mathematics

MDPI

Article

Wrapped Cauchy Robust Approach to the Circular-Circular Regression Model

Adnan Karaibrahimoglu [1], Mutlu Altuntas [2] and Hani Hamdan [3,*]

[1] Biostatistics & Medical Informatics Department, Medical School of Suleyman Demirel University, Isparta 32000, Turkey

[2] Department of Data Science and Analytics, Faculty of Arts and Science, Sinop University, Sinop 57000, Turkey; maltuntas@sinop.edu.tr

[3] Laboratoire des Signaux et Systèmes (L2S UMR CNRS 8506), CentraleSupélec, Université Paris-Saclay, CNRS, 91190 Gif-sur-Yvette, France

* Correspondence: Hani.Hamdan@centralesupelec.fr

Abstract

Circular–circular regression models are widely used to investigate relationships between angular variables in various applied fields, including biostatistics. The classical *von* Mises (vM) circular–circular regression model, however, is known to be sensitive to outliers due to its light-tailed error structure. In this study, we investigate the wrapped Cauchy (WC) circular–circular regression model as a robust alternative to the vM-based approach for analyzing circular data contaminated by outliers. Parameter estimation is performed via maximum likelihood (ML) using a modern constrained gradient-based optimization algorithm, namely the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm with box constraints (L-BFGS-B), allowing for stable estimation under natural parameter bounds. Extensive simulation studies demonstrate that, under contaminated settings, the WC model provides substantially more stable parameter estimates than the vM model, yielding markedly lower mean squared error and variability, particularly for high concentration regimes and directional outliers. The robustness advantage of the WC model is further illustrated through a real biostatistical application involving the circular relationship between the months of diagnosis and surgical intervention in gastric cancer patients. Overall, the results highlight the practical benefits of WC-based circular–circular regression for robust inference in the presence of outliers.

Keywords: circular-circular regression; wrapped Cauchy; optimization; *von* Mises; heavy-tailed data; gastric carcinoma

MSC: 90C26; 62F35

Check for updates

Academic Editor: Manuel Alberto M. Ferreira

Received: 17 December 2025
Revised: 16 January 2026
Accepted: 21 January 2026
Published: 26 January 2026

1. Introduction

Modeling functional relationships between circular variables has attracted considerable attention across a wide range of scientific disciplines. In particular, such models have been extensively applied in biology, where directional and phase-related measurements frequently arise [1,2], astronomy, where angular data naturally emerge in the analysis of arrival directions and phase information of cosmic and gamma-ray sources [3], geophysics, where angular dependencies are commonly observed in natural phenomena [4], medicine, especially in the analysis of biological rhythms and cyclic processes [5], and meteorology, notably in wind direction modeling and forecasting [6]. Traditional linear regression techniques are inadequate for modeling circular variables due to their inherent circular nature. Consequently, the development and application of statistical techniques specifically designed to address circular data have increased substantially. Circular regression is a statistical method used to model relationships between variables measured on a circular scale, such as degrees or radians. The circular nature of the data necessitates special considerations in modeling, as classical regression techniques assume linear relationships between variables. Examples of circular regression applications include the impact of wind direction on air pollution levels, the effect of leaf orientation angles on the amount of sunlight received, the relationship between the time of day (measured on a circular scale from 0 to 24) and the severity of symptoms in patients with specific medical conditions, and the relationship between comet tail directions and their orbital positions.

Circular regression models—such as circular-circular, circular-linear, and linear-circular regression—are widely used to model relationships between circular variables and other types of data or to predict circular outcomes. These models can also accommodate circular data nested within more complex data structures, such as longitudinal studies or clustered datasets. Numerous studies have focused on the theoretical foundations and methodological developments of circular regression analysis [7–11].

The circular-circular regression model aims to identify relationships between a circular response variable and a circular covariate, typically assuming circular probability distributions for both variables. A wide range of probability distributions has been proposed for modeling circular data, including uniform, cardioid, vM, wrapped normal (WN), WC, and more general wrapped families. Comprehensive treatments of circular and directional distributions can be found in classical references such as [12], as well as in more recent review studies [13,14]. These distributions differ in terms of tail behavior, modality, and robustness, offering varying degrees of flexibility depending on the characteristics of the data. Among these alternatives, the vM distribution is widely used due to its mathematical tractability and interpretability, whereas wrapped distributions—particularly the WC—exhibit heavier tails and are therefore more robust to outliers and departures from model assumptions. This diversity of available models highlights the importance of carefully selecting circular distributions based on the underlying data structure and analysis objectives.

Accordingly, circular–circular regression models based on vM-distributed error terms are known to be sensitive to outliers and deviations from model assumptions. In contrast, the WC circular–circular regression model, which assumes a WC-distributed error structure, offers a more robust alternative for modeling circular data contaminated by outliers. Despite these developments, the literature—particularly in biostatistical applications—still lacks comprehensive and systematic evaluations of robust circular–circular regression methods for analyzing relationships between two circular variables in the presence of outliers.

This study aims to address this gap by comparing the vM and WC simple circular-circular regression models, highlighting the benefits and limitations of the WC model in terms of robustness. Additionally, the study includes a real-world application in biostatistics, using data on the months of diagnosis and operation in gastric cancer patients.

In the existing literature, circular regression models assuming WC and related error structures have predominantly been estimated using iterative re-weighting algorithms within an ML framework, commonly referred to as Iteratively Reweighted Least Squares (IRWLS)—type procedures [15,16]. Although these approaches are computationally convenient, they rely on local approximations and may exhibit numerical instability or slow convergence when dealing with highly concentrated data or substantial contamination.

In contrast, the present study adopts a modern constrained optimization framework based on the L-BFGS-B to directly maximize the likelihood functions of both vM and WC circular–circular regression models. This strategy allows explicit enforcement of natural parameter bounds and efficient utilization of gradient information, leading to improved numerical stability across a wide range of concentration and contamination scenarios. To the best of our knowledge, the application of L-BFGS-B to simple circular–circular regression models has not previously been explored in the literature. Building on this estimation framework, the main contribution of this paper is not the introduction of the WC distribution itself—which is already well established in circular statistics—but rather a unified and reproducible robustness assessment of vM- and WC-based circular–circular regression models under controlled contamination scenarios. We conduct systematic comparisons across different concentration regimes and directional outlier settings and demonstrate the practical implications of these methodological choices through extensive simulation studies and a real biostatistical application.

The remainder of the article is structured as follows: Section 2 presents an overview of the development of vM and WC simple circular-circular regression models, with a focus on the relationship between a circular response and a circular predictor. This section also evaluates the theoretical properties of ML estimators for the models. Section 3 examines the sensitivity and finite-sample performance of these models through an extensive simulation study. Section 4 applies both models to real-world data from gastric cancer patients, comparing their robustness in the presence of outliers. Finally, Sections 5 and 6 conclude the study with key findings and future research directions. All computations were performed using R version 4.5.2 software [17].

2. Simple Circular-Circular Regression Analysis

Circular-circular regression is used when both the response and predictor variables are circular or cyclical in nature, often represented in angles, directions, or periodic data. This method aims to model the relationship between these circular variables. Simple circular-circular regression models are specifically designed to analyze the relationship between two circular variables. These models aim to estimate how changes in one circular variable influence or predict changes in another circular variable. There are different approaches to circular-circular regression modeling, including various statistical techniques to understand and quantify the relationship between these cyclical variables.

2.1. Von Mises Simple Circular-Circular Regression Model

One standard circular-circular regression model is based on the *von* Mises (vM) distribution, which is analogous to the Normal distribution for circular data. The vM distribution describes the circular variability and allows for the estimation of circular regression coefficients.

2.1.1. Model Specification

The vM simple circular-circular regression model assumes a linear relationship between two circular variables X and Y when projected onto a circle [18]. It is expressed as:

$$y_i = \alpha + \beta x_i + \varepsilon_i \qquad (mod\ 2\pi) \quad , \qquad i = 1, 2, \ldots, n, \tag{1}$$

where y_i represents the circular response variable, x_i is the circular predictor variable, α and β are the regression coefficients that describe the relationship between the circular variables, ε_i is the independently identically distributed (i.i.d.) circular random error term, assumed to follow a vM distribution with circular mean 0 and concentration parameter $\kappa > 0$, denoted by $\varepsilon_i \sim vM(0, \kappa)$. The parameter κ controls the concentration of the distribution around the mean direction, with larger values corresponding to more highly concentrated circular data.

2.1.2. Parameter Estimation for the vM Model

Estimating the parameters $(\alpha,\ \beta,\ \kappa)$ typically involves ML estimation or other specialized circular regression techniques. These methods aim to find the best-fitting parameters that describe the relationship between the circular variables, considering the properties of circular statistics.

The probability density function (pdf) of circular random error (ε_i) in model (1) is given by:

$$f(\varepsilon_i|\kappa) = \frac{1}{2\pi I_0(\kappa)} exp(\kappa \cos(\varepsilon_i)), \tag{2}$$

where $\varepsilon_i = y_i - \alpha - \beta x_i\ (mod\ 2\pi)$ and $I_0(\kappa)$ is the modified Bessel function of the first kind of order 0. Given n observations (y_i, x_i), the likelihood function (L) can be constructed as follows:

$$L(y, x; \alpha, \beta, \kappa) = (2\pi\ I_0(\kappa))^{(-n)} exp\left(\sum_{i=1}^{n} \kappa\ cos(y_i - \alpha - \beta\ x_i)\right) \tag{3}$$

Taking the natural logarithm gives us the log-likelihood function (log L):

$$\log L(y, x; \alpha, \beta, \kappa) = -n\ log(2\pi\ I_0(\kappa)) + \sum_{i=1}^{n} \kappa\ cos(y_i - \alpha - \beta\ x_i) \tag{4}$$

To obtain the ML estimators, we need to maximize the log-likelihood function with respect to the parameters α, β and κ. This is typically achieved through numerical optimization, as analytical solutions are generally not available.

2.1.3. Gradient Calculation for the vM Model

Calculating gradients can help in improving convergence speed in likelihood-based estimation. To improve readability and facilitate stable implementation within gradient-based constrained optimization routines, the score functions are expressed in terms of wrapped residuals. This representation avoids unnecessary trigonometric expansions while preserving the exact likelihood structure.

$$\delta_i = wrap(y_i - \alpha - \beta\ x_i) \quad , \quad \delta_i \in (-\pi, \pi] \tag{5}$$

where δ_i denotes the wrapped residual taking values in $(-\pi, \pi]$ and $wrap(\cdot)$ maps angles to the principal interval.

Based on the log L function defined in Equation (4), the corresponding score functions with respect to the model parameters are derived below. Let l denote the log L associated with the vM circular-circular regression model.

The gradient with respect to α is

$$\frac{\partial l}{\partial \alpha} = \kappa \sum_{i=1}^{n} sin(\delta_i) \tag{6}$$

The gradient with respect to β is

$$\frac{\partial l}{\partial \beta} = \kappa \sum_{i=1}^{n} x_i sin(\delta_i) \tag{7}$$

The gradient with respect to κ is

$$\frac{\partial l}{\partial \kappa} = -n\frac{I_1(\kappa)}{I_0(\kappa)} + \sum_{i=1}^{n} cos(\delta_i) \tag{8}$$

where $I_0(.)$ and $I_1(.)$ denote the modified Bessel functions of the first kind of order zero and one, respectively. Since closed-form solutions are not available, parameter estimation is carried out numerically via constrained gradient-based optimization. The score functions are not solved analytically. Instead, they are evaluated numerically and used as gradient inputs in the L-BFGS-B algorithm to maximize the log-likelihood under natural parameter constraints. Parameter estimation is therefore carried out iteratively via constrained, gradient-based optimization.

2.2. *Wrapped Cauchy Simple Circular-Circular Regression Model*

The vM simple circular-circular regression model is known to be sensitive to the presence of outliers in circular datasets, particularly when deviations from model assumptions occur. Parameter estimates obtained under vM-distributed errors may therefore be unstable in the presence of contamination. In this context, the WC simple circular–circular regression model, which assumes a heavier-tailed circular error distribution, provides a robust alternative for modeling circular data containing outliers.

2.2.1. Model Formulation

The WC simple circular–circular regression model adopts the same structural form as Equation (1), differing only in the distributional assumption of the random error term. Specifically, the random error term ε_i is assumed to be i.i.d. according to a WC distribution with location parameter zero and concentration parameter ρ, denoted by $\varepsilon_i \sim WC(0, \rho)$, where $0 < \rho < 1$. The parameter ρ controls the concentration of the distribution around the mean direction.

The WC distribution is particularly suitable for circular data as it naturally accounts for the periodic nature of angular measurements and exhibits heavier tails than the vM distribution, making it more robust to outlying observations. From a statistical perspective, the heavier tails of the WC distribution assign lower influence to extreme angular deviations, leading to a flatter likelihood surface in the presence of outliers. In contrast, the vM distribution becomes increasingly concentrated as the concentration parameter increases, resulting in a steeper likelihood surface and greater sensitivity to outliers. This fundamental difference explains the improved robustness of WC-based circular–circular regression models under data contamination.

2.2.2. Parameter Estimation for the WC Model

The pdf of the circular random error (ε_i) is given by

$$f(\varepsilon_i|\rho) = \frac{1-\rho^2}{2\pi(1+\rho^2-2\rho\cos(\varepsilon_i))}, \tag{9}$$

where $\varepsilon_i = y_i - \alpha - \beta x_i \ (mod\ 2\pi)$.

For the observed data (y_i, x_i), the corresponding L is

$$L(y, x; \alpha, \beta, \rho) = (2\pi\ )^{(-n)} \frac{\left(1-\rho^2\right)^n}{\prod_{i=1}^n (1+\rho^2-2\rho\cos(y_i-\alpha-\beta x_i))} \tag{10}$$

Taking logarithms, the $\log L$ function can be written as

$$\begin{aligned}\log L(y, x; \alpha, \beta, \rho) = \quad & -nlog(2\pi) + nlog\left(1-\rho^2\right) \\ & -\textstyle\sum_{i=1}^n log\left(1+\rho^2-2\rho\cos(y_i-\alpha-\beta x_i)\right)\end{aligned} \tag{11}$$

2.2.3. Gradient Calculation for the WC Model

Based on the log L function defined in Equation (11), the corresponding score functions with respect to the model parameters are derived below. Let l denote the log L function associated with the WC circular-circular regression model.

Define the wrapped residual and auxiliary term as

$$\begin{gathered} \delta_i = wrap(y_i - \alpha - \beta\, x_i) \quad , \quad \delta_i \in (-\pi, \pi] \\ D_i = 1 + \rho^2 - 2\rho cos(\delta_i) \end{gathered} \tag{12}$$

The gradient with respect to α is

$$\frac{\partial l}{\partial \alpha} = \sum_{i=1}^{n} \frac{2\rho sin(\delta_i)}{D_i} \tag{13}$$

The gradient with respect to β is

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^{n} \frac{2\rho x_i sin(\delta_i)}{D_i} \tag{14}$$

The gradient with respect to ρ is

$$\frac{\partial l}{\partial \rho} = -\frac{2n\rho}{1-\rho^2} - \sum_{i=1}^{n} \frac{2\rho - 2cos(\delta_i)}{D_i} \tag{15}$$

3. Simulation Study Results

This section presents the settings and results of a simulation study designed to evaluate the accuracy of parameter estimates for the two simple circular-circular regression models considered in the study.

3.1. Settings

We fix the intercept parameter $\alpha = \pi/4(\cong 0.785)$ and the slope parameter $\beta = 1.5$ as representative, non-extreme values in the simple circular–circular regression model, without loss of generality. These values avoid boundary configurations and ensure a moderate signal level, facilitating a stable and interpretable comparison between models. Additional simulation experiments conducted with alternative parameter choices yielded qualitatively similar results, indicating that the main conclusions are not sensitive to the specific values of α and β.

The sample size (n) was considered in three different choices, 50, 100 and 200. To assess the influence of outliers on the estimators, we considered adding outliers in the $x-$, $y-$ and $x-y$ directions at proportions of k% of n, where k = 5 or 10. Outliers were simulated from a Uniform distribution in the range $[5\pi, 6\pi]$ for the $x-$ direction, $[3\pi, 4\pi]$ for the $y-$ direction, and $[\pi, 2\pi]$ for the $x-y$ direction, and were added to the circular dataset. The initial values for α, β and κ were chosen as 1. For ρ parameter, initial value was selected as 0.5. Outliers were generated from uniform angular intervals to represent non-informative or randomly perturbed observations, as commonly encountered in circular data due to measurement errors, misclassification, or heterogeneous subpopulations. Such uniformly distributed contamination reflects extreme deviations from the underlying directional relationship and provides a stringent setting for evaluating robustness of the competing models. The simulation study was conducted on two scenarios as follows:

Scenario 1:

For ε_i: Generate errors from the vM distribution with a concentration parameter κ, $\varepsilon_i \sim vM(0, \kappa)$. Various values of the concentration parameter are considered: 2, 4, 8, 14.

For x_i: Simulates the predictor circular data uniformly from $[0, 2\pi]$.

For y_i: Simulates the response of circular data based on the model (1).

Defines the negative log L function to be minimized. The residuals are calculated, and then the log L for the vM simple circular-circular regression model is computed. The L-BFGS-B method, which is particularly useful in circular data analysis and can effectively handle constrained parameter optimization, was employed. The L-BFGS-B algorithm was selected due to its ability to efficiently handle box-constrained optimization problems, which is essential in the present context since the concentration parameters are naturally restricted ($\kappa > 0$ and $0 < \rho < 1$). In addition, L-BFGS-B is a gradient-based method that requires limited memory, making it well suited for large-scale simulation studies involving repeated likelihood maximization. Compared to derivative-free or unconstrained optimization methods, L-BFGS-B provides improved numerical stability and faster convergence, particularly when the likelihood surface becomes steep under high concentration or contaminated data settings.

Scenario 2:

For ε_i: Generate errors from the WC distribution with a concentration parameter ρ, $\varepsilon_i \sim WC(0, \rho)$. Various values of concentration parameter ρ are considered 0.4, 0.7, 0.8 and 0.9.

For x_i: Simulates the predictor circular data uniformly from $[0, 2\pi]$.

For y_i: Simulates the response of circular data based on the model (1).

Define the log L function for the WC simple circular-circular regression. It ensures the concentration parameter ρ is constrained between 0 and 1. The method "L-BFGS-B" is chosen for bounded optimization, where the concentration parameter ρ is constrained to be $0 < \rho < 1$.

The simulation procedure was applied for each pair of concentration parameters (κ and ρ) and sample size (n), and is repeated s =1000 times to ensure the convergence criterion. Numerical stability and convergence behavior were monitored throughout the simulation experiments. For highly concentrated settings, the likelihood surface of the vM model became increasingly steep, which occasionally resulted in slower convergence and sensitivity to starting values. In contrast, no numerical instability was observed for the WC model across the examined range of ρ values. All optimization runs terminated successfully under the specified convergence criteria.

3.2. Accuracy of Parameter Estimates

The accuracy of parameter estimates in the two considered simple circular-circular regression models depends on how well the models capture the underlying circular relationship between the predictor and response variables. Note that, in these models, parameter α is a circular quantity, whereas parameters β, κ and ρ are linear where $-\infty < \beta < \infty$, $\kappa > 0$ and $0 < \rho < 1$.

(a) Circular parameter; α:

Mean Circular Error (MCE) is appropriate here because it measures the angular discrepancy between the actual and estimated angles, unaffected by circular wrapping.

$$MCE(\alpha) = \frac{1}{s}\sum_{i=1}^{s}[1 - cos(\hat{\alpha}_i - \alpha)] \quad (16)$$

where $\hat{\alpha}_i$ is the estimated intercept angle and α is actual intercept angle [19].

(b) Linear parameters; β, κ and ρ:

β, κ and ρ are not inherently circular; they behave more like linear coefficients, so linear error measures like Mean Squared Error (MSE) are more heavily.

$$MSE(\beta) = \frac{1}{s}\sum_{i=1}^{s}\left(\hat{\beta}_i - \beta\right)^2 \tag{17}$$

where $\hat{\beta}_i$ is the estimated slope parameter and β is the proper slope parameter.

$$MSE(\kappa) = \frac{1}{s}\sum_{i=1}^{s}(\hat{\kappa}_i - \kappa)^2 \tag{18}$$

where $\hat{\kappa}_i$ is the estimated concentration parameter and κ is the proper concentration parameter.

$$MSE(\rho) = \frac{1}{s}\sum_{i=1}^{s}(\hat{\rho}_i - \rho)^2 \tag{19}$$

where $\hat{\rho}_i$ is the estimated concentration parameter and ρ is the proper concentration parameter [20].

(c) Empirical standard deviation of parameter estimates (Monte Carlo standard error of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\kappa}$ and $\hat{\rho}$)

The empirical standard deviations of the parameter estimates ($\hat{\alpha}$, $\hat{\beta}$, $\hat{\kappa}$, $\hat{\rho}$) were computed from 1000 Monte Carlo replications to assess estimator stability:

$$SD\left(\hat{\theta}\right) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\hat{\theta}^{(r)} - \overline{\hat{\theta}}\right)^2} \tag{20}$$

where $\hat{\theta}^{(r)}$ denotes the estimate of the parameter obtained from the r-th Monte Carlo replication, R indicates the total number of replications (in this study, R = 1000) and $\overline{\hat{\theta}} = \frac{1}{R}\sum_{r=1}^{R}\hat{\theta}^{(r)}$ represents the mean of the parameter estimates over all R replications [21].

3.3. Results

Parameter estimates for the vM and WC models were obtained using the “optim” function in R, which provides a unified interface for likelihood-based optimization. For the vM model, vM–distributed errors were generated using the “rvm” function from the “CircStats” package, while for the WC model, WC–distributed errors were generated using the WC function from the {circular} package. The L-BFGS-B optimization algorithm was employed due to its suitability for box-constrained likelihood optimization, which is essential for enforcing the natural parameter bounds of the concentration parameters. Data preprocessing and manipulation were performed using the {dplyr} package, and graphical representations of circular data and model fits were produced using the {ggplot2} package. Although alternative optimization routines available in R, such as “nlminb” or derivative-free methods implemented within “optim”, could also be applied to the problem, the selected combination of functions offers an efficient, flexible, and reproducible framework for the simulation and estimation procedures considered in this study.

We computed the empirical mean, standard deviation, MCE, and MSE for model parameters, and the results are presented in Tables 1–4. Circular distribution graphs for circular data sets with vM and WC distribution were given in Figure 1. Tail behaviors of vM and WC error distributions were given in Figure 2. All necessary computations were executed on the R platform (version 4.5.2).

Table 1. Parameter estimates of the vM and WC simple circular–circular regression models without outliers (s = 1000 Monte Carlo replicates).

		$\varepsilon_i \sim vM(0,\kappa)$				$\varepsilon_i \sim WC(0,\rho)$		
n	κ	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
50	2	0.788; 0.131 (0.008)	1.499; 0.130 (0.017)	2.108; 0.385 (0.159)	0.4	0.765; 0.359 (0.044)	1.465; 0.324 (0.106)	0.429; 0.078 (0.007)
	4	0.779; 0.145 (0.004)	1.498; 0.118 (0.014)	4.306; 0.858 (0.830)	0.7	0.783; 0.075 (0.003)	1.501; 0.043 (0.001)	0.709; 0.050 (0.002)
	8	0.771; 0.238 (0.008)	1.487; 0.199 (0.039)	8.600; 1.866 (3.840)	0.8	0.783; 0.047 (0.001)	1.501; 0.027 (0.001)	0.806; 0.035 (0.001)
	14	0.752; 0.380 (0.017)	1.465; 0.344 (0.119)	14.939; 3.326 (11.938)	0.9	0.784; 0.022 (0.000)	1.500; 0.012 (0.000)	0.903; 0.019 (0.000)
100	2	0.781; 0.194 (0.006)	1.494; 0.163 (0.027)	2.035; 0.271 (0.074)	0.4	0.784; 0.183 (0.013)	1.498; 0.097 (0.009)	0.414; 0.060 (0.003)
	4	0.776; 0.232 (0.008)	1.484; 0.225 (0.051)	4.090; 0.593 (0.360)	0.7	0.787; 0.052 (0.001)	1.500; 0.029 (0.000)	0.705; 0.036 (0.001)
	8	0.765; 0.279 (0.009)	1.488; 0.195 (0.038)	8.174; 1.258 (1.611)	0.8	0.786; 0.032 (0.000)	1.500; 0.018 (0.000)	0.803; 0.026 (0.001)
	14	0.781; 0.127 (0.002)	1.496; 0.111 (0.012)	14.408; 2.073 (4.460)	0.9	0.785; 0.015 (0.000)	1.500; 0.008 (0.000)	0.901; 0.013 (0.000)
200	2	0.776; 0.223 (0.006)	1.493; 0.165 (0.027)	2.012; 0.203 (0.041)	0.4	0.786; 0.108 (0.006)	1.499; 0.059 (0.003)	0.409; 0.042 (0.002)
	4	0.776; 0.179 (0.004)	1.493; 0.153 (0.023)	4.046; 0.414 (0.173)	0.7	0.785; 0.036 (0.000)	1.499; 0.020 (0.000)	0.703; 0.025 (0.000)
	8	0.761; 0.342 (0.015)	1.471; 0.310 (0.096)	8.023; 1.084 (1.174)	0.8	0.785; 0.022 (0.000)	1.500; 0.012 (0.000)	0.802; 0.018 (0.000)
	14	0.769; 0.248 (0.007)	1.486; 0.221 (0.049)	14.134; 1.597 (2.568)	0.9	0.785; 0.010 (0.000)	1.500; 0.005 (0.000)	0.901; 0.009 (0.000)

Figure 1 illustrates the distribution of circular datasets generated from simulated samples of size n = 100 under the vM and WC distributions, using the parameter settings reported in the figure caption. As shown in Figure 2, the WC distribution exhibits heavier tails than the vM distribution, which enables it to naturally accommodate outliers, even under high concentration (ρ) settings. This heavy-tailed structure reduces the influence of extreme observations and enhances robustness in circular regression models. Although classical kurtosis is not uniquely defined for circular distributions, the heavier-tailed behavior of the WC distribution is well documented through its theoretical properties and is consistently reflected in the simulated data. Although the vM distribution becomes more dispersed at low κ values, the WC distribution with low ρ still displays heavier tails, highlighting its superior robustness to outliers.

In circular datasets, observations may cluster around a central tendency, leading to a more homogeneous and concentrated structure with no apparent outliers. In such settings, standard models such as the vM and WC distributions provide accurate predictions with relatively low error rates. Table 1 reports the parameter estimates obtained from modeling a circular dataset without outliers using simple circular–circular regression models with vM and WC distributed circular errors.

Table 2. Robust and non-robust parameter estimates of the vM and WC simple circular-circular regression models with outliers (for n = 50).

		$\varepsilon_i \sim vM(0,\kappa)$						$\varepsilon_i \sim WC(0,\rho)$			
κ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
2	2	$x-$	0.867; 0.304 (0.048)	1.473; 0.096 (0.010)	2.008; 0.351 (0.123)	0.4	2	$x-$	0.982; 3.603 (0.114)	1.342; 3.334 (11.127)	0.417; 0.092 (0.009)
		$y-$	0.788; 0.256 (0.032)	1.501; 0.082 (0.006)	1.924; 0.328 (0.113)			$y-$	0.717; 1.794 (0.133)	1.471; 0.727 (0.529)	0.412; 0.083 (0.007)
		$x-y$	0.809; 0.409 (0.040)	1.515; 0.608 (0.370)	2.009; 0.361 (0.130)			$x-y$	0.980; 2.796 (0.138)	1.195; 6.714 (45.127)	0.418; 0.090 (0.008)
	5	$x-$	0.910; 0.327 (0.059)	1.452; 0.126 (0.018)	1.839; 0.323 (0.130)		5	$x-$	0.445; 15.882 (0.168)	0.514; 22.122 (489.880)	0.391; 0.102 (0.010)
		$y-$	0.784; 0.277 (0.037)	1.504; 0.076 (0.005)	1.707; 0.299 (0.175)			$y-$	0.816; 1.890 (0.155)	1.332; 3.100 (9.631)	0.386; 0.083 (0.007)
		$x-y$	0.809; 0.343 (0.042)	1.490; 0.092 (0.008)	1.866; 0.332 0.128			$x-y$	1.778; 12.867 (0.204)	1.385; 21.748 (472.522)	0.385; 0.107 (0.012)
4	2	$x-$	0.833; 0.240 (0.029)	1.483; 0.077 (0.006)	3.758; 0.709 (0.560)	0.7	2	$x-$	0.801; 0.217 (0.023)	1.495; 0.068 (0.005)	0.699; 0.052 (0.003)
		$y-$	0.791; 0.163 (0.013)	1.500; 0.045 (0.002)	3.433; 0.642 (0.733)			$y-$	0.788; 0.157 (0.012)	1.500; 0.043 (0.002)	0.696; 0.052 (0.003)
		$x-y$	0.792; 0.197 (0.019)	1.498; 0.062 (0.003)	3.730; 0.710 (0.576)			$x-y$	0.795; 0.172 (0.015)	1.497; 0.051 (0.003)	0.700; 0.051 (0.003)
	5	$x-$	0.870; 0.273 (0.040)	1.471; 0.090 (0.009)	3.129; 0.565 (1.077)		5	$x-$	0.901; 0.931 (0.036)	1.432; 1.007 (1.018)	0.673; 0.064 (0.005)
		$y-$	0.787; 0.174 (0.015)	1.503; 0.048 (0.002)	2.768; 0.484 (1.750)			$y-$	0.790; 0.169 (0.014)	1.500; 0.046 (0.002)	0.669; 0.055 (0.004)
		$x-y$	0.811; 0.222 (0.024)	1.492; 0.072 (0.005)	3.168; 0.582 (1.031)			$x-y$	0.848; 0.684 (0.028)	1.475; 0.240 (0.058)	0.676; 0.068 (0.005)

Table 2. *Cont.*

			$\varepsilon_i \sim vM(0,\kappa)$						$\varepsilon_i \sim WC(0,\rho)$		
κ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
8	2	$x-$	0.832; 0.208 (0.022)	1.484; 0.069 (0.004)	6.525; 1.591 (4.703)	0.8	2	$x-$	0.799; 0.411 (0.008)	1.500; 0.038 (0.001)	0.796; 0.038 (0.001)
		$y-$	0.787; 0.113 (0.006)	1.501; 0.031 (0.001)	5.564; 1.463 (8.070)			$y-$	0.786; 0.096 (0.005)	1.500; 0.026 (0.001)	0.794; 0.037 (0.001)
		$x-y$	0.798; 0.159 (0.012)	1.497; 0.052 (0.002)	6.462; 1.619 (4.984)			$x-y$	0.790; 0.107 (0.006)	1.498; 0.031 (0.001)	0.796; 0.037 (0.001)
	5	$x-$	0.856; 0.250 (0.033)	1.475; 0.083 (0.007)	4.542; 0.989 (12.936)		5	$x-$	0.826; 0.268 (0.016)	1.484; 0.058 (0.004)	0.775; 0.046 (0.003)
		$y-$	0.785; 0.129 (0.008)	1.502; 0.036 (0.001)	3.739; 0.797 (18.790)			$y-$	0.789; 0.101 (0.005)	1.500; 0.027 (0.001)	0.773; 0.041 (0.002)
		$x-y$	0.809; 0.190 (0.018)	1.492; 0.064 (0.004)	4.695; 1.156 (12.256)			$x-y$	0.800; 0.136 (0.009)	1.493; 0.071 (0.005)	0.778; 0.045 (0.003)
14	2	$x-$	0.819; 0.209 (0.022)	1.488; 0.069 (0.005)	9.461; 3.087 (30.116)	0.9	2	$x-$	0.785; 0.056 (0.002)	1.500; 0.016 (0.000)	0.896; 0.02 (0.000)
		$y-$	0.787; 0.090 (0.004)	1.500; 0.024 (0.000)	7.449; 2.411 (48.719)			$y-$	0.786; 0.044 (0.001)	1.500; 0.012 (0.000)	0.896; 0.02 (0.000)
		$x-y$	0.795; 0.149 (0.011)	1.497; 0.051 (0.002)	9.266; 3.078 (31.874)			$x-y$	0.787; 0.047 (0.001)	1.500; 0.014 (0.000)	0.896; 0.02 (0.000)
	5	$x-$	0.837; 0.227 (0.027)	1.481; 0.077 (0.006)	5.722; 1.584 (71.023)		5	$x-$	0.796; 0.080 (0.003)	1.496; 0.026 (0.001)	0.883; 0.025 (0.001)
		$y-$	0.788; 0.103 (0.005)	1.502; 0.029 (0.001)	4.405; 1.158 (93.391)			$y-$	0.787; 0.047 (0.001)	1.500; 0.013 (0.000)	0.883; 0.023 (0.001)
		$x-y$	0.810; 0.167 (0.014)	1.492; 0.059 (0.003)	5.765; 1.664 (70.582)			$x-y$	0.790; 0.059 (0.002)	1.498; 0.025 (0.001)	0.884; 0.026 (0.001)

[1] k is taken as 0.05 and 0.10, [2] OD = Outlier Directions.

It has been observed that when a circular data set without outliers is modeled using the vM simple circular-circular regression model, the variability of the estimator $\hat{\kappa}$ increases substantially as the true concentration parameter rises from $\kappa = 2$ to $\kappa = 14$. This behavior is a well-known property of the vM distribution. As $\hat{\kappa}$ increases, the likelihood surface becomes extremely steep, making the ML estimator highly sensitive to even very small deviations in the data. This sensitivity translates into reduced estimation stability, which can adversely affect the reliability of statistical inference for the vM model in highly concentrated settings. Even in the absence of outliers, the SD and MSE of $\hat{\kappa}$ increase with κ, and a similar pattern—although with smaller magnitude—is also observed in the estimates of $\hat{\alpha}$ and $\hat{\beta}$. As the sample size increases, parameter estimates become closer to the true parameter values, with reduced variability and error (Table 1).

Table 3. Robust and non-robust parameter estimates of the vM and WC simple circular-circular regression models with outliers (for n = 100).

$\varepsilon_i \sim vM(0,\kappa)$						$\varepsilon_i \sim WC(0,\rho)$					
κ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
2	5	$x-$	0.832; 0.239 (0.029)	1.485; 0.084 (0.007)	1.917; 0.235 (0.062)	0.4	5	$x-$	0.953; 1.794 (0.071)	1.405; 1.828 (3.349)	0.402; 0.060 (0.004)
		$y-$	0.801; 0.177 (0.015)	1.499; 0.049 (0.002)	1.841; 0.219 (0.073)			$y-$	0.790; 0.328 (0.052)	1.501; 0.090 (0.008)	0.394; 0.057 (0.003)
		$x-y$	0.815; 0.195 (0.019)	1.492; 0.060 (0.003)	1.929; 0.232 (0.059)			$x-y$	1.373; 15.146 (0.060)	1.201; 8.802 (77.480)	0.404; 0.063 (0.004)
	10	$x-$	0.846; 0.244 (0.031)	1.481; 0.078 (0.006)	1.786; 0.198 (0.085)		10	$x-$	0.985; 0.737 (0.093)	1.407; 0.551 (0.312)	0.386; 0.069 (0.005)
		$y-$	0.799; 0.187 (0.017)	1.501; 0.050 (0.002)	1.670; 0.195 (0.146)			$y-$	0.806; 0.499 (0.067)	1.504; 0.144 (0.021)	0.376; 0.057 (0.004)
		$x-y$	0.819; 0.197 (0.019)	1.491; 0.058 (0.003)	1.819; 0.209 (0.076)			$x-y$	0.956; 7.729 (0.116)	0.860; 14.680 (215.690)	0.383; 0.079 (0.007)
4	5	$x-$	0.799; 0.185 (0.017)	1.495; 0.060 (0.003)	3.441; 0.438 (0.503)	0.7	5	$x-$	0.794; 0.152 (0.011)	1.497; 0.048 (0.002)	0.690; 0.037 (0.001)
		$y-$	0.784; 0.116 (0.006)	1.502; 0.033 (0.001)	3.196; 0.385 (0.793)			$y-$	0.785; 0.107 (0.006)	1.500; 0.029 (0.001)	0.686; 0.037 (0.002)
		$x-y$	0.801; 0.143 (0.010)	1.496; 0.045 (0.002)	3.488; 0.467 (0.479)			$x-y$	0.791; 0.124 (0.008)	1.498; 0.037 (0.001)	0.690; 0.037 (0.001)
	10	$x-$	0.825; 0.208 (0.022)	1.486; 0.068 (0.004)	2.944; 0.335 (1.226)		10	$x-$	0.825; 0.172 (0.015)	1.486; 0.055 (0.003)	0.673; 0.039 (0.002)
		$y-$	0.787; 0.126 (0.008)	1.503; 0.034 (0.001)	2.665; 0.297 (1.869)			$y-$	0.787; 0.115 (0.007)	1.501; 0.031 (0.001)	0.666; 0.038 (0.003)
		$x-y$	0.812; 0.148 (0.011)	1.492; 0.047 (0.002)	3.024; 0.368 (1.087)			$x-y$	0.803; 0.138 (0.010)	1.494; 0.042 (0.002)	0.675; 0.038 (0.002)

Table 3. *Cont.*

			$\varepsilon_i \sim vM(0,\kappa)$						$\varepsilon_i \sim WC(0,\rho)$		
κ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
8	5	$x-$	0.802; 0.171 (0.014)	1.494; 0.057 (0.003)	5.539; 0.915 (6.890)	0.8	5	$x-$	0.787; 0.087 (0.004)	1.499; 0.027 (0.001)	0.790; 0.026 (0.001)
		$y-$	0.784; 0.079 (0.003)	1.502; 0.022 (0.000)	4.871; 0.761 (10.364)			$y-$	0.785; 0.065 (0.065)	1.500; 0.018 (0.018)	0.788; 0.027 (0.027)
		$x-y$	0.794; 0.115 (0.006)	1.498; 0.039 (0.001)	5.605; 0.973 (6.681)			$x-y$	0.787; 0.073 (0.002)	1.499; 0.022 (0.000)	0.790; 0.027 (0.001)
	10	$x-$	0.819; 0.190 (0.018)	1.489; 0.064 (0.004)	4.193; 0.589 (14.838)		10	$x-$	0.796; 0.106 (0.006)	1.496; 0.033 (0.001)	0.776; 0.028 (0.001)
		$y-$	0.792; 0.090 (0.004)	1.501; 0.025 (0.000)	3.599; 0.463 (19.579)			$y-$	0.787; 0.070 (0.002)	1.500; 0.019 (0.000)	0.772; 0.028 (0.002)
		$x-y$	0.813; 0.130 (0.008)	1.492; 0.045 (0.002)	4.324; 0.687 (13.986)			$x-y$	0.795; 0.087 (0.004)	1.497; 0.026 (0.001)	0.776; 0.028 (0.001)
14	5	$x-$	0.793; 0.164 (0.013)	1.497; 0.055 (0.003)	7.594; 1.618 (43.652)	0.9	5	$x-$	0.786; 0.037 (0.001)	1.500; 0.011 (0.000)	0.893; 0.014 (0.000)
		$y-$	0.788; 0.063 (0.002)	1.500; 0.018 (0.000)	6.386; 1.286 (59.628)			$y-$	0.785; 0.030 (0.000)	1.500; 0.008 (0.000)	0.892; 0.014 (0.000)
		$x-y$	0.801; 0.111 (0.006)	1.495; 0.038 (0.001)	7.658; 1.611 (42.808)			$x-y$	0.786; 0.032 (0.001)	1.500; 0.009 (0.000)	0.893; 0.014 (0.000)
	10	$x-$	0.812; 0.189 (0.018)	1.491; 0.064 (0.004)	5.107; 0.877 (79.856)		10	$x-$	0.788; 0.045 (0.001)	1.499; 0.014 (0.000)	0.884; 0.016 (0.001)
		$y-$	0.792; 0.075 (0.003)	1.501; 0.021 (0.000)	4.188; 0.655 (96.700)			$y-$	0.786; 0.033 (0.001)	1.500; 0.009 (0.000)	0.883; 0.016 (0.001)
		$x-y$	0.817; 0.112 (0.007)	1.490; 0.039 (0.001)	5.402; 1.036 (74.986)			$x-y$	0.788; 0.037 (0.001)	1.499; 0.011 (0.000)	0.884; 0.016 (0.001)

[1] k is taken as 0.05 and 0.10, [2] OD = Outlier Directions.

In contrast, when the same data set is modeled using the WC simple circular-circular regression model, a decrease in the value of ρ leads to a slight increase in the SD and MSE of the estimates of α, β and ρ. This occurs because, at low ρ values, the WC distribution becomes more dispersed and the likelihood surface flattens, increasing estimation uncertainty. As the sample size increases, however, this increase becomes more limited, indicating that larger samples partially compensate for the uncertainty associated with low concentration levels. Moreover, increases in sample size resulted in lower SD and MSE values for the parameter estimates in the WC model than in the vM model. The heavy-tailed nature of the WC distribution may have contributed to this numerical superiority by providing a better fit to the inherent variability in circular data (Table 1).

The presence of outliers in circular data sets leads to the formation of unusually extreme data points in the distribution, which can affect the estimation of model parameters. In such cases, these data sets should be modeled using heavy-tailed distributions such as WC that specifically minimize the impact of outliers to obtain robust estimates. Figure 3 visually illustrates the presence of outliers in the $x-$, $y-$ and $x-y$ directions in the circular data set for simple circular-circular regression models with vM and WC distributed circular errors.

Table 4. Robust and non-robust parameter estimates of the vM and WC simple circular-circular regression models with outliers (for n = 200).

$\varepsilon_i \sim vM(0,\kappa)$						$\varepsilon_i \sim WC(0,\rho)$					
κ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
2	10	$x-$	0.802; 0.184 (0.017)	1.495; 0.059 (0.003)	1.877; 0.158 (0.040)	0.4	10	$x-$	0.843; 0.302 (0.036)	1.514; 1.173 (1.375)	0.394; 0.041 (0.002)
		$y-$	0.790; 0.124 (0.007)	1.501; 0.034 (0.001)	1.816; 0.156 (0.058)			$y-$	0.796; 0.226 (0.025)	1.499; 0.063 (0.004)	0.386; 0.041 (0.002)
		$x-y$	0.811; 0.130 (0.009)	1.493; 0.039 (0.001)	1.897; 0.163 (0.037)			$x-y$	0.841; 0.832 (0.024)	1.490; 0.064 (0.004)	0.398; 0.039 (0.002)
	20	$x-$	0.811; 0.200 (0.020)	1.491; 0.065 (0.004)	1.742; 0.143 (0.087)		20	$x-$	0.889; 0.280 (0.043)	1.465; 0.100 (0.011)	0.382; 0.044 (0.002)
		$y-$	0.796; 0.130 (0.008)	1.501; 0.037 (0.001)	1.655; 0.142 (0.138)			$y-$	0.804; 0.240 (0.029)	1.502; 0.079 (0.006)	0.369; 0.042 (0.003)
		$x-y$	0.821; 0.133 (0.009)	1.490; 0.041 (0.001)	1.783; 0.151 (0.069)			$x-y$	0.926; 1.265 (0.045)	1.450; 0.798 (0.639)	0.384; 0.055 (0.003)
4	10	$x-$	0.794; 0.148 (0.011)	1.497; 0.048 (0.002)	3.324; 0.282 (0.536)	0.7	10	$x-$	0.789; 0.104 (0.005)	1.499; 0.032 (0.001)	0.686; 0.026 (0.001)
		$y-$	0.789; 0.084 (0.003)	1.500; 0.023 (0.000)	3.154; 0.259 (0.782)			$y-$	0.790; 0.075 (0.003)	1.499; 0.021 (0.000)	0.683; 0.026 (0.001)
		$x-y$	0.808; 0.099 (0.005)	1.492; 0.032 (0.001)	3.391; 0.295 (0.457)			$x-y$	0.798; 0.083 (0.004)	1.496; 0.025 (0.001)	0.687; 0.025 (0.001)
	20	$x-$	0.794; 0.157 (0.012)	1.497; 0.052 (0.002)	2.875; 0.219 (1.314)		20	$x-$	0.804; 0.116 (0.007)	1.494; 0.037 (0.001)	0.670; 0.027 (0.002)
		$y-$	0.786; 0.088 (0.003)	1.503; 0.024 (0.000)	2.649; 0.202 (1.865)			$y-$	0.791; 0.075 (0.003)	1.500; 0.021 (0.000)	0.663; 0.027 (0.002)
		$x-y$	0.816; 0.106 (0.006)	1.490; 0.034 (0.001)	2.959; 0.251 (1.145)			$x-y$	0.802; 0.089 (0.004)	1.495; 0.027 (0.001)	0.672; 0.027 (0.001)

Table 4. *Cont.*

			$\varepsilon_i \sim vM(0,\kappa)$						$\varepsilon_i \sim WC(0,\rho)$		
κ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\kappa}$; SD (MSE)	ρ	$n \times k$ [1]	OD [2]	$\hat{\alpha}$; SD (MCE)	$\hat{\beta}$; SD (MSE)	$\hat{\rho}$; SD (MSE)
8	10	$x-$	0.787; 0.127 (0.008)	1.499; 0.042 (0.001)	5.282; 0.569 (7.709)	0.8	10	$x-$	0.787; 0.059 (0.002)	1.499; 0.018 (0.000)	0.788; 0.019 (0.001)
		$y-$	0.790; 0.059 (0.001)	1.500; 0.016 (0.000)	4.805; 0.517 (10.473)			$y-$	0.788; 0.046 (0.001)	1.499; 0.013 (0.000)	0.786; 0.019 (0.001)
		$x-y$	0.806; 0.081 (0.003)	1.493; 0.026 (0.001)	5.495; 0.653 (6.700)			$x-y$	0.792; 0.051 (0.001)	1.498; 0.015 (0.000)	0.788; 0.018 (0.000)
	20	$x-$	0.803; 0.146 (0.010)	1.494; 0.049 (0.002)	4.033; 0.390 (15.891)		20	$x-$	0.796; 0.070 (0.003)	1.496; 0.022 (0.001)	0.774; 0.020 (0.001)
		$y-$	0.790; 0.065 (0.002)	1.502; 0.018 (0.000)	3.555; 0.318 (19.858)			$y-$	0.789; 0.046 (0.001)	1.499; 0.013 (0.000)	0.770; 0.020 (0.001)
		$x-y$	0.817; 0.089 (0.004)	1.490; 0.030 (0.001)	4.237; 0.449 (14.361)			$x-y$	0.794; 0.056 (0.002)	1.497; 0.017 (0.000)	0.775; 0.020 (0.001)
14	10	$x-$	0.795; 0.119 (0.007)	1.497; 0.040 (0.001)	7.113; 0.992 (48.405)	0.9	10	$x-$	0.786; 0.026 (0.000)	1.500; 0.008 (0.000)	0.892; 0.010 (0.000)
		$y-$	0.790; 0.045 (0.001)	1.500; 0.012 (0.000)	6.187; 0.762 (61.621)			$y-$	0.786; 0.022 (0.000)	1.500; 0.006 (0.000)	0.892; 0.010 (0.000)
		$x-y$	0.807; 0.073 (0.003)	1.493; 0.025 (0.000)	7.380; 1.114 (45.063)			$x-y$	0.787; 0.023 (0.000)	1.499; 0.007 (0.000)	0.892; 0.010 (0.000)
	20	$x-$	0.796; 0.138 (0.009)	1.496; 0.046 (0.002)	4.881; 0.568 (83.469)		20	$x-$	0.789; 0.030 (0.000)	1.499; 0.009 (0.000)	0.883; 0.011 (0.000)
		$y-$	0.787; 0.053 (0.001)	1.502; 0.014 (0.000)	4.148; 0.447 (97.262)			$y-$	0.787; 0.022 (0.000)	1.500; 0.006 (0.000)	0.882; 0.011 (0.000)
		$x-y$	0.817; 0.083 (0.004)	1.490; 0.029 (0.001)	5.195; 0.649 (77.951)			$x-y$	0.788; 0.025 (0.000)	1.499; 0.007 (0.000)	0.883; 0.011 (0.000)

[1] k is taken as 0.05 and 0.10, [2] OD = Outlier Directions.

Tables 2–4 summarize the robust and non-robust parameter estimates obtained from circular data sets contaminated with outliers in the $x-$, $y-$ and $x-y$ directions at different sample sizes, modeled using the vM and WC simple circular-circular regression models. The vM distribution with a low κ value indicates a more dispersed structure with low concentration, resulting in thicker tails. In contrast, high κ values cause data points to cluster more closely around the mean direction, producing a narrower structure with thinner tails. For circular datasets with outliers, a simple circular-circular regression model with a vM error term characterized by low κ values was employed. This approach yielded robust parameter estimates that were largely unaffected by the outliers (Table 2). As the sample size increased, the estimation errors decreased (Tables 3 and 4).

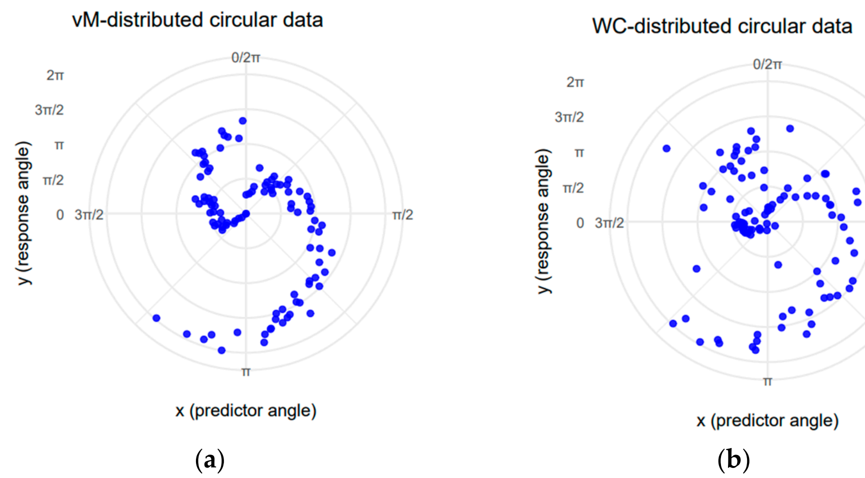


Figure 1. Circular data generated from simulated samples under the vM and WC distributions: (**a**) $vM(0, 8)$; (**b**) $WC(0, 0.7)$.

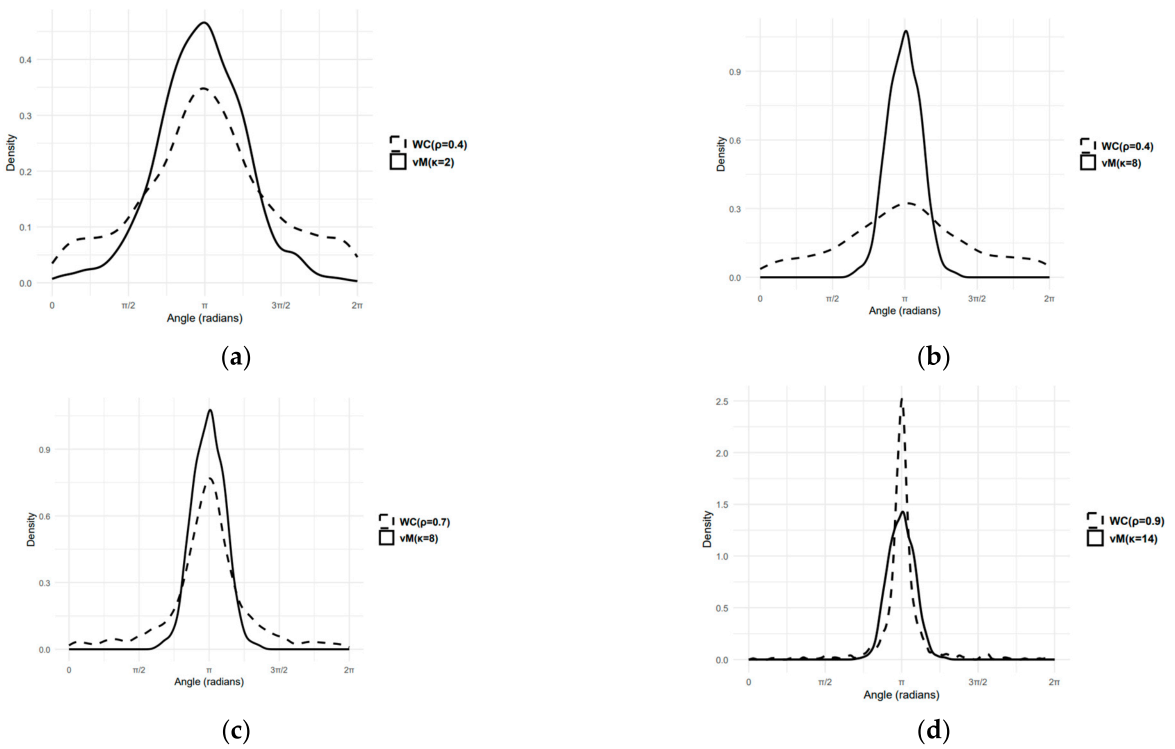
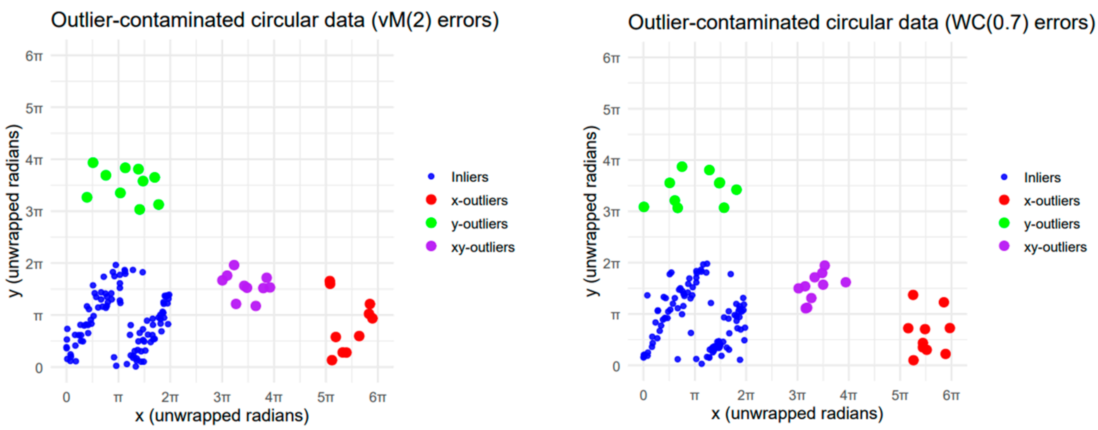


Figure 2. Visual comparisons of tails of the vM and WC distributions: (**a**) Comparison of tails of the $vM(\kappa = 2)$ and $WC(\rho = 0.4)$; (**b**) Comparison of tails of the $vM(\kappa = 8)$ and $WC(\rho = 0.4)$; (**c**) Comparison of tails of the $vM(\kappa = 8)$ and $WC(\rho = 0.7)$; (**d**) Comparison of tails of the $vM(\kappa = 14)$ and $WC(\rho = 0.9)$.

At high κ values (e.g., $\kappa = 8$ and $\kappa = 14$), vM-based estimates exhibit inflated variance and bias due to outlier sensitivity (Tables 2–4).

Increasing sample size did not mitigate high-κ instability, with non-robust behavior persisting as outliers increased (Tables 2–4).



(**a**) (**b**)

Figure 3. Outlying observations in $x-$, $y-$ and $x-y$ direction for vM and WC simple circular-circular regression models: (**a**) Outliers in $x-$, $y-$ and $x-y$ direction for simple circular-circular regression with $vM(2)$ error; (**b**) Outliers in $x-$, $y-$ and $x-y$ direction for simple circular-circular regression with $WC(0.7)$ error.

The WC distribution, due to its heavy-tailed nature, provides high tolerance to values that deviate substantially from the center (i.e., outliers) and allows these observations to contribute to the overall distribution. This feature is particularly useful when outliers are present in the modeled data. In modeling circular datasets contaminated with outliers using a simple circular-circular regression model with a WC error term, the value of the concentration parameter (ρ) plays a critical role. At low ρ levels, pronounced directional differences emerged depending on the direction of contamination. Outliers added in the $x-$ direction or $x-y$ direction produced substantially high SD and MSE values in the α and β estimates, primarily because extreme X values strongly distort the linear component $\alpha+\beta x$. Conversely, contamination in the $y-$ direction affected only the response variable, resulting in lower SD and MSE values compared to the other directions. As ρ increases, the model becomes markedly more robust to contamination; at high concentration levels, the sharper structure of the WC likelihood function stabilizes the estimation process, and the SD and MSE values of the α, β and ρ estimates remain small even under outliers from the $x-$, $y-$, or $x-y$ directions. These findings clearly illustrate that the WC model is highly robust to outliers under high-concentration scenarios. The heavy-tailed structure of the WC distribution persists even at high concentration values (high ρ), further enhancing the model's resistance to outliers (Tables 2–4).

This does not imply that the data are concentrated only at the center; rather, it accounts for the influence of outliers and allows them to be incorporated into the model without exclusion. This property plays a decisive role in correctly handling outliers and obtaining robust parameter estimates in circular regression models. It is worth noting that the numerical instability often observed for vM-based estimators under highly concentrated settings is not a limitation of the vM distribution itself but rather a consequence of the steepness of the corresponding likelihood surface in such regimes. This steep likelihood structure can pose challenges for ML estimation, particularly in the presence of contamination. Alternative robust estimation approaches, such as M-estimators and IRWLS-type procedures, have been proposed in the circular statistics literature to mitigate sensitivity to outliers. In the present study, we adopt a likelihood-based estimation framework to enable a unified and transparent comparison between vM- and WC-based circular–circular regression models under identical optimization settings.

4. Real Data Application

The real data application is based on a cohort of $n = 502$ gastric cancer patients. The circular response variable represents the month of surgical intervention, while the circular covariate corresponds to the month of diagnosis. Modeling these variables on the circle is clinically relevant, as calendar months exhibit an inherent periodic structure and may reflect seasonal patterns in disease detection and treatment scheduling.

Patient monitoring in healthcare is crucial for tracking a patient's health status and managing treatment processes. For cancer patients, paying attention to key dates, such as diagnosis, treatment, and surgery, plays a critical role in controlling the disease.

Gastric cancer stands out as the second most common malignancy among men and the fourth most common among women, both in Turkey and globally. This highlights that gastric cancer is a serious health issue worldwide. When detected early, survival rates for gastric cancer are significantly higher; however, these rates drop considerably when diagnosed at advanced stages. In the early stages, the five-year survival rate ranges from 90% to 100%. In advanced stages, this rate can drop to 15% to 25%. In Turkey, gastric cancer has an incidence rate of 7.4% among men and 6% among women. Every 100,000 men see 43 deaths, while 25 women per 100,000 die due to gastric cancer. These rates indicate that gastric cancer is more prevalent among men. The incidence of gastric cancer generally exhibits an increasing trend with age. Rarely observed before the age of 40, this disease becomes more prevalent with advancing age, reaching its peak particularly in the 60s [22]. Community screening programs in Japan play a critical role in the early detection of gastric cancer [23]. This enhances the likelihood of successful treatment and improves survival rates. Early detection of gastric cancer can increase patients' life expectancy by 90% and provide a 100% chance of recovery. Therefore, early diagnosis methods and regular screenings are crucial in the fight against gastric cancer [24]. Although there is some evidence suggesting that gastric cancer may be hereditary, only 4% of patients have a family history. The etiology of gastric cancer is complex, influenced by various dietary and environmental factors. Therefore, avoiding salty, processed, and smoked foods, abstaining from smoking, and maintaining a healthy diet are essential measures for reducing the risk of gastric cancer [25].

The data were obtained from a total of 502 patients diagnosed with gastric cancer who were presented at the Medical Oncology Department of a University Training and Research Hospital between 2000 and 2015. However, due to the absence of diagnosis or last examination dates for some patients, only 331 patients were included in the study. Although many variables were available for the patients, only four variables were considered for the circular analysis. The relationship between the diagnosis and surgery dates (in months) of patients diagnosed with gastric cancer and admitted to the Medical Oncology Department was tested using simple circular-circular regression analysis. Only the monthly portions of the dates were considered, excluding the year's information. This choice was motivated by the cyclic nature of calendar months, which naturally form a circular structure suitable for angular modeling. In contrast, the calendar year does not constitute a meaningful cycle in this dataset, as the observation period spans multiple years without repeating an annual pattern. Moreover, from a clinical perspective, the time elapsed between diagnosis and surgical intervention within the same year is of primary interest, as it reflects treatment scheduling and healthcare system responsiveness rather than long-term temporal trends across years. The data set contains outlier observations. Modeling of this data using the vM and WC simple circular-circular regression models was conducted with the open-source R software version 4.5.2. For the month's analysis, the circle was divided into 12 segments. The evaluation results were obtained by analyzing these slices.

Modeling of Gastric Carcinoma Data

In the modeling of the gastric carcinoma data (for month), a simple circular-circular regression model with vM error distribution was initially established, and parameter estimates were obtained. Due to the presence of outliers in the data set, a simple circular-circular regression model with a WC error distribution was also employed during the modeling process, yielding robust parameter estimates unaffected by outliers. For the real data application, α = 0.3 and β = 1 were used as reference parameter values for the vM and WC models, respectively, while the initial values were set to 1. Similarly, κ = 4 and ρ = 0.99 were used as reference concentration parameter values for both models, are the initial values were taken as 14 and 0.5, respectively. The results of the ML parameter estimate, along with the MCE and MSE values and the goodness of fit (GoF) values are presented in Table 5. The linear relationship between the circular dependent and circular independent variables for the vM simple circular-circular regression model was derived as follows:

$$\hat{y}_i = 0.239 + 0.990x_i \quad (mod2\pi) \quad , \quad i = 1, 2, \ldots, n, \tag{21}$$

where y_i represents the circular response variable and x_i is the circular predictor variable. GoF values for the considered vM and WC models were calculated using the following formula [26].

$$GoF = \frac{1}{n}\sum_{i=1}^{n} cos^2(y_i - \hat{y}_i) \tag{22}$$

Table 5. Modeling and comparing gastric carcinoma data using vM and WC simple circular-circular regression models.

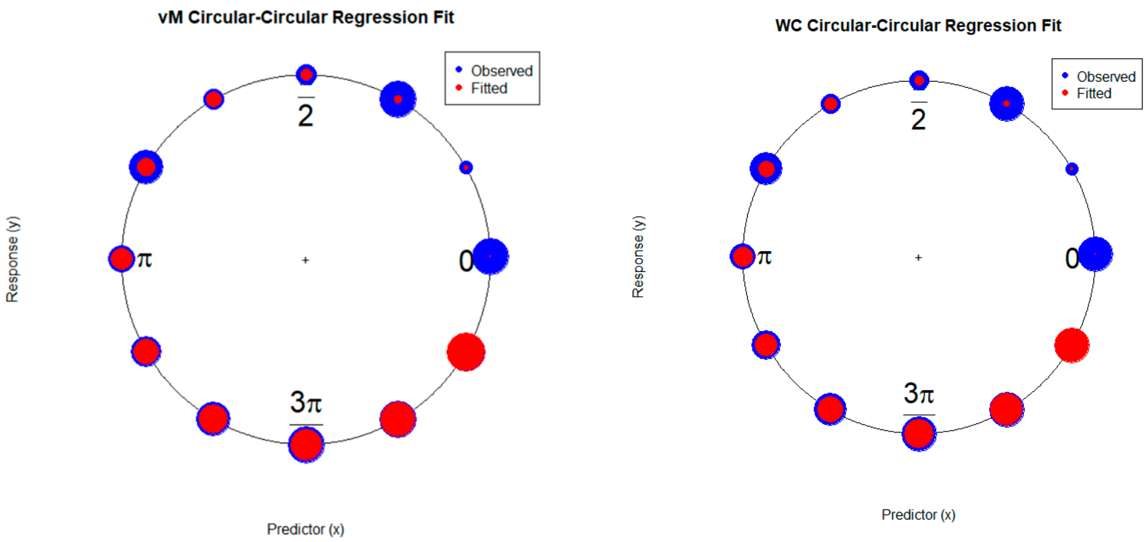
		$\varepsilon_i \sim vM(0,\kappa)$	$\varepsilon_i \sim WC(0,\rho)$
Parameter Estimates	$\hat{\alpha}$ (MCE)	0.239 (0.002)	−0.002 (0.045)
	$\hat{\beta}$ (MSE)	0.990 (9.936×10^{-5})	1.000 (1.036×10^{-8})
	concentration	$\hat{\kappa}$ = 4.180 (0.033)	$\hat{\rho}$ = 0.996 (4.408×10^{-5})
Goodness-of-fit	GoF	0.859	0.835
Mean Circular Error	MCE_model	0.109	0.136

This goodness-of-fit measure is based on circular error criteria and is therefore well suited for circular regression models, as it properly accounts for the periodic nature of angular data [27,28]. Unlike linear goodness-of-fit measures, this criterion relies on circular error measures, making it well suited for assessing model fit in circular response models.

From a clinical perspective, the slope parameter β reflects the strength of the temporal association between the month of diagnosis and the month of surgical intervention. Values close to one indicate a strong temporal alignment between diagnosis timing and subsequent treatment scheduling, suggesting that patients diagnosed in each period tend to undergo surgery within a closely related time window.

In the vM model, the estimate of the slope parameter is 0.990, which is slightly away from the true value of 1, whereas in the WC model, the estimate of the slope parameter is exactly 1.000 and exhibits a lower MSE value compared to the vM-based estimate. The estimated concentration parameters ($\hat{\kappa}$ and $\hat{\rho}$) are 4.180 and 0.996 for the vM and WC models, respectively. Notably, the concentration parameter estimate in the WC model exhibits a substantially lower MSE compared to that of the vM model. Although the GoF values of both models are close to each other, modeling the data using the WC model yields lower MCE and MSE values in parameter estimation (Table 5).

Given that the GoF measure is bounded between 0 and 1, the values reported in Table 5 indicate a reasonably good fit for both models, with only a modest difference between the vM and WC approaches. The graphical representations of both models fit to the data are shown in Figure 4.



(**a**) (**b**)

Figure 4. Fit of the vM and WC simple circular-circular regression models for gastric carcinoma data: (**a**) vM circular-circular regression fit; (**b**) WC circular-circular regression fit.

Figure 4 illustrates the fitted values obtained from the vM and WC circular–circular regression models in comparison with the observed data. While both models capture the overall circular trend, the WC model shows a closer alignment between observed and fitted points, particularly in regions affected by outliers, indicating a more robust fit.

5. Discussion

This study provides a comprehensive comparison of the vM and WC simple circular-circular regression models, which are commonly used in circular data analysis, highlighting the advantages of the WC model in terms of robustness. Notably, it has been illustrated that the WC model produces more consistent and reliable parameter estimates in data sets containing outliers. The simulation results clearly indicate the sensitivity of the vM model to outliers and deviations from model assumptions. In contrast, the WC model effectively mitigates the influence of such outliers, significantly improving estimation accuracy and model fit. In the real-world data application, the WC model outperformed the vM model in modeling the relationship between the months of diagnosis and operation for gastric cancer patients. It should be noted that the real data application focuses on the within-year cyclic relationship between diagnosis and surgery months. This modeling choice reflects the clinical relevance of seasonal and scheduling effects, rather than long-term calendar trends, and is well aligned with the circular nature of the variables under consideration. Accordingly, the exclusion of year information represents a deliberate methodological decision rather than a limitation of the modeling framework.

This finding underscores the WC model's effectiveness in fields such as biostatistics, where data structures are often complex and sensitive. The study also provides a solid methodological foundation supporting the practical applicability of the WC model. It is important to emphasize, however, that any potential increase in computational effort represents a trade-off for the improved robustness and reliability of the WC model in the presence of outliers and model deviations. In many practical applications, particularly in biomedical and observational studies where data contamination is common, the gains in estimation stability and accuracy may outweigh any additional computational cost. Additionally, the performance of the WC model may vary depending on the dataset size and data distribution. These observations suggest that further research is necessary to explore the application of this approach across various data structures and modeling scenarios.

6. Conclusions

This study establishes the WC simple circular–circular regression model as a robust alternative for analyzing circular data sets containing outliers, particularly in sensitive application areas such as biostatistical research. Both simulation results and the real data application demonstrate that the WC model yields more stable and reliable parameter estimates than the vM model under data contamination and deviations from model assumptions, while maintaining competitive performance in uncontaminated settings.

The advantages offered by the WC model make it a viable choice across various disciplines, including meteorology, astronomy, and biostatistics, where circular data structures are common and robustness is critical. The findings of this study provide a solid methodological basis for the practical use of heavy-tailed circular error distributions in regression modeling.

Future research may extend the proposed framework to more complex circular regression settings, such as multivariate or mixed circular–linear models, and explore alternative robust circular distributions. In addition, further studies could investigate computational efficiency and scalability, as well as apply the methodology to larger and more diverse real-world datasets, thereby advancing the development of robust circular regression methodologies.

Author Contributions: Conceptualization, A.K., M.A. and H.H.; Methodology, A.K. and M.A.; Software, M.A.; Validation, H.H.; Formal analysis, A.K., M.A. and H.H.; Investigation A.K., M.A. and H.H.; Resources, A.K.; Data curation, A.K., M.A. and H.H.; Writing—original draft, M.A.; Writing—review and editing, A.K. and H.H.; Visualization, M.A.; Supervision, M.A., A.K. and H.H.; Project administration, H.H., A.K. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are not publicly available due to privacy and ethical restrictions. Requests to access the data can be directed to the corresponding author (maltuntas@sinop.edu.tr) and will be subject to Institutional Review Board (IRB) approval and compliance with relevant privacy regulations.

Acknowledgments: The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

vM	*von* Mises
WC	Wrapped Cauchy
ML	Maximum Likelihood
WN	Wrapped Normal
i.i.d	Independently identically distributed

pdf	The probability density function
L	Likelihood function
L-BFGS-B	Limited-memory Broyden-Fletcher-Goldfarb-Shanno Box constraints
MCE	Mean circular error
MSE	Mean Squared error
SD	Standard deviation
GoF	The goodness-of-fit

References

1. Landler, L.; Ruxton, G.D.; Malkemper, E.P. Circular data in biology: Advice for effectively implementing statistical procedures. *Behav. Ecol. Sociobiol.* **2018**, *72*, 128. [CrossRef] [PubMed]
2. Landler, L.; Malkemper, E.P.; Painter, M.S.; Youmans, P.W.; Ruxton, G.D. Circular statistics in orientation biology. *J. R. Soc. Interface* **2019**, *16*, 20190213.
3. Protheroe, R.J. A new statistic for the analysis of circular data in gamma-ray astronomy. In Proceedings of the 19th International Cosmic Ray Conference, San Diego, CA, USA, 11–23 August 1985.
4. Rivest, L.-P. A decentred predictor for circular-circular regression. *Biometrika* **1997**, *84*, 717–729. [CrossRef]
5. Chassé, J.-L.; Théron, A. An example of circular statistics in chronobiological studies: Analysis of emergence rhythms. *Chronobiol. Int.* **1988**, *5*, 313–321. [CrossRef] [PubMed]
6. Lang, M.N.; Schlosser, L.; Hothorn, T.; Mayr, G.J.; Stauffer, R.; Zeileis, A. Circular regression trees and forests with an application to probabilistic wind direction forecasting. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2020**, *69*, 1357–1374. [CrossRef]
7. Johnson, R.A.; Wehrly, T.E. Some angular-linear distributions and related regression models. *J. Am. Stat. Assoc.* **1978**, *73*, 602–606. [CrossRef]
8. Fisher, N.I.; Lee, A.J. A correlation coefficient for circular data. *Biometrika* **1992**, *79*, 697–701. [CrossRef]
9. Downs, T.D.; Mardia, K.V. Circular regression. *Biometrika* **2002**, *89*, 683–697. [CrossRef]
10. SenGupta, A.; Ugwuowo, O. Asymmetric circular-linear multivariate regression models with applications to environmental data. *Environ. Ecol. Stat.* **2006**, *10*, 1–15. [CrossRef]
11. Alshqaq, S.; Abuzaid, A.H.; Ahmadini, A. Robust estimators for circular regression models. *J. King Saud Univ. Sci.* **2021**, *33*, 101576. [CrossRef]
12. Mardia, K.V.; Jupp, P.E. *Directional Statistics*; John Wiley & Sons: Chichester, UK, 1999.
13. Lee, D. Circular data. *WIREs Comput. Stat.* **2010**, *2*, 477–486. [CrossRef]
14. Bell, W.; Nadarajah, S. A review of wrapped distributions for circular data. *Mathematics* **2024**, *12*, 2440. [CrossRef]
15. Abuzaid, A.H.; Allahham, N.R. Simple circular regression model assuming wrapped Cauchy error. *Pak. J. Stat.* **2015**, *31*, 385–398.
16. Abuzaid, A.; Alshqaq, S.; Elkhazendar, M.; Elburai, M. Different distributional errors-in-circular-variables models. *Pak. J. Stat. Oper. Res.* **2021**, *17*, 577–589. [CrossRef]
17. R Core Team. *R: A Language and Environment for Statistical Computing*, version 4.5.2; R Foundation for Statistical Computing: Vienna, Austria, 2025. Available online: https://www.R-project.org/ (accessed on 20 January 2026).
18. Kato, S.; Shimizu, K.; Shieh, G.S. A circular–circular regression model. *Stat. Sin.* **2008**, *18*, 633–645.
19. Abuzaid, A.H.; Hussin, A.G.; Mohammed, I.B. Detection of outliers in simple circular regression model using the mean circular error statistic. *J. Stat. Comput. Simul.* **2013**, *83*, 269–277. [CrossRef]
20. Pewsey, A.; Neuhäuser, M.; Ruxton, G.D. *Circular Statistics in R*, 1st ed.; Oxford University Press: Oxford, UK, 2013; pp. 1–224.
21. Robert, C.P.; Casella, G. *Monte Carlo Statistical Methods*, 2nd ed.; Springer: New York, NY, USA, 2004; pp. 1–683.
22. Karaibrahimoglu, A.; Ayhan, S.; Karaagac, M.; Artac, M. Circular analyses of dates on patients with gastric carcinoma. *J. Appl. Stat.* **2021**, *48*, 2931–2943. [CrossRef] [PubMed]
23. Murakami, T. Early cancer of the stomach. *World J. Surg.* **1979**, *3*, 685–691. [CrossRef] [PubMed]
24. Arid, I.; Bentall, H.H.; Fraser Roberts, J.A. Relationship between cancer of stomach and the ABO blood groups. *Br. Med. J.* **1953**, *1*, 799–801. [CrossRef] [PubMed]
25. Özer, K. Mide Kanseri Olgularında Ameliyat Sonrası Sağkalıma Etki Eden Prognostik Faktörlerin Analizi. Master's Thesis, Şişli Etfal Training and Research Hospital, İstanbul, Turkey, 2008.
26. Abuzaid, A.H. Some Problems for Outliers Circular Data. Ph.D. Thesis, University of Malaya, Kuala Lumpur, Malaysia, 2010.

27. Fisher, N.I. *Statistical Analysis of Circular Data*; Cambridge University Press: Cambridge, UK, 1993.
28. Jammalamadaka, S.R.; Sengupta, A. *Topics in Circular Statistics*; World Scientific: Singapore, 2001.

https://doi.org/10.3390/math14030426