Journal of Multivariate Analysis 119 (2013) 200–208

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Inverse circular–circular regression

Ashis SenGupta [a], Sungsu Kim [a,*], Barry C. Arnold [b]

[a] Indian Statistical Institute, Kolkata, India
[b] University of California, Riverside, USA

ARTICLE INFO

Article history:
Received 10 February 2012
Available online 1 May 2013

AMS subject classifications:
62H11
62J99

Keywords:
Asymmetric circular distribution
Calibration
Circular regression

ABSTRACT

The problem of determining the values of the independent variable given a value of the dependent variable is commonly referred to as the inverse regression problem. This problem is also encountered in real life with circular data and we refer to it in that context as the inverse circular regression problem. For such a problem, we develop distance-based methods, and parametric methods, where we use the von Mises (vM) error distribution and the asymmetric generalized von Mises (AGvM) error distribution. We then present a goodness of fit comparison among distance-based and parametric methods, utilizing a new criterion called the relative circular prediction bias (RCPB) criterion, with real and simulated examples. Real data applications are given from the biological and environmental sciences.

1. Introduction

Inverse regression refers to (inversely) predicting the corresponding value of an independent variable when one only observes the value(s) of the corresponding dependent variable, using a model that has already been established for the dependence between the two variables. In all of our methods, therefore, we use a model established after having observed a number (we used n for this number) of paired data points before observing a new value (or new values) of the dependent variable. Inverse (linear) regression is typically applied in calibration settings; see [2,13,18]. For example, suppose a weighing machine is calibrated using fixed known weights. Then, later, by reading the scale, one tries to predict the unknown true weight. This corresponds to usual practice in calibration. In this paper, we illustrate our new methods in the inverse circular regression problem using two well-known bivariate circular data sets from the biological and environmental sciences, where the new inverse regression methods can be useful.

Although a number of papers on direct and inverse regression for linear variables and on direct circular regression are available, no work has previously been done on the inverse regression problem for circular variables, the inverse circular regression problem. In view of the fact that circular variables arise in many areas of investigation, a need for inverse circular regression is apparent. For inverse regression, reference can be made to [2,13,18]. For circular regression, key references are [4,6–9,16,1].

First, we present the model that will be used in this paper for circular inverse regression problems. Consider the following mapping:

$$\tan\left(\frac{\Theta-\mu_\Theta}{2}\right)=a+b\tan\left(\frac{\Phi-\mu_\Phi}{2}\right), \tag{1.1}$$

* Corresponding author.
E-mail address: dr.sungsu@gmail.com (S. Kim).

0047-259X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmva.2013.04.011

which has the unique solution (inverse mapping)

$$\Theta = \mu_\Theta + 2\arctan\left\{a + b\tan\left(\frac{\Phi - \mu_\Phi}{2}\right)\right\}. \tag{1.2}$$

Note that the arctangent has double solutions in $[0, 2/pi)$, but by restricting to half-angles $\left(\frac{\Theta}{2} \text{ or } \frac{\Phi}{2}\right)$, we have made the solution unique. This defines a one-to-one mapping between Θ and Φ provided that b is not equal to zero. Using the mapping in (1.2), we present a new circular–circular regression model, where Θ is treated as the dependent circular variable and Φ is treated as the fixed (controlled) independent variable. Specifically we assume that

$$\Theta_i = \mu_\theta + 2\arctan\left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1.3}$$

where the ε_is follow a circular distribution with 0 mean direction, and where a and b can be viewed as intercept and slope parameters appearing in (1.1), and μ_Θ and μ_ϕ are the mean directions of Θ and ϕ. Our model (1.3) is a generalization of the model introduced by Downs and Mardia [4], which excludes a from (1.3). Using the link function without the intercept parameter, it is assumed that the conditional mean direction value of the dependent circular variable is its unconditional mean direction μ_θ, i.e. $\mu_{\theta_i|\phi_i} = \mu_\theta$, when the value of the independent circular variable is its unconditional mean direction μ_ϕ. However, this is not always obviously appropriate and, therefore, need not be generally assumed. The interpretation of a in the new link function is the rotation from μ_θ, by the amount of $2\arctan(a)$, when conditioning on $\Phi_i = \mu_\phi$.

Two inverse (linear) regression methods, called the classical and the inverse methods, have been present in the literature for more than four centuries. The calibration problem has been under debate from as early as the paper by Eisenhart [5], who advocates use of the classical estimator as the only correct estimator for the corresponding value of the independent variable. On the other hand, Tallis [17] proposes a different approach to a problem of calibration in the bivariate setting when additional information in connection with the distribution of Y is available, such that the bivariate distribution is determined. Then, he shows that it estimates X for an observed y with the minimum MSE, hence providing a more suitable estimator than the classical estimator. On the other hand, inference is relatively straightforward for the classical estimator if the errors are assumed to follow a normal distribution. Then, the parameter estimators will also be normally distributed conditionally on the values of the independent variables regardless of the sample size, and they are also the maximum likelihood estimators. Our readers are referred to [15] for a comprehensive review on debates between the classical and the inverse methods. Eisenhart [5] argues that the inverse estimator violating the original assumption should not be used or preferred to the classical estimator; Kim [11,12] proposes that the inverse estimator is also justifiable when viewed in terms of minimizing prediction errors in X, i.e. the squared horizontal distances, just as the classical estimator is obtained by minimizing prediction errors in Y, i.e. the squared vertical distances. These estimators are called the distance-based methods. The difference between the inverse estimator examined in [11,12] and the one in the literature is that in [11,12], X is not treated as the dependent variable in (1.3), and Y in the right side of the equation is still treated as the variable observed with errors. In this way, we use the estimated regression of Y on X, not the estimated regression of X on Y, in both the classical and the inverse approaches, hence not violating the original postulate of y being measurement with error and x being controlled measurement.

Two of our three distance-based methods are derived as analogs to these linear methods in view of [11,12]. The other new distance-based method is influenced by methods known as the orthogonal distance or the total distance in linear regression.

Next, we introduce a new asymmetric circular distribution called the asymmetric generalized von Mises (AGvM) distribution in Section 3.2. Some of its properties are found in Section 3.3. Our new parametric methods are applied to models involving the von Mises (vM) and the AGvM error distributions. In these methods, maximum likelihood estimation is used in determining appropriate values for prediction of the independent variable. A new comparison criterion, called the relative circular prediction bias (RCPB), is introduced in Section 4.1. It is motivated by consideration of the problems associated with the use of mean square error in the inverse linear regression setting [18]. The RCPB criterion is used to compare all distance-based and parametric methods proposed in this paper. The results of these comparisons may be found in the conclusion section. All the numerical optimizations were carried out by using the "optim" function from R, where we used the "Nelder–Mead" method, which is a version of the Newton–Raphson method. In the next section, we first present the two new methods analogous to the existing two linear inverse regression methods, and then we present the other new distance-based inverse circular regression method.

Throughout this paper all circular random variables are assumed to take on values in the interval $(-\pi, \pi]$, and as a consequence, the corresponding densities are positive in this interval and zero elsewhere. Density functions in this paper will not include explicit reference to the support of the corresponding densities.

2. Three new distance-based inverse prediction methods

The circular distance between two angles α and β is given by $1 - \cos(\alpha - \beta)$. The usual measure of distance between α and β yields two distances, namely, $\alpha - \beta$ and $2\pi - (\alpha - \beta)$, depending on the sense of direction for a fixed origin. The more appropriate unique measure of distance between two angles, $1 - \cos(\alpha - \beta)$, is required for a circular statistical analysis [9]. In the following, we suppose that $\Theta_1, \ldots, \Theta_n$ are n independent circular variables observed for n fixed values $\phi_1, \ldots, \phi_n$ of the controlled variable Φ.

2.1. The distance-based method analogous to the classical method

We estimate the parameters in the model (1.3) by minimizing the sum of circular distances between observed values $\theta_1, \dots, \theta_n$ and the associated predicted values $\hat{\theta}_1, \dots, \hat{\theta}_n$. That is, the objective function to be minimized is given by

$$Q(a, b, \mu_\Theta, \mu_\phi) = \sum_{i=1}^{n} \left[1 - \cos\left(\theta_i - \mu_\theta - 2\arctan\left\{ a + b\tan\left(\frac{\phi_i - \mu_\phi}{2} \right) \right\} \right) \right].$$

Then, the estimates $\hat{a}$, $\hat{b}$, $\hat{\mu}_\theta$ and $\hat{\mu}_\phi$ obtained are used as shown below, in order to predict the corresponding unknown value of Φ after observing one or more values of Θ:

$$\hat{\phi} = \hat{\mu}_\phi + 2\arctan\left\{ \frac{\tan\left(\frac{\theta - \hat{\mu}_\theta}{2}\right) - \hat{a}}{\hat{b}} \right\}.$$

Here and in the rest of this paper, in cases in which more than one value of Θ is observed, we replace θ by $\bar{\theta}$, which represents the mean direction of the θs. In the example section and the following discussion of the results, this method will be called the "classical" method.

2.2. The distance-based method analogous to the inverse method

First of all, we rewrite (1.2) as shown below, in which it was solved for ϕ:

$$\phi = \mu_\phi + 2\arctan\left\{ c + d\tan\left(\frac{\Theta - \mu_\theta}{2} \right) \right\}, \tag{2.1}$$

where c and d are reparameterizations of $-\frac{a}{b}$ and $\frac{1}{b}$, respectively. We estimate the parameters in (2.1) by minimizing the sum of circular distances between the original values $\phi_1, \dots, \phi_n$ and the associated predicted values $\hat{\phi}_1, \dots, \hat{\phi}_n$. This means that the objective function to be minimized is given by

$$Q(c, d, \mu_\Theta, \mu_\phi) = \sum_{i=1}^{n} \left[1 - \cos\left(\phi_i - \mu_\phi - 2\arctan\left\{ c + d\tan\left(\frac{\Theta_i - \mu_\theta}{2} \right) \right\} \right) \right].$$

Then, the estimates obtained, i.e. $\hat{c}$, $\hat{d}$, $\hat{\mu}_\theta$ and $\hat{\mu}_\phi$, are used as shown below, to predict the unknown corresponding value of Φ after observing one or more values of Θ:

$$\hat{\phi} = \hat{\mu}_\phi + 2\arctan\left\{ \hat{c} + \hat{d}\tan\left(\frac{\theta - \hat{\mu}_\theta}{2} \right) \right\}.$$

In the example section and the following discussion of the results, this method will be called the "inverse" method.

2.3. The distance-based method enhanced by a new total regression method

We estimate the parameters in the model (1.3) by minimizing the sum of circular distances between observed values $\theta_1, \dots, \theta_n$ and the associated predicted values $\hat{\theta}_1, \dots, \hat{\theta}_n$, together with the sum of circular distances between the original $\phi_1, \dots, \phi_n$ and the associated predicted $\hat{\phi}_1, \dots, \hat{\phi}_n$. The objective function to be minimized is shown below:

$$\begin{aligned} Q(a, b, \mu_\Theta, \mu_\phi) = {} & \sum_{i=1}^{n} \left[1 - \cos\left(\theta_i - \mu_\theta - 2\arctan\left\{ a + b\tan\left(\frac{\phi_i - \mu_\phi}{2} \right) \right\} \right) \right] \\ & + \sum_{i=1}^{n} \left[1 - \cos\left(\phi_i - \mu_\phi - 2\arctan\left\{ \frac{\tan\left(\frac{\theta_i - \mu_\theta}{2}\right) - a}{b} \right\} \right) \right]. \end{aligned}$$

Then, the estimates obtained, i.e. $\hat{a}$, $\hat{b}$, $\hat{\mu}_\theta$ and $\hat{\mu}_\phi$, are used as shown below, to predict the unknown corresponding value of Φ after observing one or more values of Θ:

$$\hat{\phi} = \hat{\mu}_\phi + 2\arctan\left\{ \frac{\tan\left(\frac{\theta - \hat{\mu}_\theta}{2}\right) - \hat{a}}{\hat{b}} \right\}.$$

In the example section and the following discussion of the results, this method will be called the "3rd distance-based" method.

2.4. Asymptotic properties of LCD estimators of linear parameters

The estimation methods used in our distance-based predictors are called the least circular distance estimation (LCDE) methods. Asymptotic normality of the LCDE for linear parameters is established in the following theorem.

Theorem 2.1 (*Distribution of LCD Estimators of Linear Parameters*). *Make the following assumptions:*

(i) *The model is given, letting ζ denote the vector of linear parameters, i.e. $(a, b)'$, by*

$$\theta_i = m_i(\zeta') + \varepsilon_i = m_i(a, b) + \varepsilon_i = \mu_\theta + 2\arctan\left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\} + \varepsilon_i,$$

where $m_i(\cdot)$ is a continuous function of θ and ζ.

(ii) *The conditional mean directions of circular errors are zeros, and the circular errors can be heteroscedastic and correlated over i.*

(iii) *The mean function $m_i(\zeta)$ satisfies $m_i(\zeta_1) = m_i(\zeta_2)$ iff $\zeta_1 = \zeta_2$.*

(iv) *The matrix*

$$A_0 = \operatorname{plim}\frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i}{\partial\zeta}\frac{\partial m_i}{\partial\zeta'}\cos(\theta_i - m_i)|_{\zeta+} = \lim\frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i}{\partial\zeta}\frac{\partial m_i}{\partial\zeta'}E\{\cos(\theta_i - m_i)\}|_{\zeta+}$$
$$= \frac{\partial m_i}{\partial\zeta}\frac{\partial m_i}{\partial\zeta'}E\{\cos(\theta_i - m_i)\}|_{\zeta+}$$

exists and is finite and nonsingular.

(v) $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial m_i}{\partial\zeta}\sin(\theta_i - m_i)|_{\zeta_0} \xrightarrow{d} N(0, B_0)$, *where*

$$B_0 = \frac{\partial m_i}{\partial\zeta}\frac{\partial m_i}{\partial\zeta'}\operatorname{var}\{\sin(\theta_i - m_i)\}|_{\zeta_0} = \frac{\partial m_i}{\partial\zeta}\frac{\partial m_i}{\partial\zeta'}E\{\sin^2(\theta_i - m_i)\}|_{\zeta_0}.$$

(vi) *ζ_0 is an interior point in the set of its possible values.*

Then the LCD estimator $\hat{\zeta}$ of linear parameters, defined to be a root of the first-order conditions $\frac{\partial}{\partial\zeta}\sum_{i=1}^{n}\cos(\theta_i - m_i) = 0$ is consistent for ζ and

$$\hat{\zeta} \overset{a}{\sim} N(\zeta_0, n^{-1}A_0^{-1}B_0A_0^{-1}).$$

Proof. A proof of the theorem may be found in the Appendix. □

Remark 1. The assumption that the conditional means of errors are zeros in the least square estimation is replaced by the assumption that the conditional mean directions of circular errors are zeros.

Remark 2. The conditions (i)–(iii) imply that the regression function is correctly specified.

Remark 3. The probability limits in (iv) and (v) become regular limits if Φ is nonstochastic.

3. New likelihood-based inverse prediction methods

We present two new likelihood-based inverse circular regression methods, one which has the von Mises error distribution and one having the asymmetric generalized von Mises error distribution. Since the vM and AGvM distributions meet the usual regularity conditions, the MLEs of the parameters are asymptotically normally distributed. In the following, we suppose that $\Theta_1, \ldots, \Theta_n$ are n independent circular variables observed for n fixed $\phi_1, \ldots, \phi_n$.

3.1. The likelihood-based method using the von Mises (vM) error distribution

We establish the regression using (1.3) with the vM error distribution as

$$\Theta_i = \mu_\theta + 2\arctan\left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\} + \varepsilon_i, \quad i = 1, \ldots, n,$$

where the ε_is follow vM(0, 1). In this model, there are four parameters, a, b, μ_θ, and μ_ϕ. We use the MLE method to fit the model. The log likelihood function is

$$L(a, b, \mu_\theta, \mu_\phi) = \sum_{i=1}^{n}\cos\left[\theta_i - \mu_\theta - 2\arctan\left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\}\right]$$
$$- n\log\left\{\int_{-\pi}^{\pi}\exp\left(\cos\left[\theta - \mu_\theta - 2\arctan\left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\}\right]\right)d\theta\right\}.$$

3.2. The asymmetric generalized von Mises (AGvM) distribution

In this section, we present a new asymmetric circular distribution called the asymmetric generalized von Mises (AGvM) distribution. The distribution can be used to model asymmetric or bimodal data, and furthermore, it can still be asymmetric when the location parameters are zero, which is not the case in the generalized von Mises (GvM) distribution [19].

Definition 1. Θ is an AGvM random variable if its density has the following form:

$$f_{\Theta}(\theta) = \frac{\exp\left[\kappa_1 \cos(\theta-\mu) + \kappa_2 \sin 2\{\theta - (\mu-\delta)\}\right]}{\int_{-\pi}^{\pi} \exp\left[\kappa_1 \cos(\theta-\mu) + \kappa_2 \sin 2\{\theta - (\mu-\delta)\}\right] d\theta}, \tag{3.1}$$

where $\mu \in [0.2\pi)$ is a location parameter, $\delta \in [0, 2\pi)$, $\kappa_1 \in R^1$ and $\kappa_2 \in R^1$ are shape parameters.

Though δ can be different from 0 in the AGvM, we restrict our discussion of the AGvM distribution to the case in which $\delta = 0$. The mean direction of the AGvM distribution is given by $\mu = \arctan\left\{\frac{b_1}{a_1}\right\}$, where a_1 and b_1 are the first Fourier cosine and sine coefficients of (3.1), namely, $E(\cos\theta)$ and $E(\sin\theta)$, respectively. The parameter μ qualifies as being a location parameter because shifting the density, which has $\mu = 0$, i.e. AGvM$(0, \delta, \kappa_1, \kappa_2)$, by μ leads to an AGvM$(\mu, \delta, \kappa_1, \kappa_2)$ density. The restricted model, in which $\mu = 0$ and $\delta = 0$, is particularly useful in the next section when we discuss regression with AGvM errors. However, the mean direction of the AGvM distribution is non-zero for the choice of $\mu = \delta = 0$. Hence, a bias correction is needed in order to achieve a zero conditional mean direction of errors. The density is asymmetric; however, it reduces to being symmetric when $\kappa_2 = 0$. One way to see that it is asymmetric is that it becomes an asymmetric normal distribution when $\cos\theta$ and $\sin\theta$ are approximated for small θ using $1 - \frac{\theta^2}{2}$ and θ, respectively. Then, assuming $\delta = 0$ for simplicity, we get

$$\begin{aligned}
\kappa_1 \cos(\theta-\mu) + \kappa_2 \sin 2(\theta-\mu) &= \kappa_1 \cos(\theta-\mu) + 2\kappa_2 \cos(\theta-\mu)\sin(\theta-\mu) \\
&= \cos(\theta-\mu)\{\kappa_1 + 2\kappa_2 \sin(\theta-\mu)\} \simeq \left\{1 - \frac{(\theta-\mu)^2}{2}\right\}\{\kappa_1 + 2\kappa_2(\theta-\mu)\} \\
&= -\frac{\kappa_1}{2}\left\{\theta - \left(\mu + 2\frac{\kappa_2}{\kappa_1}\right)\right\}^2 - \kappa_2(\theta-\mu)^3 + 2\kappa_2\left(\frac{\kappa_2}{\kappa_1} - \frac{3}{4}\mu\right).
\end{aligned} \tag{3.2}$$

Here, the term $-\kappa_2(\theta-\mu)^3$ is shown to be responsible for the asymmetry in (3.2).

All of our computations in the example section are performed using the R function called 'optim', where the 'Nelder–Mead' optimization technique is employed. Although the normalizing constant can be written in an infinite series form [11], the numerical calculation using the R function had no difficulty in producing the MLE estimates of the parameters. For more detailed study on the AGvM model, our readers can refer to the previous work of one of the authors found in [11].

3.3. The likelihood-based method using the AGvM circular error distribution

We establish the regression using (1.3) with the AGvM error distribution as

$$\Theta_i = \mu_\theta + 2\arctan\left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\} + \varepsilon_i, \quad i = 1, \ldots, n,$$

where the ε_is are i.i.d. AGvM distributed variables with their location parameters equal to zero, i.e. $\mu = \delta = 0$. Having AGvM errors, our model estimates the conditional mean direction of θ given $\Phi = \phi$ with an added bias, which is equal to the mean direction of the errors. Therefore, in order to have zero mean direction errors, we would shift the AGvM error distribution by an amount equal to its mean direction, i.e. we would use $\varepsilon' = \varepsilon - \mu_\varepsilon$ rather than ε in our regression, where μ_ε is the mean direction of ε, which was estimated to be 0.4383, by using the rejection method of simulation, as shown in the formula below:

$$\hat{\mu_\varepsilon} = \arctan\left(\frac{\sum_{i=1}^{m} \sin\varepsilon_j}{\sum_{i=1}^{m} \cos\varepsilon_j}\right),$$

where m denotes the number of simulated values from AGvM(0, 0, 1, 1), where we used $n = 2733$. Uniqueness of arctangent mapping modulo 2π follows as illustrated in [9, p. 13]. Our modified model becomes

$$\begin{aligned}
\Theta &= \mu(\phi) + \mu_\varepsilon + \varepsilon - \mu_\varepsilon = \mu_\theta + \mu_\varepsilon + 2\arctan\left\{a + b\tan\frac{1}{2}(\phi - \mu_\phi)\right\} + \varepsilon' \\
&= \nu_\theta + 2\arctan\left\{a + b\tan\frac{1}{2}(\phi - \mu_\phi)\right\} + \varepsilon',
\end{aligned}$$

Table 1
Symmetric data simulation.

	Inverse circular predictors				
	Classic	Inverse	3rd distance-based	vM errors	AGvM errors
OERCPB	−0.5417	0.6146	−0.7356	−0.3970	−0.0734

Table 2
Asymmetric data simulation.

	Inverse circular predictors				
	Classic	Inverse	3rd distance-based	vM errors	AGvM errors
OERCPB	0.3595	1.7490	−0.3312	0.3924	0.0961

where $\nu_\theta = \mu_\theta + \mu_\varepsilon$, and ε' has an AGvM(0, 0, 1, 1) distribution. The estimate $\hat{\mu}_\theta$ of μ_θ will then be obtained by first estimating ν_θ by $\hat{\nu}_\theta$ and then setting $\hat{\mu}_\theta = \hat{\nu}_\theta - \mu_\varepsilon$. In this model, the number of parameters is 4, appearing as $a, b, \nu_\theta, \mu_\phi$. We use the MLE method to fit the model. The log-likelihood function is

$$\begin{aligned} L(a, b, \nu_\theta, \mu_\phi) = {} & \sum_{i=1}^{n} \cos\left[\theta_i - \nu_\theta - 2\arctan\left\{a + b\tan\frac{1}{2}(\phi - \mu_\phi)\right\}\right] \\ & + \sum_{i=1}^{n} \sin 2\left[\theta_i - \nu_\theta - 2\arctan\left\{a + b\tan\frac{1}{2}(\phi - \mu_\phi)\right\}\right] \\ & - \sum_{i=1}^{n} \log\int_0^{2\pi} \exp\left(\cos\left[\theta_i - \nu_\theta - 2\arctan\left\{a + b\tan\frac{1}{2}(\phi - \mu_\phi)\right\}\right]\right) \\ & \times \exp\left(\sin 2\left[\theta_i - \nu_\theta - 2\arctan\left\{a + b\tan\frac{1}{2}(\phi - \mu_\phi)\right\}\right]\right) d\theta. \end{aligned}$$

4. Examples

4.1. Comparison of the performances of the proposed prediction methods

In this section, we compare all five predictors presented in this paper using the relative circular prediction bias (RCPB) criterion. Using the RCPB criterion, we need to average estimates of prediction errors for each predictor, and use the average, called the overall estimated circular prediction bias (OERCPB), as a criterion for comparing different predictors. The OERCPB when using a simulation is given by

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{\frac{1}{N}\sum_{j=1}^{N}\sin(\phi_i - \hat{\phi}_{ij})}{\sin\phi_i}\right\},$$

where $\hat{\phi}_{ij}$ is the estimate of ϕ_i based on the jth simulated group of data, using the cross-validation method, and N and n are the number of simulations and the size of the sample, respectively.

4.2. Simulations

We present two simulation examples, one for symmetric data and the other for asymmetric data, where in both simulations, N and n are 500 and 40, respectively. For each of 500 samples, we simulated 40 observations from vM(1.1, 2) for a symmetric data set, and 40 observations from AGvM(1, 2, 1.3, 2.1) for an asymmetric data set, conditionally on 40 randomly generated independent angles from Uniform$(0, 2\pi]$. The results are shown in Tables 1 and 2. The discussion of the results is found in the next section.

4.3. Real data sets

In this subsection, we used two real data sets, which are the noisy scrub bird nest data [6] and the wind direction data [10]. The bird's nest data contain two circular variables, namely, 56 pairs of the orientation of each bird's nest and the orientation

Table 3
Bird's nest data.

	Inverse circular predictors				
	Classic	Inverse	3rd distance-based	vM errors	AGvM errors
OERCPB	−0.5611	−0.8052	−0.1469	0.9854	0.8362

Table 4
Wind direction data.

	Inverse circular predictors				
	Classic	Inverse	3rd distance-based	vM errors	AGvM errors
OERCPB	−0.9706	−0.7256	−0.7788	−2.3622	−0.4710

of the nearby creek flow. Our aim is to inversely predict the orientation of creek flow after observing an orientation of a bird nest. On the other hand, the wind direction data contain 21 pairs of two wind directions, at 6 am and 12 pm in a Milwaukee observatory. Suppose one day the wind direction at 6 am was not recorded; our aim is to inversely predict the 6 am wind direction of the same day after observing that of the wind at 12 pm. Using the formula of OERCPB, we only need to have $N = 1$, i.e. it is given by

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{\sin(\phi_i - \hat{\phi}_i)}{\sin\phi_i}\right\},$$

where all the other cases are the same as in the simulation formula. The results are shown in Tables 3 and 4. The discussion of the results is found in the following section.

5. Conclusion and recommendation

From the simulation in the last section, we conclude that the likelihood-based predictor using the AGvM error performs better than the other four predictors for asymmetric data and even for symmetric data. Using the real data examples, the 3rd distance-based predictor performed better that the other four predictors for the bird's nest data [6], which are almost symmetric, while the likelihood-based predictor using the AGvM errors performed better than the other four predictors for the wind direction data [6], which are asymmetric. Overall, on the basis of the RCPB criterion, the simulations indicate that the likelihood-based predictor using the AGvM errors performs better than the other four predictors for asymmetric data, while either the 3rd distance-based predictor or the likelihood-based predictor using the AGvM errors performs better than the other three predictors for symmetric data.

We emphasize that the likelihood method using the AGvM errors has a particular merit in that it can be flexible enough to accommodate errors in not only asymmetric but also bimodal shapes. From the example section, it was also evident that the likelihood method using the AGvM errors performs better than the new distance-based methods for asymmetric data sets.

Acknowledgments

We thank the Editor for kindly extending the time for resubmission and the referee for valuable and constructive comments and criticisms, which helped us to modify/revise our paper to achieve a better presentation.

Appendix

Proof of Theorem 1. The limiting distribution of $\zeta = (a, b)$ is obtained using an exact first-order Taylor series expansion of the first-order condition, for some ζ^+ between $\hat{\zeta}$ and ζ_0 [3]:

$$\begin{aligned}\frac{\partial Q_n(\zeta)}{\zeta} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i(\zeta)}{\partial\zeta}\sin\{\theta_i - m_i(\zeta)\}\\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i(\zeta)}{\partial\zeta}\sin\{\theta_i - m_i(\zeta)\}|\zeta_0 - \frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i(\zeta)}{\partial\zeta}\frac{\partial m_i(\zeta)}{\partial\zeta'}\cos\{\theta_i - m_i(\zeta)\}|_{\zeta^+}(\hat{\zeta} - \zeta_0) = 0.\end{aligned}$$

$$\sqrt{n}(\hat{\zeta} - \zeta_0) = \left\{\frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i}{\partial\zeta}\frac{\partial m_i}{\partial\zeta'}\cos(\theta_i - m_i)|_{\zeta_+}\right\}^{-1} \times \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial m_i}{\partial\zeta}\sin(\theta_i - m_i)|_{\zeta_0}.$$

We apply the multivariate CLT for independent random vectors [14] in the following, to obtain an asymptotic multivariate normality of $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0}$, whose proof is shown in the following section. The multivariate CLT, in Theorem 2 of [14], follows later. Then,

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0} \xrightarrow{d} N(0, B_0), \quad \text{where} \\
&B_0 = \frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}\operatorname{var}(\sin(\theta_i - m_i))|_{\zeta_0} = \frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}E(\sin^2(\theta_i - m_i))|_{\zeta_0}, \\
&\text{and} \quad \operatorname{plim}\frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}\cos(\theta_i - m_i)|_{\zeta_+} = \lim\frac{1}{n}\sum_{i=1}^{n}\frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}E\{\cos(\theta_i - m_i)\}|_{\zeta_+} \\
&\qquad = \frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}E\{\cos(\theta_i - m_i)\}|_{\zeta_+} = A_0.
\end{aligned}$$

Now, using Slutsky's theorem (or the Product Limit Normal Rule) we get

$$\sqrt{n}(\hat{\zeta} - \zeta_0) \xrightarrow{d} N(0, A_0^{-1}B_0A_0^{-1}).$$

Then, the asymptotic distributions are given by

$$\hat{\zeta} \overset{a}{\sim} N(\zeta_0, n^{-1}A_0^{-1}B_0A_0^{-1}). \quad \square$$

In the following, we prove that the $\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0}$s have an asymptotic multivariate normal distribution using the aforementioned Theorem 2. A proof of Theorem 2 is found on p. 285 of [14].

Considering independent 2×1 vectors $\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0}$, we have that

$$P\left\{\left|\frac{\partial m_i}{\partial a}\sin(\theta_i - m_i)|_{\zeta_0}\right| \le C, \left|\frac{\partial m_i}{\partial b}\sin(\theta_i - m_i)|_{\zeta_0}\right| \le C\right\} = 1, \quad \forall i \in \{1, \dots, n\}, \tag{A.1}$$

which is established below. For $\phi_i \neq \pi, 3\pi$,

$$\begin{aligned}
0 < \left.\frac{\partial m_i}{\partial a}\right|_{\zeta_0} &= \frac{2}{1 + \left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\}^2} < 2 \\
0 \le \left.\frac{\partial m_i}{\partial b}\right|_{\zeta_0} &= \frac{2\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)}{1 + \left\{a + b\tan\left(\frac{\phi_i - \mu_\phi}{2}\right)\right\}^2} < \infty.
\end{aligned}$$

This means that there exists a finite real number $2 \le C < \infty$ such that the condition (A.1) is satisfied. Next, we have $E\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0} = \mathbf{0}$, since $E\sin(\theta - m) = 0$ for all circular variables with m denoting their mean direction, and $\operatorname{cov}\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0} = \frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}E\sin^2(\theta_i - m_i)|_{\zeta_0} = \kappa_i$, where $\frac{1}{n}\sum_{i=1}^{n}\kappa_i = \kappa$ is a positive definite $\{2 \times 2\}$ matrix. Thus, following Theorem 2, we get

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial m_i}{\partial \zeta}\sin(\theta_i - m_i)|_{\zeta_0} \xrightarrow{d} N\left[(0), \frac{\partial m_i}{\partial \zeta}\frac{\partial m_i}{\partial \zeta'}E\{\sin^2(\theta_i - m_i)\}|_{\zeta_0}\right].$$

Theorem A.1 (*Multivariate CLT for Independent Vectors*)**.** *Let $\{X_i\}$ be a sequence of independent $\{k \times 1\}$ random vectors such that $P(|X_{i1}| \le C, |X_{i2}| \le C, \dots, |X_{ik}| \le C) = 1$ for all i, where $C \in (0, \infty)$. Let $E(X_i) = \mu_i$, $\operatorname{cov}(X_i) = \Psi_i$, and suppose that $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Psi_i = \Psi$ is a positive definite $\{k \times k\}$ matrix. Then*

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_i) \to N(0_k, \Psi) \quad \textit{in distribution},$$

where 0_k represents the $k \times 1$ 0 vector.

Proof. See [14, pp. 274–276]. $\square$

References

[1] S. Bhattacharya, A. SenGupta, Bayesian analysis of semiparametric linear–circular models, Journal of Argicultural, Biological, and Environmental Statistics 14 (2009) 33–65.
[2] P.J. Brown, Measurement, Regression, and Calibration, Oxford University Press, Oxford, 1993.
[3] A. Cameron, P. Trivedi, Microeconometrics, Methods and Applications, Cambridge University Press, New York, 2005.
[4] T.D. Downs, K.V. Mardia, Circular regression, Biometrika 89 (2002) 683–698.
[5] H. Eisenhart, The interpretation of certain regression methods and their use in biological and industrial research, The Annals of Mathematical Statistics 10 (1939) 162–186.
[6] N.I. Fisher, Statistical Analysis of Circular Data, Cambridge University Press, Cambridge, 1993.
[7] N.I. Fisher, A.J. Lee, Regression models for an angular response, Biometrics 48 (1992) 665–677.
[8] A.L. Gould, A regression technique for angular variates, Biometrics 25 (1969) 683–700.
[9] S.R. Jammalamadaka, A. SenGupta, Topics in Circular Statistics, in: Series in Multivariate Analysis, World Scientific Publishing Co., Inc., River Edge, NJ, 2001.
[10] R.A. Johnson, T.E. Wehrly, Bivariate models for dependence of angular observations and a related Markov process, Biometrika 66 (1978) 255–256.
[11] S. Kim, Inverse circular regression with possibly asymmetric error distribution, Ph.D. Dissertation, University of California, Riverside, 2009.
[12] S. Kim, New calibration methods with asymmetric data, Korean Journal of Applied Statistics 23 (2010) 759–765.
[13] R.G. Krutchkoff, Classical and inverse regression methods of calibration, Technometrics 9 (1967) 425–439.
[14] R.C. Mittelhammer, Mathematical Statistics for Economics and Businesses, Springer-Verlag New York Inc., 1999.
[15] C. Osborne, Statistical calibration: a review, International Statistical Review 59 (1991) 309–336.
[16] A. SenGupta, F. Ugwuowo, Asymmetric circular–linear multivariate regression models with applications to environmental data, Environmental and Ecological Statistics 13 (2006) 299–309.
[17] G.M. Tallis, Note on a calibration problem, Biometrika 56 (1969) 505–508.
[18] E.J. Williams, A note on regression methods in calibration, Technometrics 11 (1969) 189–192.
[19] E.A. Yfantis, L.E. Borgman, Extension of the von Mises distribution, Communications in Statistics—Theory and Methods 11 (1982) 1695–1706.