Sains Malaysiana 53(4)(2024): 935-951
http://doi.org/10.17576/jsm-2024-5304-16

Detection of Outliers in Circular Regression Model via *DFBETAc*$_{IS}$ Statistic

(Pengesanan *Outlier* dalam Model Regresi Bulat melalui Statistik *DFBETAc*$_{IS}$)

Intan Mastura Ramlee[1,2], Safwati Ibrahim[1,2,*], Leow Wai Zhe[3] & Mohd Irwan Yusoff[3]

[1]*Institute of Engineering Mathematics, Universiti Malaysia Perlis, Pauh Putra Main Campus, 02600 Arau, Perlis, Malaysia*
[2]*Centre of Excellence for Social Innovation and Sustainability (COESIS), Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia*
[3]*Faculty of Electrical & Technology Engineering, Universiti Malaysia Perlis, Pauh Putra Main Campus, 02000 Arau, Perlis, Malaysia*

Received: 4 July 2023/Accepted: 1 March 2024

ABSTRACT

The outlier issues in circular regression models have recently received much attention. The presence of outliers may cause the sign and magnitude of regression coefficients to vary, resulting in inaccurate model development and incorrect prediction. Many methods for detecting outliers in a circular regression model have been proposed in previous studies such as *COVRATIO*, *D*, *M*, *A,* and *Chord* statistics, but it is suspected that they are not very successful in the presence of multiple outliers in a data set since the masking and swamping is not considered in their studies. This study aimed to develop an outlier detection procedure using *DFBETAc* statistic for circular cases, where this new statistic will investigate and identify multiple outliers in the Jammalamadaka and Sarma circular regression model (JSCRM) by considering masking and swamping effect. Monte Carlo simulations are used to determine the corresponding cut-off point and the power of performance is investigated. The performance of the proposed statistic is evaluated by the proportion of detected outliers and the rate of masking and swamping. The simulation procedure is applied at 10% and 20% contamination levels for varying sample sizes. The results show that the proposed *DFBETAc*$_{IS}$ statistic for JSCRM successfully detect the outliers. For illustration purposes, this process is applied to wind direction data.

Keywords: Circular regression model; *DFBETAc*; outlier

ABSTRAK

Isu data terpencil dalam model regresi bulat baru-baru ini banyak mendapat perhatian. Kehadiran data terpencil boleh menyebabkan tanda dan magnitud pekali regresi berubah, mengakibatkan pembangunan model yang tidak tepat dan ramalan yang salah. Banyak kaedah untuk mengesan data terpencil dalam model regresi bulat telah dicadangkan dalam kajian sebelum ini seperti statistik *COVRATIO*, *D*, *M*, *A* dan *Chord* tetapi dipercayai bahawa kaedah tersebut tidak begitu berjaya dengan kehadiran berbilang data terpencil dalam set data kerana litupan dan limpahan tidak diambil kira dalam kajian mereka. Kajian ini bertujuan untuk membangunkan prosedur pengesanan data terpencil menggunakan statistik *DFBETAc* untuk kes bulatan dengan statistik baharu ini akan mengkaji dan mengenal pasti berbilang data terpencil dalam model regresi bulat Jammalamadaka dan Sarma (JSCRM) dengan mengambil kira kesan litupan dan limpahan. Simulasi Monte Carlo digunakan untuk menentukan titik potong yang sepadan dan kuasa prestasi dikaji. Prestasi statistik yang dicadangkan dinilai oleh perkadaran data terpencil yang dikesan dan kadar litupan dan limpahan. Prosedur simulasi digunakan pada tahap pencemaran 10% dan 20% untuk sampel saiz yang berbeza. Keputusan menunjukkan statistik *DFBETAc*$_{IS}$ yang dicadangkan untuk JSCRM berjaya mengesan data terpencil. Untuk tujuan ilustrasi, proses ini digunakan pada data arah angin.

Kata kunci: Data terpencil; *DFBETAc*; model regresi bulat

Introduction

An outlier is a data point that differs so much from the rest of the sample that neglecting it can lead to significantly inaccurate estimations (Chambers, Hentges & Zhao 2004). Outliers are common in real-world data, yet they typically go unrecognized since so much data is now handled by computers without proper review or filtering. Keypunch mistakes, missed decimal points, recording or transmission problems, unusual events such as earthquakes or strikes, or individuals of a distinct group sneaking into the sample can all cause outliers. A single observation that is also significantly different from all others can have a significant impact on regression analysis findings. Outliers in regression can result in an overestimation of the coefficient of determination, erroneous slope, and intercept values, and, in many cases, incorrect model conclusions. Outliers in regression are typically found using graphical approaches such as residual plots with erased residuals. The common methods for identifying outliers can misidentify good samples as outliers and fail to find true outliers when there are multiple outliers.

There are often two problems with methods of detecting outliers which are masking and swamping problems. Masking is the inability of the procedure to detect correct outliers and swamping is the identification of inliers as outliers. It is now evident that the presence of outliers causes misleading conclusions to be drawn from the results. Thus, researchers are interested in improving the ways of detecting outliers in statistical data. In linear regression models, many researchers have proposed methods to identify outliers. However, in circular regression models, there are only few methods in the study that develop methods for detecting outliers.

Circular regression is used to represent the connection between a circular dependent and a collection of circular independent variables. Circular regression methods have been used in a variety of fields, including crystallography by MacKenzie (1957), vector cardiography by Downs (1974), making a prediction of the direction of ground movement during an earthquake by Rivest (1997), and research of circadian biological rhythms, where a 24-hour clock has deemed a circle (Binkley 1990; Downs 1974; Moore-Ede, Sulzman & Fuller 1982). Medical imaging (Jones & Silverman 1989; Weir & Green 1994) and circadian timing of cancer therapy to reduce the number and severity of toxic side effects (Hrushesky 1985) are two examples of medical applications. According to recent studies on the genetic and molecular aspects of mammalian circadian rhythms, stronger circular regression models are required (Lowrey et al. 2000; Shearman et al. 2000). According to new studies on the genetic and molecular aspects of mammalian circadian rhythms, stronger circular regression models are required (Lowrey et al. 2000; Shearman et al. 2000).

One of the first angular-linear regression models was proposed by Gould (1969). Mardia (1975) devised a nonparametric rank correlation coefficient for circular data. Johnson and Wehrly (1978) modified the Gould model by limiting the range of the independent variables to the half-open interval $(0, 2\pi)$. They may be found in a variety of scientific domains, including meteorology and biology. To analyze data from spatial rock magnetism, Stephens (1979) employed a directed regression strategy. By mapping the real line onto the unit circle, Fisher and Lee (1992) employed a link function to generalize Johnson and Wehrly's model. Follman and Proschan (1999) employed correlated successive seizure timings on the same persons to test for circadian circular uniformity of epileptic seizure durations, whereas Lund (1999) proposed a regression model with one circular variable and a set of linear variables as independent variables.

Jammalamadaka and Sarma (1993) investigated the conditional expectation of the vector e^{iv} given u, describing it in terms of Fourier series expansions with errors assumed to follow a normal distribution. In the meanwhile, Hussin, Fieller and Stillman (2004) proposed a simple approach for determining the linear relationship between two circular variables. Outliers in bivariate circular data, like outliers in linear data, can make parameter estimation and forecasting more difficult. By studying the changes in the covariance matrix of parameters obtained by deleting one data at a time, Abuzaid et al. (2012) discovered outliers in a simple circular regression model. Mohamed et al. (2016) suggested a new discordancy test approach based on the A, C, M statistics and the idea of spacing to find outliers in circular data and patches of outliers.

Meanwhile, a single-case deletion statistic, such as *DFBETAS*, *DFFITS*, externally studentized residual, Cook's distance, and the criterion (Belsley, Kuh & Welsch 1980; Barnett & Levis 1994; Chatterjee & Hadi 1988; Cook 1977), are one of the conventional ways for detecting outliers. However, a single-case deletion method is possible to have masking or swamping problems, and tests based on it will disappear their power significantly when there are multiple outliers in the data set. Outlier detection in circular data requires a different method than outlier detection in linear data. Two previous works, Ibrahim et al. (2013) and Abuzaid, Mohamed and Hussin (2009) focus on the detection of outliers in circular regression models using *COVRATIOc*, *DMCEc* and *DMCEs* statistics. Ibrahim (2013) and Ibrahim et al. (2013) explored the issue of outliers in JSCRM. Then, Alkasadi et al. (2019, 2018, 2016) investigated these issues and proposed various statistics for detecting outliers in multiple circular regression models (MCRM). Abuzaid (2020) extends the local outlier factor (LOF) application to the detection of potential outliers in circular samples, where the angles of the circular data are represented in two Cartesian coordinates and treated as bivariate data. Recently, Meilán-Vila, Crujeiras and Francisco-Fernández (2021) considers a spatial regression model with a circular response and several real-valued predictors and proposes nonparametric estimators of the circular trend surface that account for spatial correlation. Jha, Biswas and Cheng (2022) propose a paper that uses the Möbius transformation-based link function to address robustness issues in circular-circular regression. An exact polynomial time algorithm is then suggested for calculating the maximum trimmed cosine estimator in this context, as well as the estimator's breakdown point. However, there is a lack of study on the problem of numerous outliers in JSCRM. The *DFBETAc* of MCRM is extended in this work to detect possible outliers in the JSCRM. The main goal of this study is to identify outliers for a single independent circular variable using the *DFBETAc* statistic that has not been investigated before.

The focus of the study was on the effect of outliers and detection methods on circular regression models. The circular regression model is discussed, as well as the least-squares approach for parameter estimation. The *DFBETAc* statistic for the JSCRM is then demonstrated. The cut-off points are then obtained, and the suggested statistic's performance is examined. Detecting outliers in a real-world data set is crucial for increasing the quality of the original data and reducing the impact of outliers. Finally, identifying outliers in wind direction data is discussed as an example.

THE JAMMALAMADAKA AND SARMA REGRESSION MODEL (JSCRM)

In circular regression, there are a variety of methods for detecting outliers. Ibrahim (2013) introduced JSCRM for two circular random variables $\boldsymbol{U}$ and $\boldsymbol{V}$. To predict v for a given u, consider the conditional expectation of vector e^{iv} given u

$$E\left(e^{iv}|u\right)=\rho(u)e^{iv(\boldsymbol{u})}=g_1(u)+g_2(u) \tag{1}$$

where $e^{iv} = \cos v + i \sin v$, $\mu(u)$ is the conditional mean direction of v given u and $\rho(u)$ is the conditional concentration parameter. Then, estimate the parameter $\mu(u)$ and $\rho(u)$ such

$$\mu(u)=\hat{v}=\arctan^{*}\frac{g_2(u)}{g_1(u)}$$

$$=\begin{cases}\tan^{-1}\dfrac{g_2(u)}{g_1(u)} & \text{if } g_1(u)\leq 0\\ \pi+\tan^{-1}\dfrac{g_2(u)}{g_1(u)} & \text{if } g_2(u)\leq 0\\ \text{underfined} & \text{if } g_1(u)=g_2(u)=0.\end{cases} \tag{2}$$

The values of $g_1(u)$ and $g_2(u)$ estimated using the following trigonometric polynomials of a suitable degree (m),

$$\begin{aligned} g_1(u)&\approx\sum_{k=0}^{m}\left(A_k\cos ku+B_k\sin ku\right)\\ g_2(u)&\approx\sum_{k=0}^{m}\left(C_k\cos ku+D_k\sin ku\right). \end{aligned} \tag{3}$$

So, the following two observational regression-like models according to equation (3);

$$\begin{aligned} V_{1j}&=\cos v_j=\sum_{k=0}^{m}\left(A_k\cos ku_j+B_k\sin ku_j\right)+\varepsilon_{1j}\\ V_{2j}&=\sin v_j=\sum_{k=0}^{m}\left(C_k\cos ku_j+D_k\sin ku_j\right)+\varepsilon_{2j} \end{aligned} \tag{4}$$

where $\boldsymbol{\varepsilon}=\varepsilon_1+\varepsilon_2$ is the vector of random error following the bivariate normal distribution with mean $\boldsymbol{0}$ and the unknown dispersion matrix $\boldsymbol{\Sigma}$. The generalized least-squares method can be used to estimate the parameters A_k, B_k, C_k and D_k for $k = 0, 1, ..., m$, the standard error, as well as the dispersion matrix $\boldsymbol{\Sigma}$. Assume that $A_0 = C_0 = 0$ to ensure identifiability.

Therefore, the summary of observational equation (4) as

$$\begin{aligned} \boldsymbol{V}^{(1)}&=\boldsymbol{U}\boldsymbol{\lambda}^{(1)}+\boldsymbol{\varepsilon}^{(1)}\\ \boldsymbol{V}^{(2)}&=\boldsymbol{U}\boldsymbol{\lambda}^{(2)}+\boldsymbol{\varepsilon}^{(2)} \end{aligned} \tag{5}$$

Thus, the least-square estimation turns out to be given by

$$\begin{aligned} \hat{\boldsymbol{\lambda}}^{(1)}&=\left(\boldsymbol{U'U}\right)^{-1}\boldsymbol{U'V}^{(1)}\\ \hat{\boldsymbol{\lambda}}^{(2)}&=\left(\boldsymbol{U'U}\right)^{-1}\boldsymbol{U'V}^{(2)} \end{aligned} \tag{6}$$

where the matrix combination cosine and sine came from $\boldsymbol{U}$. The covariance matrix $\boldsymbol{\Sigma}$ is estimated as follows:

$$\hat{\boldsymbol{\Sigma}} = \left[n - 2(2m+1)\right]^{-1} R_0 \tag{7}$$

where $R_0 = (p,q) = \boldsymbol{V}^{(p)}\boldsymbol{V}^{(q)} - \boldsymbol{V}^{(p)}\boldsymbol{U}(\boldsymbol{U'U})\boldsymbol{U'V}^{(q)}$ and $R_0 = \left(R_0(p,q)\right)_{p,q=1,2}$ is unbiased estimation of $\boldsymbol{\Sigma}$. According to Belsley, Kuh and Welsch (1980), any observation is considered an outlier if it meets certain conditions, such as $\left|DFBETAc_{IS\,j,i}\right| > 2/\sqrt{n}$ where n is the sample size (Cousineau & Chartier 2010; Rousseeuw & Leroy 2005).

$\boldsymbol{DFBETAC_{IS}}$ STATISTIC IN JSCRM

There are numerous techniques for identifying outliers in circular regression, such as the *DFBETAS* statistic, which shows how much the regression coefficient, would change if the i^{th} observation were deleted. Belsley, Kuh and Welsch (1980) proposed a measure for calculating how much an observation changed the estimation of the *DFBETAS* regression coefficient in a linear case. In the circular case, the $DFBETAc_{IS}$ statistic is presented as follows:

$$DFBETAc_{IS\,j,i} = \frac{\hat{\boldsymbol{\lambda}}_j - \hat{\boldsymbol{\lambda}}_{j(i)}}{\sqrt{S_{(i)}^2 C_{jj}}} \tag{8}$$

For i = 1, 2, ..., n, where $\hat{\boldsymbol{\lambda}}_j$ is the prediction from the full regression model for the i^{th} observation, and $\hat{\boldsymbol{\lambda}}_{j(i)}$ when the i^{th} observation is deleted. $S_{(i)}^2$ denoted the standard error computed without the point i and $\boldsymbol{C}_{jj}$ is the j^{th} diagonal element of $(\boldsymbol{U'U})^{-1}$ where matrix $\boldsymbol{U}$ is the combination of cosine and sine functions,

$$\underset{n\times(4m+1)}{\boldsymbol{U}} = \begin{bmatrix} \underset{n\times 1}{1} & \underset{n\times m}{\boldsymbol{CC}} & \underset{n\times m}{\boldsymbol{CS}} & \underset{n\times m}{\boldsymbol{SC}} & \underset{n\times m}{\boldsymbol{SS}} \end{bmatrix} \tag{9}$$

The relationship between a circular independent variable and a circular dependent variable is examined in this study (Ibrahim 2013). The row deletion approach was used to construct the procedure for detecting outliers in JSCRM. If there are outliers in the data, the character of parameter estimates, the variance of the residual and covariance matrix, and the standard error will be impacted in regression.

DETERMINATION OF CUT-OFF POINT BY $\boldsymbol{DFBETAC_{IS}}$ STATISTIC

The cut-off points of the $DFBETAC_{IS}$ statistic to identify the outliers in JSCRM were obtained by simulation study using SPlus statistical software. In this section, the number of simulations is set to 1000 for each sample size, n, and standard deviation ($\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$). Let's consider the case when m = 1 to obtain the parameter estimates. Fifteen different sample sizes of n = 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140 and 150 with the combination of standard deviation ($\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$) in the range of **[0.03, 0.3]**, respectively, are used (Alkasadi et al. (2018)). For various combinations of ($\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$), a set of random errors from the bivariate Normal distribution with mean $\boldsymbol{0}$. The following are the entire processes for obtaining the cut-off points:

Step 1. Generate a variable $\boldsymbol{U}$ of size n from VM (π, 3) and VM (π, 2), respectively.

Step 2. Generate $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ of size n from $N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}\right)$. For a fixed $\alpha = n$ obtain the true values of $\lambda = A_0$, A_1, B_1, C_0, C_1 and D_1. Here, let the true values of A_0 and C_0 to be zero. Then, calculate $\boldsymbol{V}_{1j}$ and $\boldsymbol{V}_{2j}$ j = 1, ..., n using Equation (1).

Step 3. Obtain the circular variable $v_j = \arctan\left(\frac{\boldsymbol{V}_{2j}}{\boldsymbol{V}_{1j}}\right)$, j = 1, ..., n using equation (2).

Step 4. Fit the generated circular data using JSCRM to give the parameter estimates of $\lambda = \hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1$ and $\hat{D}_1$.

Step 5. Exclude the i^{th} row from the generated circular data, where i = 1, ..., n. For each i, repeat step 4 for the reduced data set to obtain $\hat{\lambda}_{j(i)}$.

Step 6. Calculate $|DFBETAc_{IS\,j,i}|$ for each i from Equation (8).

Step 7. State the maximum value of $|DFBETAc_{IS\,j,i}|$. The upper percentiles of the maximum values $|DFBETAc_{IS\,j,i}|$ are computed at the 1%, 5%, and 10% levels. For the JSCRM, the upper percentiles are used as cut-off points to identify outliers.

POWER PERFORMANCE OF $\boldsymbol{DFBETAC_{IS}}$ STATISTIC

A Single Outlier

The performance of $|DFBETAc_{IS\,j,i}|$ detecting outliers in the JSCRM based on Equation (8) was investigated in a simulated study. There are six sample sizes to consider: n = 20, 30, 40, 50, and 130. To create the data, the same process is used as in the previous section. Then, at position d, say v_d the observation becomes contaminated as follows:

$$v_d^* = v_d + \lambda\pi \quad \mod(2\pi) \tag{12}$$

where v_d^* denotes the value after contamination and τ denotes the degree of contamination, which can range from $0 \le \lambda \le 1$. The parameter estimates of $\hat{A}_0, \hat{A}_1, \hat{B}_1, \hat{C}_0, \hat{C}_1$ and $\hat{D}_1$ was calculated using the generated data of $\boldsymbol{U}$ and $\boldsymbol{V}$. $|DFBETAc_{IS\,j,i}|$ is also determined using an Equation (8). As a result, j = 1, ..., n remove the j^{th} row from the sample and fit the remaining data using Equation (1). If the $|DFBETAc_{IS\,j,i}|$ values are larger than the cut-off points calculated from the relevant cut-off point, the product has effectively detected the outlier in the data. 1000 times the process is repeated. By calculating the percentage of the correct detection of the contaminated observation at position d, the procedure's power performance is evaluated. Table 1 presents the cut-off points of 5% upper percentile for different n with standard deviation (σ_1, σ_2) = (0.3, 0.3) at α = 2.

Figure 1 shows the $DFBETAc_{IS}$ detection method's power of performance for n = 70 and four different values of (σ_1,σ_2) = (0.03, 0.03), (0.05, 0.05), (0.1, 0.1) and (0.3, 0.3). The procedure's performance improves as σ_1 and σ_2 grow smaller. This is predicted because when σ_1 and σ_2 are approach zero, V_{1j} and V_{2j} in Equation (4) move closer to the horizontal axis, increasing the chance of detecting the outlier even when λ is small.

On the other hand, Figure 2 shows a plot of the $DFBETAc_{IS}$ detection method power of performance for fixed (σ_1, σ_2) = (0.1, 0.1) and different values of n = 30, 50, 70, 100, 120, 150. When n is high enough, the performance increases as a function of n, but the curves are quite near to each other. Similar trends may be seen in other cases as well.

TABLE 1. The 5% upper percentiles of $|DFBETAc_{IS\,j,i}|$ statistics for(σ_1, σ_2) = (0.3, 0.3) at α = 2

Sample size, n	$\hat{A}_0$	$\hat{A}_1$	$\hat{B}_1$	$\hat{C}_0$	$\hat{C}_1$	$\hat{D}_1$
20	2.0980	1.7896	3.0366	3.4346	2.7441	4.3725
30	2.1954	1.6869	2.7129	3.5201	2.8039	4.5642
40	2.6095	2.2507	2.6054	3.6402	3.0005	4.6363
50	2.8614	2.4148	2.7564	4.8738	4.1269	4.5967
60	3.1022	2.6265	2.9515	4.3298	3.4112	4.7651
70	2.9871	2.6772	3.0660	4.9623	4.2618	4.8137
80	3.0021	2.6065	3.2255	4.2042	3.5514	4.5776
90	2.9236	2.4559	3.2702	4.7140	3.9039	4.6479
100	3.1213	2.6280	3.1146	4.7774	4.0793	5.1480
110	3.5851	3.0646	3.2382	4.4520	3.6242	5.0003
120	3.8318	3.1921	3.2617	5.1182	4.2121	5.1819
130	3.3899	2.8531	2.5251	4.3032	3.5472	4.3054
140	3.5386	2.9997	2.8332	4.9036	4.1402	5.0691
150	3.5165	2.9737	2.8610	4.6200	3.8163	4.7701

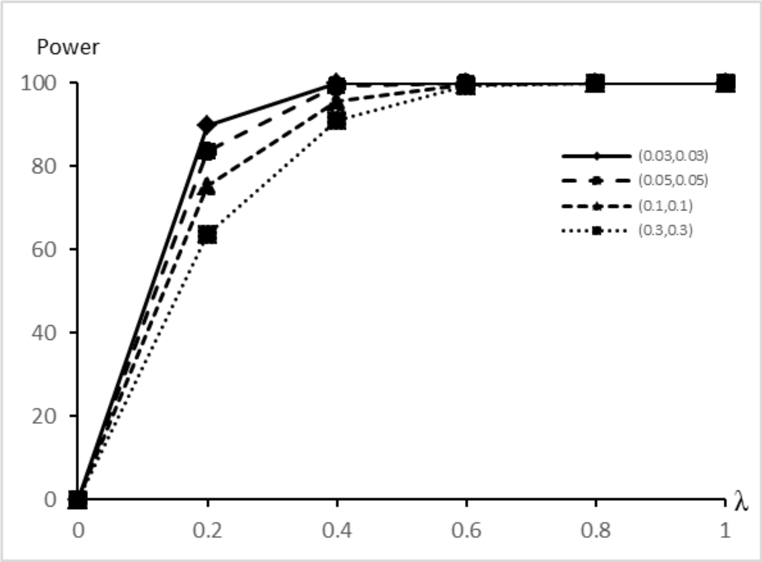


FIGURE 1. Power performance of the $|DFBETAc_{IS\,j,i}|$ statistic for $n = 70$ at $\alpha = 2$

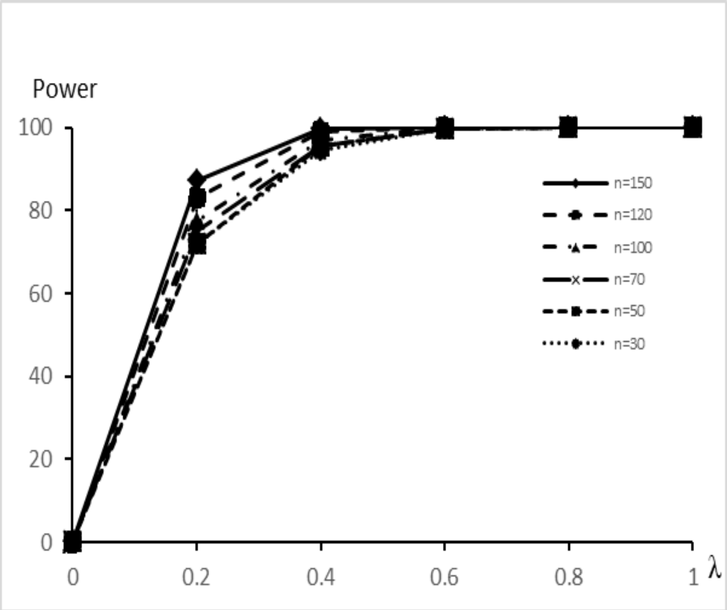


FIGURE 2. Power performance of the $|DFBETAc_{IS\,j,i}|$ for $(\sigma_1, \sigma_2) = (0.1, 0.1)$ at $\alpha = 2$

Multi-Outlier

Two contamination ratios (10% and 20%) were chosen, repeated 1000 times for the same combinations of sample sizes and concentration parameters, with 5% of the $DFBETAc_{IS}$ statistics cut-off points were used to assess the performance of statistics for various contamination ratios. Three measurements are taken into consideration to determine the performance of all the statistics: the proportion of detected outliers, the masking and swamping rates. In every replication cycle, it is noted that the number of real outliers is found. As a result, the proportion of outliers is determined as follows:

$$\text{Proportion of outliers} = \frac{\text{Sum of detected true outliers}}{P \times n \times 1000}$$

where P is the percentage of contamination. Similarly, to calculate the rate of masking and the rate of swamping, it is observed that the number of generated outliers detected as inlier (clean observation) and the number of inliers detected as outlier, respectively as follows:

$$\text{Rate of masking} = \frac{\text{Sum of detected outliers as inliers}}{P \times n \times 1000}$$

$$\text{Rate of swamping} = \frac{\text{Sum of detected outliers as outliers}}{\left(n - \left(P \times n\right)\right) \times 1000}$$

The approach with the lowest rates of masking and swamping and the highest rate of outlier detection is considered the good method. In Figures 3(a)-3(f) and 4(a)-4(f), the proportion of detected outliers is displayed together with the rate of masking and swamping with a 5% upper percentile for n = 70 and 100. For all statistics, the rates of swamping are zero or very nearly zero, according to Figures 3 and 4. However, for every combination of sample sizes, degree of contamination, and outlier proportion, the rates of masking of the $DFBETAc_{IS}$ statistic are very high and the proportions of outliers detected are very low. It was notice that the proportion of outliers detected of the $DFBETAc_{IS}$ statistics is low when the λ less than 0.6 with 10% contamination and this proportion significantly decrease with 20% contamination. Consequently, it has a high rate or masking. The (σ_1,σ_2) = (0.03, 0.03) relatively has a higher proportion of outliers detected than the (σ_1,σ_2) = (0.1, 0.1) and the proportion increase with the λ but it is low at 20% contaminated. The proportion of outliers detected of the proposed (σ_1,σ_2) = (0.03, 0.03) is relatively low for small value of the λ. This is acceptable as when the λ is low, the circular data will be dispersed over the circle's circumference. Consequently, it is very difficult to identify outliers in this case (Collett 1980). The (σ_1,σ_2) = (0.03, 0.03) gives a greater proportion of outliers detected than the others (σ_1,σ_2), as expected. The proportion is an increasing function of the λ and increase to 100% for values of the λ greater than 0.4. Therefore, the rate of masking is very low and is a decreasing function of the λ, decreasing down to 0%.

In general, the proposed $DFBETAc_{IS}$statistics is very successful in detection of outliers. The (σ_1,σ_2) = (0.03, 0.03) is the best when compared to the other three measures. It has the highest proportion of outliers detected and the lowest rate both masking and swamping.

PRACTICAL EXAMPLE: WIND DIRECTION DATA

In Hussin, Fieller and Stillman (2004), 129 observations of wind direction were recorded using two different instruments: an HF radar system and an anchored wave buoy along the Holderness coastline (the Humberside coast of the North Sea, United Kingdom). The least-square estimates of the parameters are $\hat{A}_0 = 0.0674$, $\hat{A}_1 = 0.7559$, $\hat{B}_1 = -0.0948, \hat{C}_0 = -0.047, \hat{C}_1 = 0.1049$, $\hat{D}_1 = 0.9762$, $\hat{\sigma}_1 = 0.3$ and $\hat{\sigma}_2 = 0.3$ and thus the fitted model gives $\hat{g}_1(u)$ and $\hat{g}_2(u)$ are follow

$$\hat{g}_1(u) = 0.0674 + 0.7559\cos u - 0.0948\sin u$$
$$\hat{g}_2(u) = -0.047 + 0.1049\cos u + 0.9762\sin u.$$

Further, the model $E\left(e^{iv}|u\right) = \rho(u)e^{i\mu(u)} = g_1(u) + ig_2(u)$ is obtained such that

$$\mu(u) = \hat{v}_j = \arctan^* \frac{-0.047 + 0.1049\cos u_j + 0.9762\sin u_j}{0.0674 + 0.7559\cos u_j - 0.0948\sin u_j},$$
$$j = 1,\ldots,n,$$

and the concentration parameter $\rho(u)$ is obtained using Equation (5). The estimated concentration parameter is given by $\hat{\rho}(u) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\rho^2(u)} = 0.9322.$

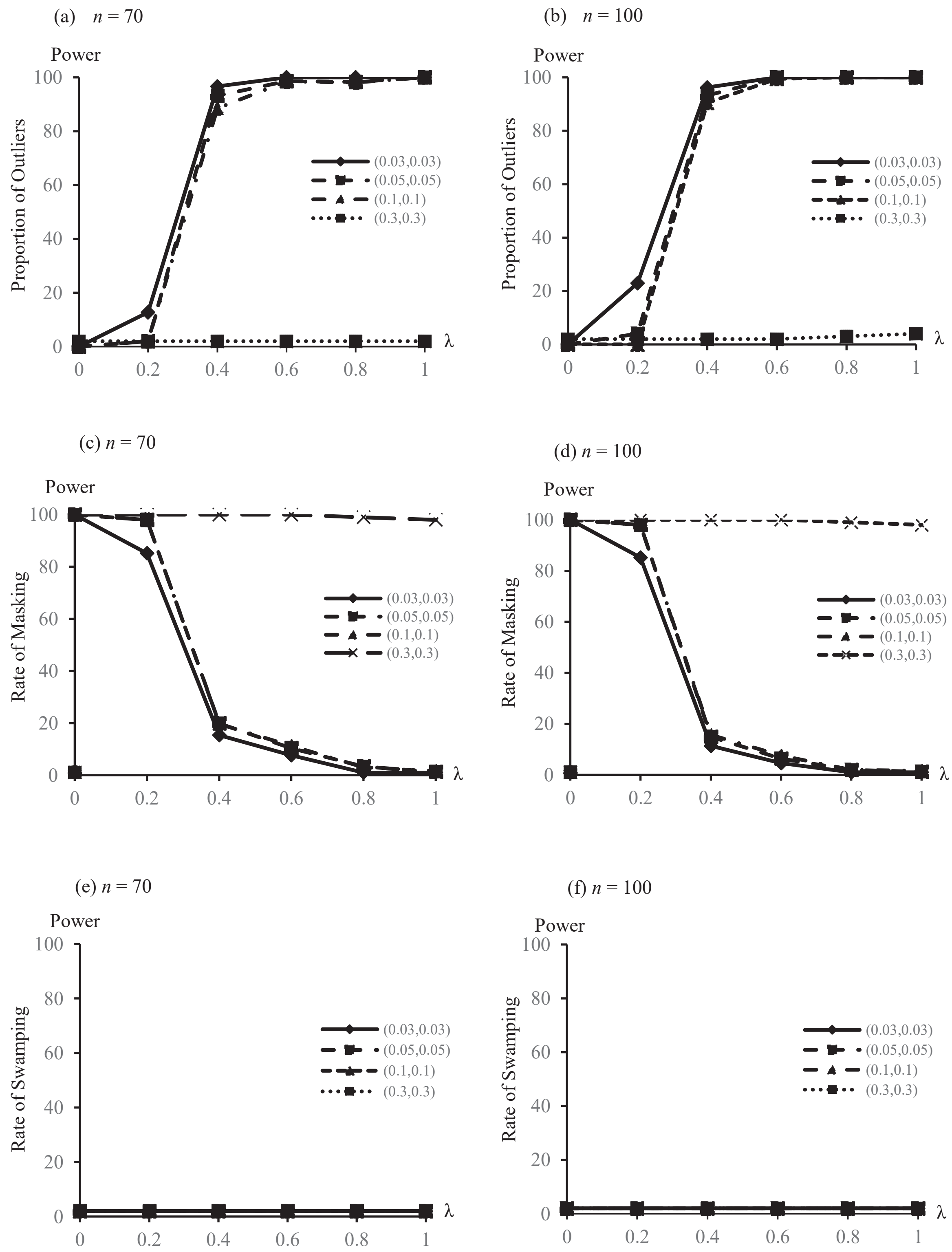


FIGURE 3. (a-f). The proportion of outliers detected and rate of masking and swamping with 5% cut-off points and 10% contamination for *n* =70 and 100

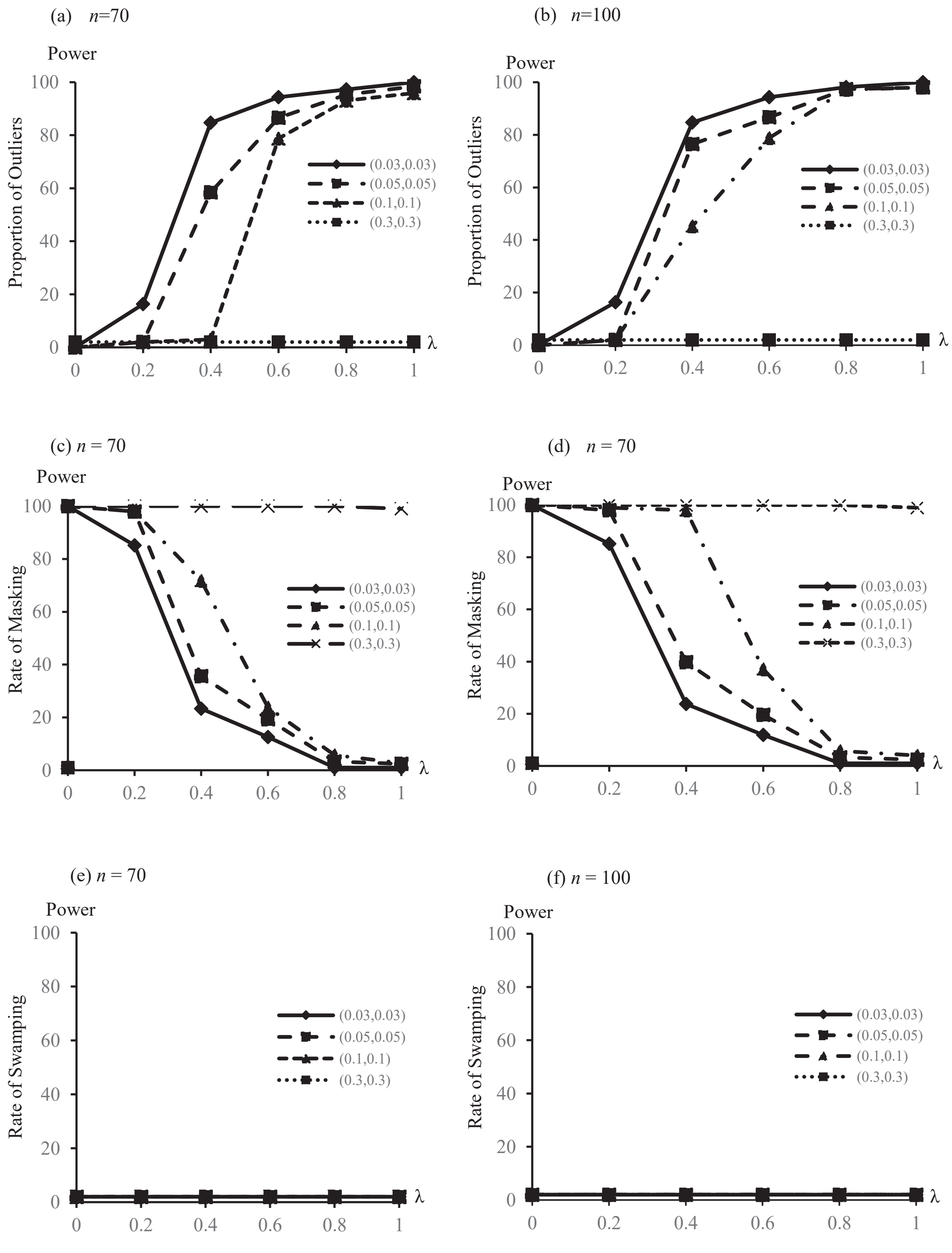


FIGURE 4. (a-f). The proportion of outliers detected and rate of masking and swamping with 5% cut-off points and 20% contamination for *n*=70 and 100

$\boldsymbol{DFBETAc_{IS}}$ STATISTIC

The $DFBETAc_{IS}$ statistic is used to detect outliers in wind direction data in this case. The cut-off point for this was constructed again and utilized the simulated program since the wind direction data was 129. Table 2 summarizes the 5% upper percentile cut-off points for wind direction data.

A part of simulation results are displayed in Appendix. The table last two columns show the number of outliers and the percentage of observations that exceed the cut-off point, indicating outlier possibilities. The observation numbers 38 and 111 are identified as outliers using the $|DFBETAc_{IS\ j,i}|$ statistic since the test values are more than 0.176 percent. Figure 5(a), 5(b), and 5(c) indicates that the equivalent value for observation number 38 is different from the others, but the line for observation 111 in Figure 5(d), 5(e), and 5(f) cuts at the parameter estimates of $\hat{C}_0, \hat{C}_1$ and $\hat{D}_1$.

TABLE 2. The cut-off point value for wind direction data

Parameter estimate	$\hat{A}_0$	$\hat{A}_1$	$\hat{B}_1$	$\hat{C}_0$	$\hat{C}_1$	$\hat{D}_1$
Cut-off point value	3.8318	3.1921	3.2617	5.1182	4.2121	5.1819

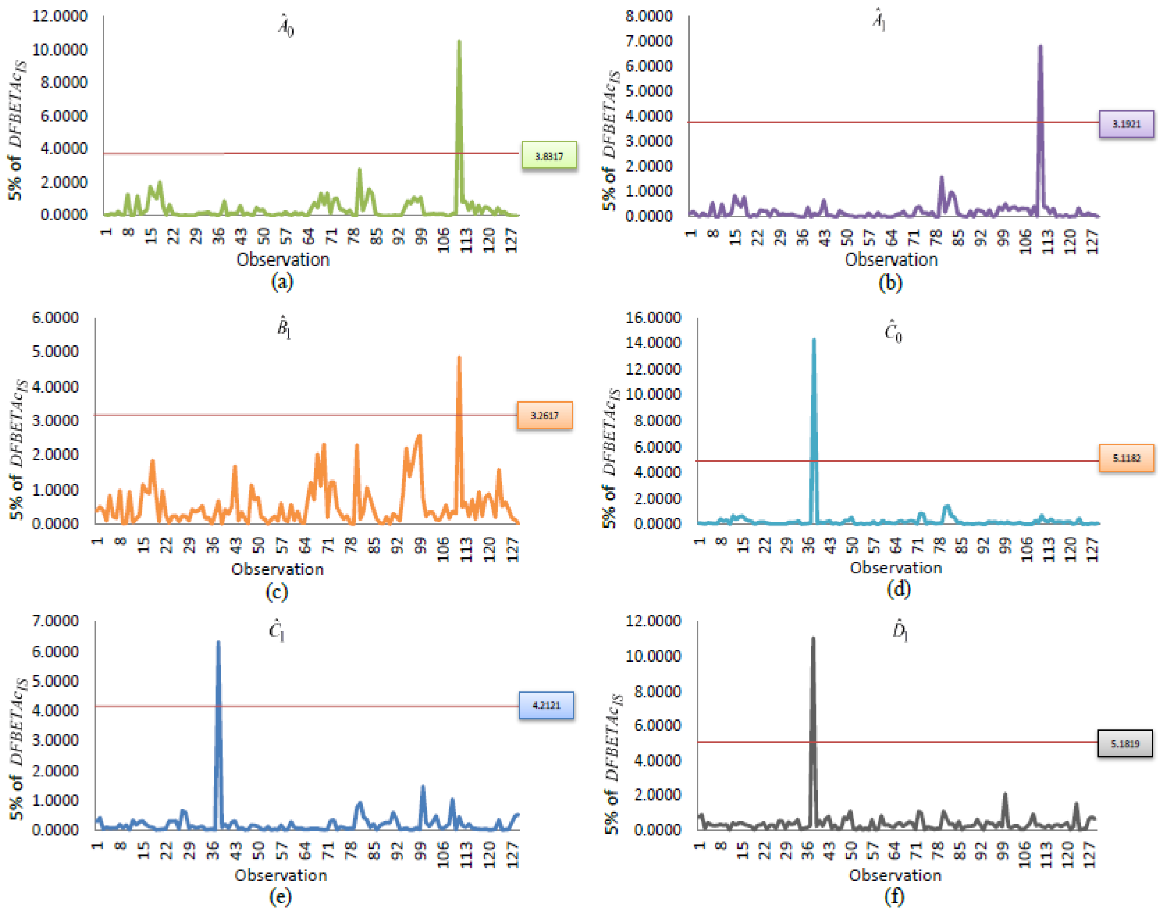


FIGURE 5. $DFBETAc_{IS}$ for parameter estimate of wind direction data with 5% cut-off points

Table 3 shows performance measures for the 5% cut-off points and the 10% and 20% contamination for wind direction data. From the table, two observations that seem to be outlier from the other observations, namely observations 38 and 111. In addition, no observations were identified in the data set. Therefore, the percentage of contamination unsuccessfully identify outliers based on swamping and masking methods in wind direction data set.

Then, the effect of deleting outliers on parameter estimations is shown in Table 4. The removal of observations 38 and 111 has a significant impact on certain of JSCRM's estimate parameters. For clean data, the standard error for parameter estimates $\hat{A}_0, \hat{A}_1$ and $\hat{B}_1$ does not very much, but the values of $\hat{C}_0, \hat{C}_1$ and $\hat{D}_1$ are less than for contaminated data. The concentration parameter, $\hat{\rho}$ on the other hand, increased from 0.9322 to 0.9474.

Figure 6 shows the Q-Q plots of the resulting residuals corresponding to the observational regression-like models after eliminating observations number 38 and 111 from the wind direction data set. The points on the plots ε_1 are now closer to the straight line than they were in Figure 6(b). Deleting observations 38 and 111 from the analysis allows the reduced data to be better fitted using the JSCRM, though it does point out that there is a point in Figure 6(b) that needs to be investigated further. This result show better than *CORATIO* statistic (Ibrahim et al. 2013) when only one outlier is detected (observation 38) in the wind direction data using the same model.

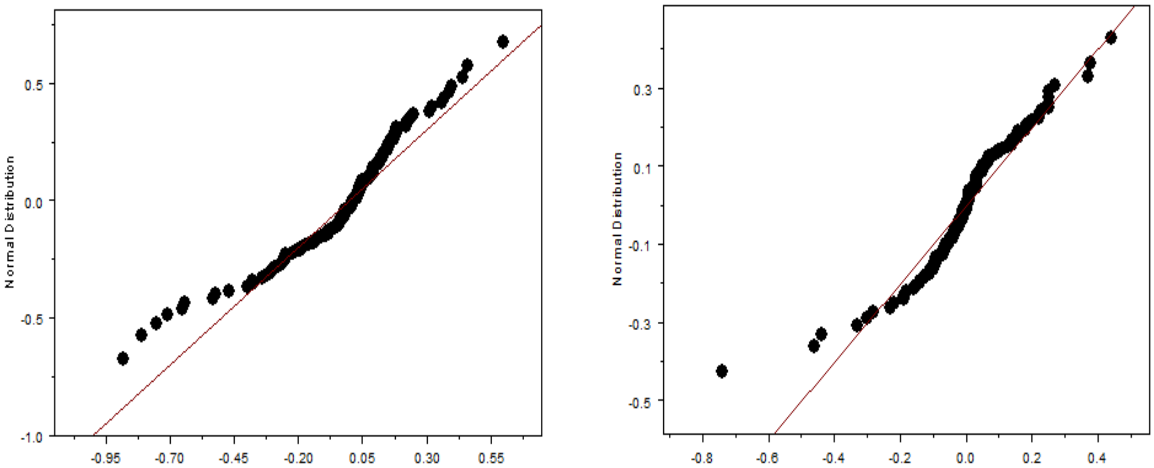
However, it is also supported that the *COVRATIO* statistic in Rambli et al. (2015) and Mokhtar et al. (2019), *D*, *M*, *A*, and *Chord* Statistics (Abuzaid 2010) show the same results when applied to wind direction data. Therefore, the exclusion of these two observations from the original data set improves the goodness-of-fit for the model.

TABLE 3. Performance measure for 5% cut-off points with 10% and 20% of contaminations for wind direction data

Percentage of contamination	Outlier observation in the data	Outlier observations being identified	Number of observations swamped	Number of observations masked
10	38 and 111	0	0	0
20		0	0	0

TABLE 4. $DFBETAc_{IS}$ statistic of wind direction data without observations number 38 and 111, $n = 117$

Parameter estimates	Contaminated data	Standard error	Clean data (case 38 and 111 deleted)	Standard error
$\hat{A}_0$	0.0674	0.0361	0.0633	0.0365
$\hat{A}_1$	0.7557	0.0598	0.7609	0.0602
$\hat{B}_1$	-0.0948	0.0323	-0.0974	0.0325
$\hat{C}_0$	-0.0470	0.0291	-0.0114	0.0196
$\hat{C}_1$	0.1049	0.0483	0.0620	0.0324
$\hat{D}_1$	0.9762	0.0261	0.9981	0.0175
$\hat{\sigma}_1$	0.3000	0.2849	0.2800	0.2853
$\hat{\sigma}_2$	0.3000	0.2300	0.1500	0.1533
$\hat{\rho}$	0.9322	-	0.9474	-



(a) ε_1 (b) ε_2

FIGURE 6. Q-Q plot for circular residuals without observations number 38 and 111

Conclusion

The aim of this study was to extend *DFBETAc* statistic in multiple circular case to identify multiple outliers in JSCRM. The $DFBETAc_{IS}$ statistic was developed to determine how much observation has influenced the JSCRM estimate of the regression coefficient. Simulation studies are used to determine the cut-off points and assess the power's performance for single and multi-outlier. The proposed method performance is evaluated using the proportion of detected outliers and the rate of masking and swamping. The results show that the proposed $DFBETAc_{IS}$ statistic is very successful in identifying genuine outliers for different sample sizes and with very low masking and swamping. Furthermore, the performance improves when the sample size gets larger and the residual dispersion small.

This study also found that all $DFBETAc_{IS}$statistics using JSCRM successfully detected all outliers in wind direction data with low rate of masking and no rate of swamping effect at 5% cut-off point. Moreover, the proposed $DFBETAc_{IS}$ statistics can detect outliers in real data with a high level of λ. Also, the proposed statistic is successful in detecting outliers in a large data set.

When the $DFBETAc_{IS}$ statistic is applied to wind direction data, observations 38 and 111 are identified as outliers, which similar observations detected as found in Abuzaid (2010), Mokhtar et al. (2019) and Rambli et al. (2015). The exclusion of these two observations from the original data set improves the fitted JSCRM. In conclusion, the proposed $DFBETAc_{IS}$ statistics using JSCRM is a practical and promising approach for detecting outliers in circular data.

Acknowledgments

The authors are most grateful to the Associate Editor and referees for their thorough reading and valuable suggestions, which led to an improvement of the article.

References

Abuzaid, A.H. 2020. Detection of outliers in univariate circular data by means of the outlier local factor (LOF). *Statistics in Transition New Series* 21(3): 39-51.

Abuzaid, A.H. 2010. Some problems of outliers in circular data. Doctoral dissertation, University of Malaya.

Abuzaid, A.H., Mohamed, I.B. & Hussin, A.G. 2009. A new test of discordancy in circular data. *Communications in Statistics-Simulation and Computation* 38: 682-691.

Abuzaid, A.H., Hussin, A.G., Rambli, A. & Mohamed, I. 2012. Statistics for a new test of discordance in circular data. *Communications in Statistics-Simulation and Computation* 41: 1882-1890.

Alkasadi, N.A., Ibrahim, S., Abuzaid, A. & Yusoff, M.I. 2019. Outlier detection in multiple circular regression model using DFFITC statistic. *Sains Malaysiana* 48(7): 1557-1563.

Alkasadi, N.A., Abuzaid, A.H., Ibrahim, S. & Yusoff, M.I. 2018. Outliers detection in multiple circular regression model via DFBETAc statistic. *International Journal of Applied Engineering Research* 13: 9083-9090.

Alkasadi, N.A., Ibrahim, S., Ramli, M.F. & Yusoff, M.I. 2016. A comparative study of outlier detection procedures in multiple circular regression. In *AIP Conference Proceedings* 1775: 030032.

Barnett, V. & Lewis, T. 1994. *Outliers in Statistical Data.* New York: John Wiley and Sons.

Belsley, D.A., Kuh, E. & Welsch, R.E. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley & Sons.

Binkley, S.A. 1990. *The Clockwork Sparrow: Time, Clocks, and Calendars in Biological Organisms*. New Jersey: Prentice Hall.

Chambers, R., Hentges, A. & Zhao, X. 2004. Robust automatic methods for outlier and error detection. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 167: 323-339.

Chatterjee, S. & Hadi, A.S. 1988. Impact of simultaneous omission of a variable and an observation on a linear regression equation. *Computational Statistics & Data Analysis* 6: 129-144.

Collett, D. 1980. Outlier in circular data. *Journal of the Royal Statistical Society Series C: Applied Statistics* 29(1): 50-57.

Cook, R.D. 1977. Detection of influential observation in linear regression. *Technometrics* 19:1 5-18.

Cousineau, D. & Chartier, S. 2010. Outliers detection and treatment: A review. *International Journal of Psychological Research* 3: 58-67.

Downs, T. 1974. Rotational angular correlation. In *Biorhythms and Human Reproduction*, edited by Ferin, M., Halberg, F. & van der Wiele, L. New York: Wiley. pp. 97-104.

Fisher, N.I. & Lee, A.J. 1992. Regression models for an angular response. *Biometrics* 48(3): 665-677.

Follmann, D.A. & Proschan, M.A. 1999. A simple permutation-type method for testing circular uniformity with correlated angular measurements. *Biometrics* 55(3): 782-791.

Gould, A.L. 1969. A regression technique for angular variates. *Biometrics* 25(4): 683-700.

Hrushesky, W.J.M. 1985. Circadian timing of cancer chemotherapy. *Science* 228: 73-75.

Hussin, A.G., Fieller, N.R.J. & Stillman, E.C. 2004. Linear regression model for circular variables with application to directional data. *Journal of Applied Science and Technology* 9(1 & 2): 1-6.

Ibrahim, S. 2013. Some outlier problems in a circular regression model. Doctoral dissertation, Fakulti Sains, Universiti Malaya.

Ibrahim, S., Rambli, A., Hussin, A.G. & Mohamed, I. 2013. Outlier detection in a circular regression model using *COVRATIO* statistic. *Communications in Statistics-Simulation and Computation* 42(10): 2272-2280.

Jammalamadaka, S. & Sarma, Y. 1993. Circular regression. *Statistical Science and Data Analysis* 34: 109-128.

Jha, J., Biswas, A. & Cheng, T.C. 2022. Trimmed estimator for circular–circular regression: Breakdown properties and an exact algorithm for computation. *Statistics* 56(2): 375-395.

Johnson, R.A. & Wehrly, T.E. 1978. Some angular-linear distributions and related regression models. *Journal of the American Statistical Association* 73: 602-606.

Jones, M.C. & Silverman, B.W. 1989. An orthogonal series density estimation approach to reconstructing positron emission tomography images. *Journal of Applied Statistics* 16: 177-191.

Lowrey, P.L., Shimomura, K., Antoch, M.P., Yamazaki, S., Zemenides, P.D., Ralph, M.R., Menaker, M. & Takahashi, J.S. 2000. Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. *Science* 288(5465): 483-492.

Lund, U. 1999. Least circular distance regression for directional data. *Journal of Applied Statistics* 26: 723-733.

Mackenzie, J.K. 1957. The estimation of an orientation relationship. *Acta Crystallographica* 10: 61-62.

Mardia, K. 1975. Statistical of directional data (with discussion). *Journal of the Royal Statistical Society* 37: 390.

Meilán-Vila, A., Crujeiras, R.M. & Francisco-Fernández, M. 2021. Nonparametric estimation of circular trend surfaces with application to wave directions. *Stochastic Environmental Research and Risk Assessment* 35(4): 923-939.

Mohamed, I.B., Rambli, A., Khaliddin, N. & Ibrahim, A.I.N. 2016. A new discordancy test in circular data using spacings theory. *Communications in Statistics-Simulation and Computation* 45: 2904-2916.

Mokhtar, N.A., Zubairi, Y.Z., Hussin, A.G. & Moslim, N.H. 2019. An outlier detection method for circular linear functional relationship model using *covratio* statistics. *Malaysian Journal of Science* 38(Special Issue 2): 46-54.

Moore-Ede, M.C., Sulzman, F.M. & Fuller, C.A. 1982. *The Clocks that Time Us: Physiology of the Circadian Timing System.* Massachusetts: Harvard University Press.

Rambli, A., Yunus, R.M., Mohamed, I. & Hussin, A.G. 2015. Outlier detection in a circular regression model. *Sains Malaysiana* 44(7): 1027-1032.

Rivest, L.P. 1997. A decentered predictor for circular-circular regression. *Biometrika* 84: 717-726.

Rousseeuw, P.J. & Leroy, A.M. 2005. *Robust Regression and Outlier Detection*. New York: John Wiley & Sons.

Shearman, L.P., Sriram, S., Weaver, D.R., Maywood, E.S., Chaves, I., Zheng, B., Kume, K., Lee, C.C., van der, G.T.J, Horst, Hastings, M.H. & Reppert, S.M. 2000. Interacting molecular loops in the mammalian circadian clock. *Science* 288(5468): 1013-1019.

Stephens, M.A. 1979. Vector correlation. *Biometrika* 66(1): 41-48.

Weir, I.S. & Green, P.J. 1994. Modelling data from single-photon emission computerized tomography. *Journal of Applied Statistics* 21: 313-337.

*Corresponding author; email: safwati@unimap.edu.my

APPENDIX

The value of the $|DFBETAc_{IS\,j,i}|$ statistic for wind direction data, n = 129 and the number of influenced parameters

Obs.	Parameter Estimate						Influenced Parameter	
	$\hat{A}_0$	$\hat{A}_1$	$\hat{B}_1$	$\hat{C}_0$	$\hat{C}_1$	$\hat{D}_1$	Number of outliers	Percentage detection
1	0.0567	0.1527	0.3903	0.1075	0.2893	0.7393	0	0
2	0.0411	0.2337	0.5024	0.0722	0.4109	0.8834	0	0
3	0.1245	0.0887	0.4029	0.0322	0.0230	0.1043	0	0
4	0.0339	0.0251	0.1115	0.1329	0.0981	0.4365	0	0
5	0.2752	0.1724	0.8447	0.0954	0.0597	0.2928	0	0
6	0.0572	0.0674	0.2352	0.0596	0.0703	0.2455	0	0
7	0.0531	0.0473	0.1896	0.0748	0.0666	0.2673	0	0
8	1.2789	0.5659	0.9878	0.4249	0.1880	0.3281	0	0
9	0.0160	0.0044	0.0193	0.2202	0.0599	0.2660	0	0
10	0.0510	0.0215	0.0425	0.3821	0.1606	0.3185	0	0
11	1.1892	0.5131	0.9549	0.0125	0.0054	0.0101	0	0
12	0.0976	0.0481	0.0610	0.6741	0.3320	0.4215	0	0
13	0.2048	0.0905	0.1587	0.3992	0.1763	0.3093	0	0
14	0.3949	0.1845	0.2768	0.5997	0.2803	0.4204	0	0
15	1.7575	0.8474	1.1547	0.6462	0.3115	0.4245	0	0
16	1.2399	0.5438	0.9711	0.4095	0.1796	0.3207	0	0
17	1.0324	0.4171	0.9058	0.2853	0.1153	0.2503	0	0
18	2.0360	0.7969	1.8544	0.2317	0.0907	0.2110	0	0
19	0.5710	0.0981	0.8218	0.0110	0.0019	0.0158	0	0
20	0.0351	0.0027	0.0692	0.1355	0.0103	0.2675	0	0
21	0.6606	0.1005	0.9798	0.1948	0.0296	0.2890	0	0
22	0.1186	0.0288	0.2747	0.1802	0.0438	0.4173	0	0
23	0.0864	0.2824	0.0478	0.0935	0.3055	0.0517	0	0
24	0.0190	0.2579	0.2399	0.0223	0.3030	0.2819	0	0
25	0.0222	0.2577	0.2466	0.0265	0.3083	0.2950	0	0
26	0.0322	0.0309	0.0864	0.0328	0.0315	0.0881	0	0
27	0.0093	0.3082	0.2571	0.0194	0.6456	0.5385	0	0
28	0.0150	0.2669	0.2380	0.0329	0.5867	0.5231	0	0
29	0.0267	0.0730	0.1115	0.0197	0.0539	0.0823	0	0
30	0.1690	0.1333	0.4273	0.1790	0.1412	0.4526	0	0
31	0.1578	0.0916	0.3683	0.1384	0.0803	0.3230	0	0
32	0.1815	0.0972	0.4159	0.1749	0.0936	0.4007	0	0
33	0.2368	0.1135	0.5297	0.2719	0.1303	0.6081	0	0
34	0.0667	0.0566	0.1723	0.0109	0.0093	0.0282	0	0
35	0.0883	0.0404	0.1957	0.0344	0.0157	0.0763	0	0
36	0.0037	0.0020	0.0085	0.0620	0.0338	0.1427	0	0
37	0.1763	0.0102	0.3179	0.1369	0.0079	0.2467	0	0

38	0.8847	0.3914	0.6833	14.3172	6.3349	11.0576	3	0.5
39	0.0025	0.0030	0.0073	0.0731	0.0858	0.2100	0	0
40	0.1605	0.1527	0.4295	0.2056	0.1957	0.5502	0	0
41	0.1279	0.0860	0.3097	0.0958	0.0644	0.2318	0	0
42	0.1772	0.2184	0.5182	0.2149	0.2649	0.6284	0	0
43	0.6009	0.6668	1.6924	0.2814	0.3122	0.7925	0	0
44	0.0321	0.0550	0.1071	0.0245	0.0420	0.0818	0	0
45	0.1511	0.0824	0.3476	0.1250	0.0682	0.2876	0	0
46	0.0723	0.0321	0.1593	0.0151	0.0067	0.0332	0	0
47	0.0102	0.0022	0.0200	0.0583	0.0124	0.1148	0	0
48	0.4971	0.2662	1.1389	0.3257	0.1744	0.7463	0	0
49	0.3067	0.1642	0.7027	0.2733	0.1463	0.6262	0	0
50	0.3958	0.0888	0.7843	0.5431	0.1219	1.0761	0	0
51	0.0807	0.0674	0.2075	0.0302	0.0253	0.0777	0	0
52	0.0639	0.0472	0.1586	0.0057	0.0042	0.0142	0	0
53	0.0029	0.0048	0.0095	0.0705	0.1178	0.2326	0	0
54	0.0594	0.0464	0.1497	0.0003	0.0002	0.0007	0	0
55	0.1055	0.0482	0.2338	0.0562	0.0257	0.1245	0	0
56	0.0546	0.0087	0.1045	0.0055	0.0009	0.0106	0	0
57	0.2525	0.1627	0.6045	0.2521	0.1625	0.6037	0	0
58	0.0647	0.0316	0.1454	0.0059	0.0029	0.0132	0	0
59	0.0326	0.0196	0.0768	0.0321	0.0193	0.0754	0	0
60	0.2180	0.1867	0.5647	0.3167	0.2711	0.8201	0	0
61	0.0455	0.0350	0.1142	0.1145	0.0881	0.2876	0	0
62	0.1793	0.0557	0.3712	0.1513	0.0470	0.3132	0	0
63	0.0238	0.0160	0.0577	0.0923	0.0620	0.2233	0	0
64	0.0204	0.0054	0.0411	0.0714	0.0188	0.1443	0	0
65	0.4179	0.0370	0.6824	0.2254	0.0200	0.3681	0	0
66	0.7798	0.1193	1.2115	0.2736	0.0419	0.4251	0	0
67	0.5082	0.1392	0.7076	0.1497	0.0410	0.2085	0	0
68	1.3334	0.2340	2.0333	0.2229	0.0391	0.3399	0	0
69	0.6702	0.0478	1.1084	0.2934	0.0209	0.4852	0	0
70	1.3622	0.0312	2.3304	0.2347	0.0054	0.4014	0	0
71	0.1640	0.0584	0.2083	0.0244	0.0087	0.0310	0	0
72	0.9858	0.3679	1.2257	0.8532	0.3184	1.0608	0	0
73	1.0410	0.4300	1.2253	0.8166	0.3373	0.9611	0	0
74	0.4173	0.1761	0.4846	0.0186	0.0079	0.0216	0	0
75	0.3344	0.1805	0.3076	0.2107	0.1137	0.1938	0	0
76	0.1870	0.1048	0.1618	0.1474	0.0827	0.1276	0	0
77	0.3351	0.2005	0.2505	0.0568	0.0340	0.0425	0	0
78	0.0301	0.0174	0.0244	0.0897	0.0519	0.0728	0	0
79	0.0038	0.0022	0.0031	0.0984	0.0568	0.0805	0	0
80	2.7892	1.6010	2.3066	1.3056	0.7494	1.0797	0	0

81	0.3239	0.2085	0.1631	1.4157	0.9112	0.7130	0	0
82	0.8271	0.5331	0.4085	0.6851	0.4416	0.3384	0	0
83	1.6165	1.0011	1.0671	0.5637	0.3491	0.3721	0	0
84	1.3014	0.8261	0.7465	0.1361	0.0864	0.0781	0	0
85	0.1512	0.1691	0.4271	0.1763	0.1972	0.4979	0	0
86	0.0559	0.0640	0.1591	0.0008	0.0010	0.0024	0	0
87	0.0077	0.0297	0.0391	0.0385	0.1486	0.1955	0	0
88	0.0025	0.0253	0.0249	0.0221	0.2246	0.2209	0	0
89	0.0144	0.2484	0.2226	0.0150	0.2583	0.2314	0	0
90	0.0008	0.0054	0.0059	0.0359	0.2429	0.2643	0	0
91	0.0370	0.3018	0.3123	0.0719	0.5866	0.6070	0	0
92	0.0068	0.2756	0.2266	0.0088	0.3566	0.2932	0	0
93	0.0238	0.2177	0.1098	0.0028	0.0258	0.0130	0	0
94	0.4403	0.0625	0.9289	0.2318	0.0329	0.4890	0	0
95	0.8897	0.2896	2.2069	0.1626	0.0529	0.4033	0	0
96	0.7410	0.0215	1.3894	0.2021	0.0059	0.3789	0	0
97	1.0970	0.0617	1.8576	0.0925	0.0052	0.1567	0	0
98	0.8610	0.4013	2.3748	0.1588	0.0740	0.4379	0	0
99	1.0906	0.2949	2.5851	0.0824	0.0223	0.1953	0	0
100	0.0725	0.5334	0.7577	0.2002	1.4737	2.0933	0	0
101	0.0859	0.2858	0.2403	0.0755	0.2513	0.2113	0	0
102	0.1098	0.3841	0.3521	0.0359	0.1255	0.1151	0	0
103	0.1230	0.4107	0.3476	0.0844	0.2820	0.2386	0	0
104	0.0934	0.2810	0.1443	0.1581	0.4758	0.2443	0	0
105	0.1137	0.3462	0.1263	0.0332	0.1011	0.0369	0	0
106	0.1017	0.3438	0.2979	0.0141	0.0475	0.0412	0	0
107	0.0185	0.3403	0.5545	0.0067	0.1224	0.1994	0	0
108	0.0014	0.1008	0.1806	0.0030	0.2162	0.3875	0	0
109	0.1271	0.4311	0.3759	0.3081	1.0454	0.9115	0	0
110	0.1720	0.0196	0.3207	0.1326	0.0151	0.2473	0	0
111	10.5281	6.8163	4.8585	0.6872	0.4450	0.3171	3	0.5
112	0.8281	0.4116	0.5063	0.2998	0.1490	0.1833	0	0
113	0.8633	0.3988	0.6186	0.1937	0.0895	0.1388	0	0
114	0.3444	0.1604	0.2429	0.4230	0.1970	0.2984	0	0
115	0.8286	0.3398	0.7135	0.1538	0.0631	0.1325	0	0
116	0.1177	0.0246	0.1595	0.2093	0.0438	0.2835	0	0
117	0.6177	0.0789	0.9498	0.2036	0.0260	0.3131	0	0
118	0.1638	0.0217	0.2500	0.1322	0.0175	0.2017	0	0
119	0.5348	0.0888	0.7766	0.2084	0.0346	0.3026	0	0
120	0.4690	0.0161	0.8846	0.2269	0.0078	0.4280	0	0
121	0.3010	0.0391	0.6277	0.0101	0.0013	0.0210	0	0
122	0.0820	0.0338	0.2175	0.0841	0.0347	0.2231	0	0

123	0.4866	0.3592	1.5979	0.4649	0.3432	1.5267	0	0
124	0.1756	0.1060	0.5313	0.0011	0.0007	0.0034	0	0
125	0.2564	0.0858	0.6407	0.0250	0.0084	0.0625	0	0
126	0.0644	0.1692	0.4356	0.0104	0.0273	0.0704	0	0
127	0.0211	0.0791	0.1852	0.0735	0.2759	0.6459	0	0
128	0.0054	0.0894	0.1443	0.0286	0.4709	0.7599	0	0
129	0.0043	0.0219	0.0275	0.1007	0.5179	0.6486	0	0