 2010) derived the variances of model parameters by deriving the asymptotic covariance matrix and developed an outlier detection procedure. (Satari, et al 2014) developed a new functional relationship model for circular variables by extending the (Downs and Mardia 2002) circular-circular regression model.

The heavily tailed property of the wrapped Cauchy distribution motivated (Abuzaid and Allahham 2015) to propose the simple linear regression for circular variables in the following form:

$$y_i = \alpha + \beta x_i + \varepsilon_i \pmod{2\pi},$$

where  $\varepsilon_i$  circular random error having a wrapped Cauchy distribution with circular mean 0 and concentration parameter  $\rho$ . The estimates of the parameters were obtained based on maximizing the log-likelihood function via iterative procedure. The Results show that their model is more robust to the deviation of assumptions compare to (Hussin et al 2004).

The rest of this paper is organized as follows, Section 2 formulates the proposed model and derives parameter estimators. Section 3 performs an extensive simulation studies to investigate the properties of estimators. Section 4 applies the proposed model on wind directions data from Palestine.

## 2. Circular Functional Relationship Model with Wrapped Cauchy Errors

The robustness of circular regression model proposed by (Abuzaid and Allahham 2015) motivates to extend it to the functional case, where the relationship is linear. Such situation can be found in the calibration between two instruments for measuring circular variables

Suppose  $x_i$  and  $y_i$  for  $i=1,2,\dots,n$  where  $0 < x_i, y_i \leq 2\pi$  are observed values of the true values of circular variables  $X$  and  $Y$  respectively. Assuming that there is a linear relationship between these two variables with known slope parameter equals one. For any fixed values  $X_i$  and  $Y_i$ , we assume that the observations  $x_i$  and  $y_i$  are measured with errors  $\delta_i$  and  $\varepsilon_i$ , respectively and thus the full model can be written as

$$\begin{aligned} x_i &= X_i + \delta_i, \quad \text{and} \quad y_i = Y_i + \varepsilon_i, \\ Y_i &= \alpha + X_i \pmod{2\pi}, \quad i=1,\dots,n, \end{aligned} \tag{1}$$

we also assume that  $\delta_i$  and  $\varepsilon_i$  are independently distributed with wrapped Cauchy distributions with mean zero and concentration parameters  $\rho$  and  $\kappa$ , respectively, that is,  $\delta_i \sim WC(0, \rho)$  and  $\varepsilon_i \sim WC(0, \kappa)$ . We also defined  $\lambda = \frac{\rho}{\kappa}$  (i.e.  $\rho = \lambda\kappa$ ) as a ratio

of error concentration parameters to be known for the proposed circular functional relationship model.

In the context of functional relationship models there are strong reasons for fixing  $\beta = 1$ , the first reason is related to the desired symmetry of the functional relationship model, where the conclusions should be independent of which quantity is chosen to be  $X$  or  $Y$ . Furthermore, the function  $X \rightarrow \beta X + \alpha \pmod{2\pi}$  does not vary continuously when going from  $2\pi$  to  $0$ , for further discussion see (Caires and Wyatt 2003).

## 2.1 Parameter estimation:

The probability density function of error in (1) is given by:

$$f(\alpha, \rho, \kappa, X_i; x_i, y_i) = \frac{1}{(2\pi)^2} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(x_i - X_i)} \frac{1 - \kappa^2}{1 + \kappa^2 - 2\kappa \cos(y_i - \alpha - X_i)}$$

$$= \frac{1}{(2\pi)^2} \left( \frac{1 - \lambda^2 \kappa^2}{1 + \lambda^2 \kappa^2} \right) \left( 1 - \frac{2\lambda \kappa \cos x_i \cos X_i}{1 + \lambda^2 \kappa^2} - \frac{2\lambda \kappa \sin x_i \sin X_i}{1 + \lambda^2 \kappa^2} \right)^{-1} \left( \frac{1 - \kappa^2}{1 + \kappa^2} \right) \left( 1 - \frac{2\kappa \cos \alpha \cos(y_i - X_i)}{1 + \kappa^2} \right)^{-1}. \quad (2)$$

By introducing the following re-parameterization

$$\alpha_1 = \frac{2\kappa \cos \alpha}{1 + \kappa^2}, \quad \alpha_2 = \frac{2\kappa \sin \alpha}{1 + \kappa^2}, \quad \beta_1 = \frac{2\lambda \kappa \cos X_i}{1 + \lambda^2 \kappa^2}, \quad \text{and} \quad \beta_2 = \frac{2\lambda \kappa \sin X_i}{1 + \lambda^2 \kappa^2}.$$

The probability density function in (2) can be written in the form

$$f(\alpha, \rho, \kappa, X_i; x_i, y_i) = \frac{1}{(2\pi)^2} \frac{1}{d(1 - \beta_1 \cos x_i - \beta_2 \sin x_i)} \frac{1}{c(1 - \alpha_1 \cos(y_i - X_i) - \alpha_2 \sin(y_i - X_i))}, \quad (3)$$

$$\text{where } c = \frac{1}{\sqrt{1 - \alpha_1^2 - \alpha_2^2}} \quad \text{and} \quad d = \frac{1}{\sqrt{1 - \beta_1^2 - \beta_2^2}}.$$

Letting  $\eta_1 = c\alpha_1$ ,  $\eta_2 = c\alpha_2$ ,  $\tau_1 = d\beta_1$ , and  $\tau_2 = d\beta_2$ , then errors in model (1) can be formulated as follows:

$$f(\alpha, \rho, \kappa, X_i; x_i, y_i) = \frac{1}{(2\pi)^2} \frac{1}{d - \tau_1 \cos x_i - \tau_2 \sin x_i} \frac{1}{c - \eta_1 \cos(y_i - X_i) - \eta_2 \sin(y_i - X_i)}, \quad (4)$$

$$\text{where } d = \sqrt{1 + \tau_1^2 + \tau_2^2} \quad \text{and} \quad c = \sqrt{1 + \eta_1^2 + \eta_2^2}.$$

Then the log-likelihood function for model (1) is given by:

$$\log L = \log \prod_{i=1}^n f(\eta_1, \eta_2, \tau_1, \tau_2, X_i; x_i, y_i)$$

$$= -2n \log 2\pi - \sum_{i=1}^n \log(d - \tau_1 \cos x_i - \tau_2 \sin x_i) - \sum_{i=1}^n \log(c - \eta_1 \cos(y_i - X_i) - \eta_2 \sin(y_i - X_i)). \quad (5)$$

Assuming the equality of error concentration parameters, i.e.  $\lambda = 1$ , there are  $(n + 2)$  parameters to be estimated which are  $\alpha$ ,  $\kappa$  and  $X_i$  for  $i = 1, 2, \dots, n$ . The maximum likelihood estimates of  $\eta_1$  and  $\eta_2$  are obtained by differentiating the log-likelihood

function in (5) with respect to  $\eta_1$  and  $\eta_2$ , respectively and equating them with zero as follows:

$$\frac{\partial \log L}{\partial \eta_1} = \frac{1}{c} \sum w_i (\cos(y_i - X_i) - \alpha_1) = 0 \quad (6)$$

where  $w_i = \frac{1}{1 - \alpha_1 \cos(y_i - X_i) - \alpha_2 \sin(y_i - X_i)}$ , similarly for  $\eta_2$ , we have

$$\frac{\partial \log L}{\partial \eta_2} = \frac{1}{c} \sum w_i (\sin(y_i - X_i) - \alpha_2) = 0. \quad (7)$$

#### a) Estimation of $\alpha$ :

From Equations (6) and (7) we have

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^n w_i \cos(y_i - \hat{X}_i)}{\sum_{i=1}^n w_i} \text{ and } \hat{\alpha}_2 = \frac{\sum_{i=1}^n w_i \sin(y_i - \hat{X}_i)}{\sum_{i=1}^n w_i}. \quad (8)$$

From the definition of  $\alpha_1$  and  $\alpha_2$  we get  $\tan \hat{\alpha} = \frac{\hat{\alpha}_2}{\hat{\alpha}_1}$ , then, the estimate of  $\alpha$  is obtained as follows:

$$\hat{\alpha} = \begin{cases} \tan^{-1} \left( \hat{\alpha}_2 / \hat{\alpha}_1 \right), & \hat{\alpha}_1 > 0, \hat{\alpha}_2 > 0 \\ \tan^{-1} \left( \hat{\alpha}_2 / \hat{\alpha}_1 \right) + \pi, & \hat{\alpha}_1 < 0 \\ \tan^{-1} \left( \hat{\alpha}_2 / \hat{\alpha}_1 \right) + 2\pi, & \hat{\alpha}_1 > 0, \hat{\alpha}_2 < 0 \\ \text{undefined}, & \hat{\alpha}_1 = 0, \hat{\alpha}_2 = 0 \end{cases} \quad (9)$$

#### b) Estimation of $\kappa$ :

The estimate of  $\kappa$  can be obtained after easy mathematical formulation as given below,

$$\hat{\kappa} = \frac{1 - \sqrt{1 - \hat{\alpha}_1^2 - \hat{\alpha}_2^2}}{\sqrt{\hat{\alpha}_1^2 + \hat{\alpha}_2^2}} \quad (10)$$

#### c) Estimation of $X_i$ :

The maximum likelihood estimates of  $\tau_1$  and  $\tau_2$  are obtained by differentiating the log-likelihood function in (5) with respect to  $\tau_1$  and  $\tau_2$ , respectively and equating them with zero as follows:

$$\frac{\partial \log L}{\partial \tau_1} = \frac{1}{d} \sum m_i (\cos x_i - \beta_1) = 0 \quad (11)$$

where  $m_i = \frac{1}{1 - \beta_1 \cos x_i - \beta_2 \sin x_i}$ , similarly for  $\tau_2$ , we have

$$\frac{\partial \log L}{\partial \tau_2} = \frac{1}{d} \sum m_i (\sin x_i - \beta_2) = 0. \quad (12)$$

From Equations (11) and (12) we have

$$\beta_1 = \sum_{i=1}^n m_i \cos x_i / \sum_{i=1}^n m_i \text{ and } \beta_2 = \sum_{i=1}^n m_i \sin x_i / \sum_{i=1}^n m_i. \quad (13)$$

The first derivative of the log-likelihood function in (5) with respect to  $X_i$  is given by:

$$\begin{aligned} \frac{\partial \log L}{\partial X_i} &= - \sum \frac{\frac{2\kappa \sin X_i \cos x_i}{1-\kappa^2} - \frac{2\kappa \cos X_i \sin x_i}{1-\kappa^2}}{d - \tau_1 \cos x_i - \tau_2 \sin x_i} - \sum \frac{-\eta_1 \sin(y_i - X_i) + \eta_2 \cos(y_i - X_i)}{c - \eta_1 \cos(y_i - X_i) - \eta_2 \sin(y_i - X_i)} \\ &= - \left( \frac{2\kappa}{1-\kappa^2} \right) \left( \frac{1}{d} \right) \sum m_i \sin(X_i - x_i) - \frac{1}{c} \sum w_i (-\eta_1 \sin(y_i - X_i) + \eta_2 \cos(y_i - X_i)), \end{aligned}$$

where  $m_i = \frac{1}{1 - \beta_1 \cos x_i - \beta_2 \sin x_i}$ .

$$\frac{\partial \log L}{\partial X_i} = - \left( \frac{2\kappa}{1-\kappa^2} \right) \sum m_i \sin(X_i - x_i) + \frac{1}{c} \sum w_i (\eta_1 \sin(y_i - X_i) - \eta_2 \cos(y_i - X_i)) = 0. \quad (14)$$

The previous equation can be solved iteratively given some initial guesses for  $X_i$ .

Suppose  $\hat{X}_{i0}$  is an initial estimate for  $\hat{X}_i$ .

$$X_i - x_i = X_i - X_{i0} + X_{i0} - x_i = X_{i0} - x_i + \Delta_i,$$

where,  $\Delta_i = X_i - X_{i0}$ . Also we have  $y_i - X_i = y_i - X_{i0} + X_{i0} - X_i = y_i - X_{i0} - \Delta_i$ .

Since,

$$\sin(X_i - x_i) = \sin(X_{i0} - x_i) \cos \Delta_i + \cos(X_{i0} - x_i) \sin \Delta_i, \quad (15)$$

$$\sin(y_i - X_i) = \sin(y_i - X_{i0}) \cos \Delta_i - \cos(y_i - X_{i0}) \sin \Delta_i. \quad (16)$$

$$\text{and } \cos(y_i - X_i) = \cos(y_i - X_{i0}) \cos \Delta_i + \sin(y_i - X_{i0}) \sin \Delta_i, \quad (17)$$

For small  $\Delta_i$ , we have  $\cos \Delta_i \approx 1$  and  $\sin \Delta_i \approx \Delta_i$ .

So, Equations (15, 16 and 17) become

$$\sin(X_i - x_i) = \sin(X_{i0} - x_i) + \cos(X_{i0} - x_i) \Delta_i, \quad (18)$$

$$\sin(y_i - X_i) = \sin(y_i - X_{i0}) - \cos(y_i - X_{i0})\Delta_i, \quad (19)$$

$$\text{and} \quad \cos(y_i - X_i) = \cos(y_i - X_{i0}) + \sin(y_i - X_{i0})\Delta_i. \quad (20)$$

Hence, (14) is simplified to:

$$\begin{aligned} \frac{\partial \log L}{\partial X_i} = & -\left(\frac{2\lambda\kappa}{1+\kappa^2}\right) \sum m_i \sin(X_{i0} - x_i) - \Delta_i \left(\frac{2\lambda\kappa}{1+\kappa^2}\right) \sum m_i \cos(X_{i0} - x_i) \\ & - \Delta_i \frac{1}{c} \sum w_i [\eta_1 \cos(y_i - X_{i0}) + \eta_2 \sin(y_i - X_{i0})] + \frac{1}{c} \sum w_i [\eta_1 \sin(y_i - X_{i0}) - \eta_2 \cos(y_i - X_{i0})] = 0 \end{aligned}$$

Then,

$$\hat{X}_i = \hat{X}_{i0} + \frac{\left(\frac{2\kappa}{1+\kappa^2}\right) \sum m_i \sin(\hat{X}_{i0} - x_i) - \frac{1}{c} \sum w_i [\hat{\eta}_1 \sin(y_i - \hat{X}_{i0}) - \hat{\eta}_2 \cos(y_i - \hat{X}_{i0})]}{\left(\frac{-2\kappa}{1+\kappa^2}\right) \sum m_i \cos(\hat{X}_{i0} - x_i) - \frac{1}{c} \sum w_i [\hat{\eta}_1 \cos(y_i - \hat{X}_{i0}) + \hat{\eta}_2 \sin(y_i - \hat{X}_{i0})]} \quad (21)$$

Possible initial guesses for iteration are  $\hat{X}_{i0} = x_i$  in (21) for  $i = 1, 2, \dots, n$ . The iterative re-weighting algorithm for maximum likelihood estimation obtained step by step as follows:

**Step 1:** Initialize  $\alpha_1^{[0]}$ ,  $\alpha_2^{[0]}$  and  $\hat{X}_{i0} = x_i$  with  $\alpha_1^{[0]} + \alpha_2^{[0]} < 1$  and calculate  $w^{[0]}$  using

$$w_i = \frac{1}{1 - \alpha_1 \cos(y_i - X_i) - \alpha_2 \sin(y_i - X_i)}$$

**Step 2:** Given  $\alpha_1^{[k]}$ ,  $\alpha_2^{[k]}$ ,  $X^{[k]}$  and  $w^{[k]}$  at iteration  $k$ , calculate  $\alpha_1^{[k+1]}$ ,  $\alpha_2^{[k+1]}$  and  $X^{[k+1]}$  using Equations (8) and (21), respectively.

**Step 3:** Repeat step 2 until algorithm converges.

**Step 4:** Obtain the values of  $\hat{\alpha}$  and  $\hat{\kappa}$  by solving the Equations (9) and (10) respectively.

## 2.2 Asymptotic variance of circular functional relationship model parameters:

It is difficult to derive the sampling variance based on the expectation of the second partial derivative of the log-likelihood function of the proposed model in (1). Therefore, we used bootstrapping method (Chernick 1999) as explained below:

For any  $n$  pairs of circular observations  $(x_1, y_1), \dots, (x_n, y_n)$  of two circular variables  $X$  and  $Y$  with linear relationship.

**Step 1:** Select  $m$  pairs of the observations such that  $(m < n)$ .

**Step 2:** For the selected  $m$  observations, obtain the estimates of  $\alpha$  and  $\kappa$  and label them  $\hat{\alpha}_{(1)}$  and  $\hat{\kappa}_{(1)}$

**Step 3:** Repeat Step 1 and Step 2,  $B$  times.

**Step 4:** Obtain the variance for the parameters  $\hat{\alpha}$  and  $\hat{\kappa}$  as follows:

$$\text{var}(\hat{\alpha}) = \frac{1}{B-1} \sum_{j=1}^B (\hat{\alpha}_j - \bar{\alpha})^2 \text{ and } \text{var}(\hat{\kappa}) = \frac{1}{B-1} \sum_{j=1}^B (\hat{\kappa}_j - \bar{\kappa})^2,$$

where  $\bar{\alpha}$  and  $\bar{\kappa}$  are the sample mean of  $\hat{\alpha}_{(j)}$  and  $\hat{\kappa}_{(j)}$ , respectively for  $j = 1, \dots, B$ .

### 2.3 Confidence intervals of model parameters:

The percentile or bootstrap-p method is used to construct the  $100(1 - \alpha)\%$  confidence intervals for parameters' estimates, since it is the most widely used method due to its simplicity and natural appeal, using the first three steps mentioned in Subsection 2.2 and then continued by the following step:

**Step 4:** Arrange the bootstrap estimates in an increasing order  $\hat{\alpha}_{(1)}, \dots, \hat{\alpha}_{(B)}$  and  $\hat{\kappa}_{(1)}, \dots, \hat{\kappa}_{(B)}$ , then the  $(1 - \alpha)\%$  confidence intervals of  $\alpha$  and  $\kappa$  are given by  $(\hat{\alpha}_{(\frac{\alpha}{2})}, \hat{\alpha}_{(1-\frac{\alpha}{2})})$  and  $(\hat{\kappa}_{(\frac{\alpha}{2})}, \hat{\kappa}_{(1-\frac{\alpha}{2})})$ , respectively.

### 3. Simulation Study

The main objective of this section is to assess the accuracy and biasness of the parameters of the proposed model (1) via simulation. The simulation results are obtained based on 1000 generated samples for each set of parameters values are shown in Table 1 and Table 2. The values of  $X$  follows the  $\nu M(0.5, 2)$  and the intercept parameter is fixed at  $\alpha = 0$ . Ten different choices of concentration parameter  $\kappa = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$  and  $0.99$  for random error generated from wrapped Cauchy with mean 0 and concentration parameters  $\kappa = \nu$ . Six choices of sample size  $n = 10, 30, 50, 70, 100$  and  $150$  have been considered. The simulation study is conducted as follows:

1. Two random samples  $\delta_i$  and  $\varepsilon_i$  of size  $n$  are generated from the wrapped Cauchy with mean 0 and concentration  $\kappa$  for the errors.
2. A random sample of size  $n$  is generated from the von Mises with mean 0.5 and concentration parameter 2 for the independent variable  $X$ .
3. Obtain  $x_i$  based on formula  $x_i = X_i + \delta_i \pmod{2\pi}$ .
4. Obtain  $Y$  variable as given in (1), then obtain  $y_i$  based on the formula  $y_i = Y_i + \varepsilon_i$ .
5. Estimate the models parameter using the iterative procedure as derived in Section 2.
6. For each combination of sample size  $n$  and concentration parameter  $\kappa$ , the process is repeated  $s=1000$  times.

#### 3.1 Biasness of estimators

Results of simulation study are tabulated in Table 1, and it shows that  $\hat{\alpha}$  is a good estimator of  $\alpha$ , where its bias generally decreases with the increasing of the sample size  $n$  or the concentration parameter  $\kappa$ . Similar conclusions may also be drawn from Table 2

as the mean of  $\hat{\kappa}$  close to the true value of  $\kappa$  with the increase of sample size  $n$  or the concentration parameters of circular random errors.

Also, we note from Table 1 and Table 2 that there is an inverse relationship between the sample size  $n$  and the values of the  $MCE = \frac{1}{s} \sum_{i=1}^s (1 - \cos(\hat{\alpha}_i - \alpha))$  and

$MSE = \sqrt{\frac{1}{s} \sum_{j=1}^s (\hat{\kappa}_j - \kappa)^2}$  of the estimators  $\hat{\alpha}$  and  $\hat{\kappa}$ , respectively, these findings indicate the consistency of the estimators.

**Table 1: Simulation Results for  $\hat{\alpha}$  (True value of  $\alpha = 0.0$ ).**

$\kappa$	$n$						
	Measure	10	30	50	70	100	150
<b>0.2</b>	Mean	-0.025	0.008	-0.017	-0.015	0.001	0.009
	Bias	0.025	0.008	0.017	0.015	0.001	0.009
	MCE	0.489	0.253	0.152	0.108	0.067	0.02
<b>0.3</b>	Mean	-0.019	-0.029	0.006	0.006	0.005	0.007
	Bias	0.019	0.029	0.006	0.006	0.005	0.007
	MCE	0.308	0.102	0.05	0.034	0.022	0.016
<b>0.4</b>	Mean	-0.001	-0.008	-0.014	0.003	0.000	-0.003
	Bias	0.001	0.008	0.014	0.003	0.000	0.003
	MCE	0.178	0.045	0.025	0.016	0.012	0.007
<b>0.5</b>	Mean	-0.013	0.002	-0.001	-0.001	-0.003	0.000
	Bias	0.013	0.002	0.001	0.001	0.003	0.000
	MCE	0.100	0.021	0.012	0.009	0.006	0.004
<b>0.6</b>	Mean	-0.006	-0.003	0.000	-0.000	0.006	-0.001
	Bias	0.006	0.003	0.000	0.000	0.006	0.001
	MCE	0.041	0.010	0.006	0.005	0.003	0.002
<b>0.7</b>	Mean	0.007	-0.003	0.003	-0.001	-0.001	-0.001
	Bias	0.007	0.003	0.003	0.001	0.001	0.001
	MCE	0.019	0.005	0.003	0.002	0.001	0.001
<b>0.8</b>	Mean	-0.008	0.001	0.001	-0.002	-0.001	-0.000
	Bias	0.008	0.001	0.001	0.002	0.001	0.000
	MCE	0.007	0.002	0.001	0.001	0.001	0.0003
<b>0.9</b>	Mean	-0.005	0.002	0.001	0.000	-0.000	-0.000
	Bias	0.005	0.002	0.001	0.000	0.000	0.000
	MCE	0.001	0.0004	0.0002	0.0001	0.0001	0.000
<b>0.99</b>	Mean	0.000	0.000	0.000	0.000	0.000	0.000
	Bias	0.000	0.000	0.000	0.000	0.000	0.000
	MCE	0.000	0.000	0.000	0.000	0.000	0.000



### 3.2 Coverage probability of confidence intervals

The coverage probability of a confidence interval is the proportion of the time that the interval contains the true value of interest, see (Dodge; 2003). In this simulation study we construct intervals at 0.95 confidence level. Hence, the good indicator must give the coverage probability close to 0.95. The estimation of the parameters were repeated for different sample sizes ( $n = 30, 50, 100$ , and  $150$ ) and concentration parameter  $\kappa = 0.2, 0.4, 0.6, 0.8$  and  $0.99$ . The results of the simulations are given in Table 3 based on 1000 repetition times.

Results in Table 3 show that the observed coverage appears to be approximately equivalent to the confidence level, which reveals an excellent coverage of the confidence intervals.

**Table 2: Simulation Results for  $\hat{\kappa}$**

$\kappa$	Measure	$n$					
		10	30	50	70	100	150
<b>0.2</b>	Mean	1.574	0.466	0.339	0.299	0.251	0.230
	Bias	1.374	0.266	0.139	0.099	0.051	0.030
	MSE	94.031	1.206	0.343	0.037	0.016	0.005
<b>0.3</b>	Mean	0.890	0.418	0.36	0.342	0.331	0.322
	Bias	0.590	0.118	0.06	0.042	0.031	0.022
	MSE	5.138	0.052	0.015	0.010	0.006	0.003
<b>0.4</b>	Mean	0.800	0.471	0.440	0.431	0.418	0.411
	Bias	0.200	0.071	0.0340	0.031	0.018	0.011
	MSE	2.283	0.023	0.01	0.007	0.004	0.002
<b>0.5</b>	Mean	0.704	0.541	0.529	0.522	0.513	0.508
	Bias	0.204	0.041	0.029	0.022	0.013	0.008
	MSE	0.318	0.013	0.007	0.005	0.003	0.002
<b>0.6</b>	Mean	0.717	0.629	0.613	0.6103	0.609	0.605
	Bias	0.117	0.029	0.013	0.0103	0.009	0.005
	MSE	0.109	0.009	0.005	0.003	0.002	0.001
<b>0.7</b>	Mean	0.755	0.715	0.710	0.704	0.705	0.702
	Bias	0.055	0.015	0.010	0.004	0.005	0.002
	MSE	0.019	0.005	0.003	0.002	0.001	0.001
<b>0.8</b>	Mean	0.825	0.806	0.804	0.803	0.802	0.801
	Bias	0.025	0.006	0.004	0.003	0.002	0.001
	MSE	0.007	0.002	0.001	0.001	0.001	0.000
<b>0.9</b>	Mean	0.905	0.900	0.901	0.901	0.900	0.900
	Bias	0.005	0.000	0.001	0.001	0.000	0.000
	MSE	0.002	0.000	0.000	0.000	0.000	0.000
<b>0.99</b>	Mean	0.991	0.990	0.990	0.990	0.991	0.990
	Bias	0.001	0.001	0.000	0.001	0.000	0.000
	MSE	0.000	0.000	0.000	0.000	0.000	0.000

**Table 3: Results for coverage percentage of confidence interval for  $\alpha$  and  $\kappa$**

$\kappa$	$n$									
	30		50		70		100		150	
	$\alpha$	$\kappa$	$\alpha$	$\kappa$	$\alpha$	$\kappa$	$\alpha$	$\kappa$	$\alpha$	$\kappa$
<b>0.2</b>	95.5	98.9	93.3	99.7	94.2	99.6	95.1	93.6	94.8	93.1
<b>0.4</b>	93.0	99.6	95.6	92.5	96.4	92.1	93.6	94.0	95.1	94.5
<b>0.6</b>	93.9	95.2	95.2	92.3	95.0	93.5	94.1	94.8	94.8	94.7
<b>0.8</b>	94.7	95.5	95.5	95.4	95.8	90.7	95.0	94.5	95.3	94.5
<b>0.99</b>	95.9	96.6	95.2	95.9	94.7	97.8	95.1	96.5	95.0	96.5

### 3.3 Robustness of the estimates

Robustness of an estimator is a useful property which gives a fair assurance that the existence of any possible outlier or violation of model assumptions will not have much effect on the parameters estimates. To assess the robustness, based on simulation study, where 5% of the generated data are contaminated start from observation  $d$  as follows:

$$y^*[d] = y[d] + \gamma\pi \pmod{2\pi},$$

where  $y^*[d]$  and  $y[d]$  are the contaminated and original values of the dependent variable  $Y$  respectively, and  $\gamma$  is the contamination level, where  $0 \leq \gamma \leq 1$ .

The results of simulation study show that, for data with highly concentrated error  $\kappa \geq 0.6$ , the bias of  $\hat{\alpha}$  is less than 0.3 for small samples ( $n = 10$ ) and less than 0.15 for moderate and large samples ( $n \geq 30$ ). Regardless the concentration of error and the sample size, the optimum values of bias of  $\alpha$  are obtained for moderate levels of contamination  $\gamma \approx 0.6$ . It may be referred to the nature of the circular data, where the shift of 5% of the generated data by more than  $\pi/2$  of its original values may lead to get closer to the majority of the data. Almost similar conclusion can be drawn for the concentration parameter.

## 4. Real Data Analysis

As an illustration of the proposed model, this section considers the measurements of wind directions that have been collected from two meteorological stations in two main cities in the Gaza Strip, namely Gaza and Khan Younis in Palestine. The data were provided by the Palestinian Metrological Authority, in 2007. It represent the monthly average wind directions every three hours a day viz, 0:00 mid night, 3:00 am, 6:00 am, 9:00 am, 12:00 noon, 15:00 pm, 18:00 pm and 21:00 pm. (Badawi 2013) analyzed the Palestinian wind data in order to establish a wind farm to reach the optimum electricity energy. Recently, (Abuzaid and Allahham 2015) modelled the data by using the simple circular regression model assuming the wrapped Cauchy error. The linear circular regression model between dependent and independent variables is given by

$$Y = 0.747 + 0.842X \pmod{2\pi},$$

where dependent variable  $Y$  is the wind direction data for Gaza while the wind direction data for Khan Younis is the independent variable  $X$ .

#### 4.1 Fitting the functional circular regression model with wrapped Cauchy error

Since the relationship between the wind directions of Gaza and Khan Younis is reasonably linear as shown in Figure 1; the simple functional circular regression model (1) is suggested to fit the data. The parameters estimates are obtained by applying the iterative procedure and after converting the data into radians. The initial values of  $\alpha_1$  and  $\alpha_2$  are taking to be 0.3.

Table 4 presents the estimates of parameters, their standard error and the 95% confidence intervals, where the convergence is occurred after 16 iterations. Hence, the estimate relationship for wind directions of Gaza and Khan Younis data set is given by

$$Y = 0.237 + X \quad (\text{mod } 2\pi),$$

where variable  $Y$  is the wind directions data for Gaza and the wind directions data for Khan Younis is the variable  $X$ .

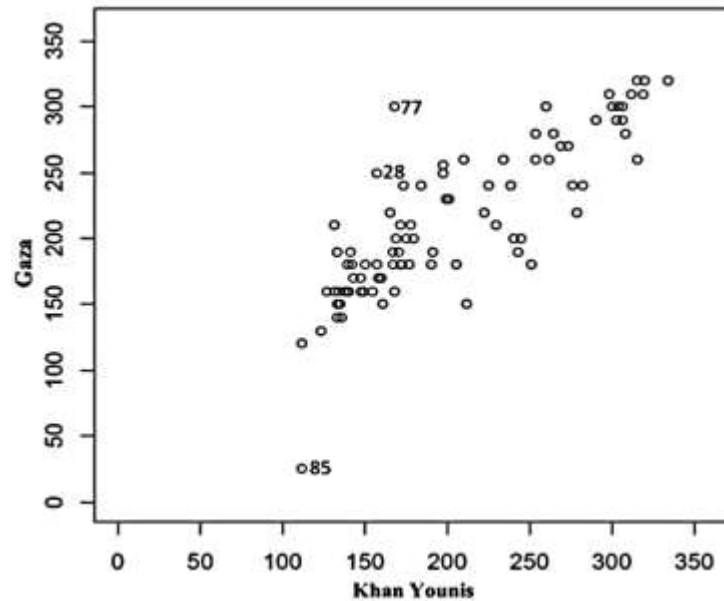


Figure 1: Scatter plot of wind directions in Gaza versus KhanYounis

**Table 4: Parameter estimation of circular functional model with wrapped Cauchy Errors**

Parameter	Estimate	Standard Error	Confidence Interval
$\hat{\alpha}$	0.237	0.0024	(0.140, 0.333)
$\hat{\kappa}$	0.843	0.0018	(0.759, 0.927)

(Abuzaid 2010) suggested a statistic to test the goodness-of-fit of circular models,  $A^*(\kappa)$  to be analogues the coefficient of determination in the linear regression models as given below:

$$A^*(\kappa) = \frac{1}{n} \sum \cos^2(y_i - \hat{y}_i),$$

where  $0 \leq A^*(\kappa) \leq 1$ . Therefore, the closer  $A^*(\kappa)$  to 1 indicates a better fitting the model. Thus, the goodness-of-fit for the model is  $A^*(\kappa) = 0.799$ .

The obtained residuals were tested to follow the wrapped Cauchy distribution via Kolmogorov-Smirnov test, where the values of test for the errors of and are 0.0401 and 0.0321 with P-values 0.835 and 0.895, respectively.

## 5. Conclusions

A new linear functional relationship model for circular variables with a wrapped Cauchy errors has been proposed due to the attractive properties of wrapped Cauchy distribution. The maximum likelihood estimates of parameter has been obtained assuming equality of concentration parameters of errors. Estimation has been obtained iteratively since the closed-form expression for estimates are not available, by choosing a suitable initial values, the standard error of estimates as well as their confidence intervals are obtained by bootstrapping methods. Moreover, the proposed angular regression model has been applied on a real data set of wind directions at two cities in the Gaza strip.

## References

1. Abuzaid, A. H. (2010). *Some Problems for Outliers Circular Data*. PhD Thesis, University of Malaya, Kuala Lumpur.
2. Abuzaid, A. and Allahham, N. (2015). Simple circular model assuming wrapped Cauchy error. *Pakistan Journal of Statistics*, 31 (4), 385-398.
3. Adcock, R. J. (1877). Note on the method of least squares, *Analyst*, 4, 183-184.
4. Adcock, R. J. (1878). A problem in least squares, *Analyst*, 5, 53-54.
5. Badawi, A. S. (2013). An Analytical Study for Establishment of Wind Farms in Palestine to Reach the Optimum Electrical Energy. Master thesis. The Islamic University of Gaza, Palestine.
6. Caires, S. and Wyatt, L. R. (2003). A linear functional relationship model for circular data with an application to the assessment of ocean wave measurements, *J. Agric. Biol. Environ. Stat.*, 8 (2), 153-169.
7. Chernick, M. R. (1999). *Bootstrap Method: A practitioner's Guide*. Canada: John Wiley & Sons.
8. Dodge, Y. (2003). The Oxford Dictionary of Statistical Terms, *OUP*.
9. Downs, T. D. and Mardia, K. V. (2002). Circular regression. *Biometrika*, 89 (3), 683-697.
10. Hussin, A. G. (1997). *Pseudo-replication in functional relationship with environmental application*. Ph.D. Thesis. University of Sheffield, England.
11. Hussin, A. G. (2003). On estimating error concentration parameter for circular functional model. *Bull. Malaysian. Math. Soc. (Second Series)*, 26, 181-188.

12. Hussin, A. G. (2005). Approximating Fisher's information for the replicated linear circular functional relationship model. *Bull. Malays. Math. Sci. Soc.*, 28,(2), 1-9
13. Hussin, A. G., Abuzaid, A. H., Zulkifli, F. and Mohamed, I. B. (2010). Asymptotic covariance and detection of influential observations in a linear relationship model for circular data with application to the measurements of wind directions. *ScienceAsia*, 36, 249-253.
14. Hussin, A. G., Fieller, N. R. J. and Stillman, E. C. (2004). Linear regression for circular variables with application to directional data. *Journal of Applied Science & Technology*, 9, (1 & 2), 1-6.
15. Kendall, M. G. (1951). Regression, structural and functional relationship I, *Biometrika*, 38, 11-25.
16. Kendall, M. G. (1952). Regression, structural and functional relationship II, *Biometrika*, 39, 96-108.
17. Satari, S. Z., Hussin, A. G., Zubairi, Y. Z. and Hassan, S. F. (2014). A new functional relationship model for circular variables. *Pak. J. Statist.*, 30 (3), 397-410.