Journal of Economics and Administrative Sciences
2024; 30(144), pp. 473-484

P-ISSN 2518-5764
E-ISSN 2227-703X

Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq
DOI: https://doi.org/10.33095/2cbwj529

Nadaraya-Watson Estimation of a Circular Regression Model on Peak Systolic Blood Pressure Data

Rana Sadiq Nazer*
Department of Statistics
College of Administration and Economics
University of Baghdad, Iraq
E-mail: rana.musa2101@coadec.uobaghdad.edu.iq
Orcid: https://orcid.org/0009-0006-8048-8326
***Corresponding author**

Omar Abdulmohsin Ali
Department of Statistics
College of Administration and Economics
University of Baghdad, Iraq
E-mail: dromar72@coadec.uobaghdad.edu.iq
Orcid: https://orcid.org/0000-0003-0274-9325

Received:12/3/2024 Accepted: 6/10/2024 Published Online First: 1 /12/ 2024

Abstract:
Purpose: The research aims to estimate models representing phenomena that follow the logic of circular (angular) data, accounting for the 24-hour periodicity in measurement.
Theoretical framework: The regression model is developed to account for the periodic nature of the circular scale, considering the periodicity in the dependent variable y, the explanatory variables x, or both.
Design/methodology/approach: Two estimation methods were applied: a parametric model, represented by the Simple Circular Regression (SCR) model, and a nonparametric model, represented by the Nadaraya-Watson Circular Regression (NW) model. The analysis used real data from 50 patients at Al-Kindi Teaching Hospital in Baghdad.
Findings: The Mean Circular Error (MCE) criterion was used to compare the two models, leading to the conclusion that the Nadaraya-Watson (NW) circular model outperformed the parametric model in estimating the parameters of the circular regression model.
Research, Practical & Social Implications: The recommendation emphasized using the Nadaraya-Watson nonparametric smoothing method to capture the nonlinearity in the data.
Originality/value: The results indicated that the Nadaraya-Watson circular model (NW) outperformed the parametric model.
Paper type *Research paper.*
Keywords: Circular regression, Möbius transformation, Von Mises distribution, angular measurement, circular kernel, mean circular error.
JEL Classification: C10, C13, C14.
Authors' individual contribution: the Methodology and Writing — Rana Sadiq Nazer.; Reviewing & Editing — Supervision Omar Abdulmohsin Ali
Declaration of conflicting interests: The Authors declare that there is no conflict of interest.

P-ISSN 2518-5764
E-ISSN 2227-703X

1.Introduction:

Over the past two decades, the use of modified regression analysis has expanded significantly to address the complexities of circular or angular data, particularly in disciplines like medicine, agriculture, social sciences, finance, and more. In these contexts, the regression model largely relies on the values taken by the dependent variable. Consequently, many applied fields handling circular data—such as crystallography, biology, meteorology, and medicine—have grown more complex. Various methods have been developed for detecting linear circular regression models, such as Gould's approach (1969). In 1989, Bai et al. proposed a kernel estimator for directional data, $y = f(x)$, under the conditions of pointwise and uniform strong consistency. Expanding on this area, Choi and Lim (2000) presented a generalization of the Möbius transformation on the complex plane.

The maximum likelihood estimator for the Cauchy regression model was derived using harmonic patterns and the Möbius transformation within a fractional regression framework (McCullagh, 1996). This approach highlights the suitability of Möbius transformation for regression curve fitting, particularly when both the dependent and explanatory variables are angular, as demonstrated in wind direction data applications (Kato et al., 2008). Additionally, a notable extension to angular regression was introduced through a link function addressing multiple cases (Downs & Mardia, 2002).

Abuzaid et al. (2008) proposed a novel method for outlier detection using simulation experiments, while Abuzaid et al. (2011) developed cut-off points for the COVARATIO statistical test, facilitating both outlier detection and model estimation. Further, Abuzaid and Allahham (2015) analyzed wind direction data from two major cities in the Gaza Strip, Palestine, estimating a new model via the maximum likelihood method and comparing it to an iterative suggested approach.

Follmann and Proschan (1999) explored Rayleigh null hypothesis testing for circular data using simulation, while Hornik and Grün (2014) detailed the implementation of fitting functions in R, employing the EM algorithm for maximum likelihood estimation and addressing normalization of the von Mises-Fisher distribution constant. Abbas and Abood (2022) compared the circular S-estimator to circular least squares using simulation experiments under three contamination scenarios: covariate contamination, dependent variable contamination, and both. Their comparison criteria included the median standard error, the median mean square error, and the median mean cosine of circular residuals.

Finally, Meilán-Vila et al. (2024) conducted a study on climate change, focusing on the Atlantic region across four seasons. By monitoring daily temperatures, they employed nonparametric kernel estimation and used asymptotic bias as the comparison criterion.

2.Methodology:

2.1 Mobius Transformations:

Transformations, in general, refer to processes that alter the position, size, or shape of two-dimensional figures or three-dimensional objects. In engineering mathematics, common types of transformations include translations, rotations, reflections, and glide reflections. These transformations involve repositioning shapes or objects over a specified distance and in a particular direction. Rotation involves turning shapes or objects around a specific point or axis, while reflection entails flipping shapes or objects across a defined line or plane. Glide reflections combine translation and reflection along a given direction. Such transformations are essential in engineering mathematics for analysing, comparing, and classifying shapes and objects.

The Möbius transformation, also known as Möbius rotation, can be described as a process that fully rotates a plane onto itself by moving a single point. Alternatively, it can be understood as a transformation of two-dimensional spaces in different planes by shifting the plane’s axis. Möbius transformations have widespread applications in mathematical and engineering fields, including design, technical drawing, and three-dimensional modelling.

P-ISSN 2518-5764
E-ISSN 2227-703X

They also play a significant role in disciplines such as physics and advanced mathematics. Named after the German mathematician August Ferdinand Möbius, these transformations are sometimes referred to as homogeneous or homogeneity transformations, as well as partial linear or rotational transformations (McCullagh, 1996).
The general form of Möbius transforms is:

- $f(z) = \frac{az+b}{cz+d}$

Since:

- Z is a nodal variable.
- a, b, c, d are complex numbers.
- ad - bc $\neq 0$.

By applying the Möbius transformation geometrically, it is possible to achieve three-dimensional projection transformations. This process is defined as a specific function that projects a sphere onto a plane, excluding one point on the sphere, known as the point of projection. The result of this transformation is angle-preserving, meaning it maintains the angles formed by intersecting curves. Stereographic projection, a key application of this concept, is widely utilized in various fields, including complex analysis, cartography, geology, and photography.

Practically, the projection can be carried out either computationally or manually, often using specialized graph paper designed for such tasks. In the case of the two-unit sphere, the transformation involves rotating the sphere and relocating it to a new position and orientation in space. The stereographic projection is then performed from the sphere's new position. These transformations preserve angles, map straight lines to lines or circles, and convert circles into either lines or other circles.

2.2 Circular Normal Distribution:

The circular normal distribution, also known as the Von Mises distribution, is a continuous probability distribution analogous to the normal distribution, except that it is defined on a circular domain spanning $[0,2\pi]$. It serves a central role in analyzing circular data, much like the normal distribution does for linear data. The Von Mises distribution can also be viewed as a specific case of the Von Mises-Fisher distribution, which extends to multi-dimensional or multi-domain contexts.
This distribution is widely favored due to its flexibility regarding parameter influence and its ease of interpretation. Originally described by Richard Von Mises in 1918 to model the distribution of atomic weights (Hornik & Grün, 2014), the Von Mises distribution has since found applications in modelling a variety of phenomena across different fields. An example is provided below:

- Rotational motion (physics).
- Epidemiology (spread of disease). The general form of the probability density function of the Von Mises distribution is (Mahmood, E.A. et al. , 2019):
- $f(\theta, \mu, k) = \frac{1}{2\pi I_0(k)} \exp[k\cos(\theta - \mu)]\ 0 \leq \theta, \mu < 2\pi, K > 0,$

where:

- θ represents a circular observation,
- μ is the mean direction,
- k is the concentration parameter,
- $I_0(k)$ denotes the modified Bessel function of the first kind and order zero.

The modified Bessel function, $I_0(k)$, is defined as:

P-ISSN 2518-5764
E-ISSN 2227-703X

$$I_0(k) = \frac{1}{2\pi}\int_0^{2\pi} e^{k\cos(\vartheta)}\,d\vartheta$$

This formulation establishes the mathematical foundation for the Von Mises distribution, which is critical for modelling circular data. When the values of k are large and approach infinity, the data becomes increasingly clustered, and the distribution converges toward the envelope Cauchy distribution with a mean and variance of 1/k. In this case, a high concentration parameter (k) corresponds to low variance, while a low concentration indicates higher variance. Conversely, if k=0, the distribution becomes uniform.

The Von Mises distribution was pioneering in enabling scientists to model circular response data using linear predictors. Researchers proposed that a circular response (θ), which follows the Von Mises distribution, creates a distinctive pattern resembling a "barber's pole"—spirals wrapped around an infinite unit cylinder. This cylinder represents the average direction of the circular response (θ) conditioned on a real-valued predictor (x).

2.3 Trigonometric Transformation Functions:

Trigonometric transformation functions, also known as trigonometric, angular, or circular functions, are a set of real functions that relate the angles of a right triangle to the ratios of its sides. The most well-known basic trigonometric functions include the sine function (sin), the cosine function (cos), and the tangent function (tan). Additionally, their reciprocal functions cosecant (csc), secant (sec), and cotangent (cot) are also classified as trigonometric function:

- Cosecant (csc).
- Categor (sec).
- Cotangent (cot).

The reciprocal of the sine function is the cosecant (csc), and the reciprocal of the cosine function is the secant (sec). Trigonometric functions can generally be defined as the ratios between the sides of a right triangle containing the given angle or, more broadly, as coordinates on the unit circle.

When referring to triangles, it typically implies triangles on a flat, Euclidean surface, where the sum of the interior angles is always (180˚). Trigonometric functions can also be defined using integrals, power series, and differential equations, each of which has its specific applications. The variable for trigonometric functions can be an angle or a real number. These functions exhibit unique properties, including being even or odd, periodic, continuous, and orthogonal. Their primary application is in calculating the lengths of sides, angles, and related factors in triangles. This capability is widely utilized across various fields, such as surveying, navigation, and physics.

In surveying, triangulation is used to calculate the coordinates of specific points, a method now commonly employed in optical measurement. In navigation, trigonometric functions are applied to determine ship coordinates, plot routes, and calculate distances. In geography, they are used to compute distances between two points on the Earth's surface and to determine the Qibla direction by calculating its angle relative to the north. In optics, these functions are integral in studying light refraction.

Trigonometric functions are periodic, meaning their values repeat over regular intervals. This characteristic makes them essential for representing cyclical phenomena such as waves and forms the basis of the Fourier transform resulting in a mathematical process that transforms a function with real variables and complex values into another function of the same type. Additional applications include the study of alternating currents and sine wave estimations in the electric power and communications industries.

All trigonometric functions are periodic, with the smallest period being (2π), except for tangent (tan) and cotangent (cot), whose smallest period is (π).

P-ISSN 2518-5764
E-ISSN 2227-703X

The equation , $c^2 = a^2 + b^2$, represents a fundamental law used to calculate the length of the third side of a triangle when the other two sides and the enclosed angle are known. This relationship is a cornerstone of trigonometry and geometry.

2.4 Simple Circular Regression (SCR) Model

Several simple linear cyclic regression models have been proposed. In 2004, Hussin et al. introduced a basic cyclic regression model, which is conceptually similar to the simple linear regression model, with its framework adapted from earlier work (A. Abuzaid et al., 2011; Mahmood et al., 2019), as previously mentioned. The model can be expressed as:

- $y_i = \alpha + \beta x_i + \varepsilon_i (\mathrm{mod}\ 2\pi)$ (1)

Where:

- y_i represents the dependent variable.
- x_i represents the independent variable
- α represents the constant term of the model
- β represents the marginal slope of the model
- ε_i represents the random error that follows a Von Mises distribution with a circular mean of zero and a concentration parameter (k).

For a sample size n, the dependent variable y has the following observations: $(y_1, y_2, \ldots, y_n)$. The probability density function (p.d.f.) for all observations of the dependent variable is expressed as:

- $l(y_1) = \frac{1}{2\pi I_0(k)} e^{\{k\cos(y_1-\alpha-\beta x_1)\}}$ (2)
- $l(y_2) = \frac{1}{2\pi I_0(k)} e^{\{k\cos(y_2-\alpha-\beta x_2)\}}$ (3)
- $l(y_n) = \frac{1}{2\pi I_0(k)} e^{\{k\cos(y_n-\alpha-\beta x_n)\}}$ (4)

Assuming that the observations of the dependent variable are independent, the joint probability density function (j.p.d.f.) is the product of the individual p.d.f.s of each observation, which is written as:

- $l(y_1, y_2, \ldots, y_n) = \frac{1}{(2\pi I_0(k))^n} e^{\sum_{i=1}^{n}\{k\cos(y_i-\alpha-\beta x_i)\}}$ (5)

Taking the logarithm of this equation yields:

- $\log l\,(\alpha, \beta, k, x_1, \ldots \ldots x_n, y_1, \ldots \ldots y_n) = -n\log(2\pi) - n\log I_0(k) + k\sum(\cos(y_i - \alpha - \beta x_i))$ (6)

Where $I_0(k)$ is the zero-order modified Bessel function.

To estimate the parameters, we take the first derivatives of the log-likelihood function with respect to each parameter:

- $\frac{\partial \log l\,(,y_1,\ldots\ldots y_n)}{\partial \alpha} = \sum_{i=1}^{n}\{k\sin(y_i - \alpha - \beta x_i)\}$ (7)
- $\frac{\partial \log l\,(,y_1,\ldots\ldots y_n)}{\partial \beta} = \sum_{i=1}^{n}\{k\sin(y_i - \alpha - \beta x_i)x\}$ (8)
- $\frac{\partial \log l\,(,y_1,\ldots\ldots y_n)}{\partial k} = \frac{-n\,I_{1(k)}}{I_{0(k)}} + \sum_{i=1}^{n}\{k\cos(y_i - \alpha - \beta x_i)\}$ (9)

Where $I_1(k)$ is the first derivative of $I_0(k)$.

Setting the derivatives equal to zero results in the following equations:

- $k\sum_{i=1}^{n}\{\sin(y_i - \alpha - \beta x_i)\} = 0$ (10)
- $k\sum_{i=1}^{n}\{\sin(y_i - \alpha - \beta x_i)x\} = 0 \frac{-nI_1(k)}{I_0(k)} + \sum_{i=1}^{n}\{k\cos(y_i - \alpha - \beta x_i)\} = 0$ ... (11)

P-ISSN 2518-5764
E-ISSN 2227-703X

From these equations, the estimation of $\widehat{\alpha}$ is derived as:

- $\hat{a} = \begin{cases} \tan^{-1}(s/c) & \text{if } s \geq 0, c > 0 \\ \tan^{-1}(s/c) + \pi & \text{if } c < 0 \\ \tan^{-1}(s/c) + 2\pi & \text{if } s < 0, c \geq 0 \end{cases}$ (12)

Where:

- $S = \sum \sin(y_i - \beta_0 x_i)$
- $C = \sum \cos(y_i - \beta_0 x_i)$
- β_0 is the initial estimate of β.

The parameter $\widehat{\beta}$ can be estimated as:

- $\widehat{\beta}_1 \approx \widehat{\beta}_0 + \frac{\sum x_i \sin(y_i - \widehat{\alpha} - \widehat{\beta}_0 x_i)}{\sum x_i^2 \cos(y_i - \widehat{\alpha} - \widehat{\beta}_0 x_i)}$ (13)

Finally, the concentration parameter k is estimated by:

- $\widehat{K} = A^{-1}\left(\frac{1}{n} \sum \cos\left(y_i - \widehat{\alpha} - \widehat{\beta} x_i\right)\right)$ (14)

Where A^{-1} is the inverse of the Bessel function of the first kind of order 1, approximated by:

- $A^{-1}(w) \approx \frac{9 - 8w + 3w^2}{8(1-w)}$ (15)

Here, w is a real number and is defined as:

- $w = \frac{1}{n} \sum \cos\left(y_i - \widehat{\alpha} - \widehat{\beta} x_i\right)$ (16)

2.5 Circular Nonparametric Regression Model:

Circular data refers to measurements taken on a circle in either degrees or radians, and it exhibits periodicity in various applied fields, such as biology (e.g., animal movement direction), meteorology (wind direction), and oceanography (ocean currents). In this context, the circular nonparametric regression estimator plays a crucial role, as traditional parametric circular models may lack the flexibility to capture complex data distributions. These distributions are represented as points on the circumference of a unit circle. Specifically, we present an estimator derived from asymptotic precision measures similar to those in Euclidean space. Nonparametric regression methods, such as the Nadaraya-Watson (NW) estimator (Rasheed et al., 2012), are widely used in this context.

The objective is to propose and study a nonparametric regression estimator for a model involving a circular response variable and a covariate. When the response variable is circular, the regression function is defined as the minimizer of the circular hazard function. It has been shown that the minimizer of this risk function corresponds to the inverse tangent of the ratio between the conditional expectations of the sine and cosine of the response variable (Bai et al., 1989). We introduce two regression models: one for the sine and one for the cosine of the response variable. Subsequently, a nonparametric estimator of the regression function is obtained by calculating the inverse tangent of the ratio of the NW estimators for the sine and cosine functions.

Let $\{(x_i, \Theta_i)\}$ be a simple random sample from the vector $\{(x, \Theta)\}$,

Where:

(Θ) is a circular random variable taking values in $\mathbb{T} = [0, 2\pi)$, and X is a random variable on E. We assume that the circular random variable Θ depends on X via the following regression model:

- $\Theta_i = [m(x_i) + \varepsilon_i](\text{mod } 2\pi),\ i = 1, \dots, n$ (17)

Here, (m) is the regression function mapping E onto T. The circular regression function is defined as the minimizer of the risk function:

- $E[\sin(\varepsilon)|X = x| = 0$
- $E\{1 - \cos[\Theta - m(x)]\}|X = x|$

The sine function is minimized by the following equation:

- $m(x) = \operatorname{atan} 2[m_1(x), m_2(x)]$ (18)

Where:

- $m_1(x) = E[\sin(\Theta)|X - x|$
- $m_2(x) = E[\cos(\Theta)\ |X - x|$

The function $\operatorname{atan2}(y, x)$ returns the angle between the x-axis and the vector from the origin to the point (x,y). Using this formula, m1(x) and m2(x) act as regression functions for the sine and cosine models, respectively. The model assumes:

- $\sin(\Theta_i) = m_1(x_i) + \xi_i$ (19)
- $\cos(\Theta_i) = m_2(x_i) + \zeta_i$ (20)
- $i = 1, \dots, n$

where m1 and m2 are regression functions over the interval [−1,1] and, ξ_i and ζ_i are the independent angular errors. Assuming that both models hold simultaneously with Equation (18) leads to relationships between the variances and covariances of the errors in these models. Using the sine and cosine formulas from Equation (18), we derive:

- $\sin(\Theta_i) = \sin[m(x_i)]\cos(\varepsilon_i) + \cos[m(x_i)]\sin(\varepsilon_i)$ (21)
- $\cos(\Theta_i) = \cos[m(x_i)]\cos(\varepsilon_i) + \sin[m(x_i)]\sin(\varepsilon_i)$ (22)

Thus, the functions $f_1(x) = \sin[m(x)]$ and $f_2(x) = \cos[m(x)]$are defined, leading to:

- $m_1(x) = f_1(x)\ell(x)$, $m_2(x) = f_2(x)\ell(x)$ (23)

Where $f_1(x)$and $f_2(x)$ correspond to $m_1(x)$ and $m_2(x)$, and $\ell(x)$ is defined as:

- $\ell(x) = [m^2{}_1(x) + m^2{}_2(x)]^{1/2}$

The error terms in Equations (19) and (20) are expressed in terms of the conditional variances and the covariance of the Cartesian coordinates of ε:

- $s_1^2(x) = f_1^2(x)\sigma_2^2(x) + 2f_1(x)f_2(x)\sigma_{12}(x) + f_2^2(x)\sigma_1^2(x)$ (24)
- $s_2^2(x) = f_2^2(x)\sigma_2^2(x) - 2f_2(x)f_1(x)\sigma_{12}(x) + f_1^2(x)\sigma_1^2(x)$ (25)

The variance between the error terms is given by:

- $c(x) = f_1(x)f_2(x)\sigma_2^2(x) - f_1^2(x)\sigma_{12}(x) + f_2^2(x)\sigma_{12}(x) - f_1(x)f_2(x)\sigma_1^2(x)$ (26)

The NW estimator for the circular regression function in Equation (18) is presented as:

- $\widehat{m}_h(x) = \operatorname{atan} 2\left[\widehat{m}_{1,h}(x), \widehat{m}_{2,h}(x)\right]$ (27)

where $(\widehat{m}_{1,h}(x))$ and $(\widehat{m}_{2,h}(x))$ are the NW estimators for $m_1(x)$ and $m_2(x)$, respectively.
The asymptotic properties of this estimator, including bias, variance, and asymptotic normality, are derived assuming that Equations (19) and (20) hold. The NW estimator is given by:

- $$\widehat{m}_{j,h}(x) = \begin{cases} \dfrac{\sum_{i=1}^{n} K(h^{-1}\|X_i - x\|)\sin(\Theta_i)}{\sum_{i=1}^{n} K(h^{-1}\|X_i - x\|)} & \text{if } j = 1 \\ \dfrac{\sum_{i=1}^{n} K(h^{-1}\|X_i - x\|)\cos(\Theta_i)}{\sum_{i=1}^{n} K(h^{-1}\|X_i - x\|)} & \text{if } j = 2 \end{cases} \quad (28)$$

Where k represents the symmetric kernel function, while h (or $h = h_n$) is the smoothing parameter, also referred to as the bandwidth. The bandwidth is a positive real value that determines the smoothness of the estimator. While the choice of the kernel function is generally of secondary importance, the bandwidth plays a critical role in the performance of the Nadaraya-Watson (NW) estimators.

For the regression estimator described in Equation (18), if the bandwidth h is too large, an excessive number of observations are included in the estimation process, resulting in an oversmoothed estimator. Conversely, if h is too small, only a limited number of observations are considered, leading to a highly variable estimator. Hence, bandwidth selection is essential for obtaining reliable estimates. In practice, data-driven methods are employed to determine the optimal bandwidth. For consistency, the same bandwidth is used to calculate ($\widehat{m}_{1,h}$) and ($\widehat{m}_{2,h}$) in Equation (19). Using different bandwidths for the sine and cosine components would result in inconsistent regression estimators, as the Cartesian coordinates correspond directly to the same angular measure.

To derive the bias and variance of the estimator in Equation (18), the properties of NW estimators are utilized. From these properties, the following definitions and results are obtained:

- φ_x , $\varphi_{j,x}$, $j = 1,2$

are the functions defined for each ($s \in \mathbb{R}$) Vintage:

- $\varphi_x(s) = \mathbb{E}\{[m(X) - m(x)] \mid \|X - x\| = s\}$ (29)
- $\varphi_{j,x}(s) = \mathbb{E}\{[m_j(X) - m_j(x)] \mid \|X - x\| = s\}$ (30)

We denote the cumulative distribution functions of the random variable, $(X - x)$ as:

- $F_x(t) = \mathbb{p}(\|X - x\| \le t),\ t \in \mathbb{R}$
- $\tau_{x,h}(s) = \frac{F_x(hs)}{F_x(h)} = \mathbb{p}(\|X - x\| \le hs \mid \|X - x\| \le h)$

Which leads to:

- $\mathbb{E}[\widehat{m}_{j,h}(x) - m_j(x)] = \acute{\varphi}_{j,x}(0)\frac{M_{x,0}}{M_{x,1}}h + \mathcal{O}\left[\frac{1}{nF_x(h)}\right] + o(h)$
- $\mathbb{V}ar[\widehat{m}_{j,h}(x)] = \frac{s_j^2(x)}{nF_x(h)}\frac{M_{x,2}}{M_{x,1}^2} + o\left[\frac{1}{nF_x(h)}\right]$
- $\mathbb{C}ov[\widehat{m}_{1,h}(x) - \widehat{m}_{2,h}(x)] = \frac{c(x)}{nF_x(h)}\frac{M_{x,2}}{M_{x,1}^2} + o\left[\frac{1}{nF_x(h)}\right]$
- $M_{x,0} = K(1) - \int_0^1 [sK(s)]'\,\tau_{x,0}(s)\,ds$
- $M_{x,1} = K(1) - \int_0^1 K'(s)\,\tau_{x,0}(s)\,ds$
- $M_{x,2} = K^2(1) - \int_0^1 (K^2)'(s)\,\tau_{x,0}(s)\,ds$

From the above equations, we end up with the two equations:

- $\mathbb{E}[\widehat{m}_{,h}(x) - m(x)] = \acute{\varphi}_{,x}(0)\frac{M_{x,0}}{M_{x,1}}h + \mathcal{O}\left[\frac{1}{nF_x(h)}\right] + o(h)$ (31)
- $\mathbb{V}ar[\widehat{m}_{,h}(x)] = \frac{1}{nF_x(h)}\frac{\sigma_1^2(x)}{\ell^2(x)}\frac{M_{x,2}}{M_{x,1}^2} + o\left[\frac{1}{nF_x(h)}\right]$ (32)

As for the bandwidth parameter or smoothing parameter (h), it can be calculated from the equation:

$$CV(h) = \sum_{i=1}^{n}\left\{1 - \cos\left[\Theta_i - \widehat{m}_h^{(i)}(x_i)\right]\right\} \qquad (33)$$

Where ($\widehat{m}_h^{(i)}$)denotes estimator (NW).

The bandwidth parameter h is typically determined using automated software routines, which iteratively search within a specified range to converge on the optimal value, eliminating subjective bias in its selection.

P-ISSN 2518-5764
E-ISSN 2227-703X

2.6 Comparison Criterion

After reviewing the various estimation methods for circular regression, their performance can be compared using the **Mean Circular Error (MCE)** criterion, defined as:

- $MCE_s = \frac{1}{n}\sum_{j=1}^{n} \sin\left(\frac{d_j}{2}\right)$ (34)

Where:

- (MCE_s) represents the average circular error.
- $d_j = \pi - \left|\pi - \left|v_j - \hat{v}_j\right|\right|$
- $MCE_s \in [0,1]$

3. Data Collection:

The data were collected from Al-Kindi Teaching Hospital in Baghdad during April 2023. A sample of 50 patients with systolic blood pressure readings was used. The time of blood pressure measurement was recorded at peak times, with the study period divided into two sections. The first section corresponds to the time of blood pressure measurement during the first week, while the second section corresponds to the second week. The estimation methods presented in the theoretical section were applied, and the parameters of both parametric and nonparametric circular regression models were estimated using these methods.

4. Results:

The data were collected in the form of two daily readings for each patient in the sample (n=50). Systolic blood pressure readings were recorded separately for each patient in degrees, twice per day during rest periods. The independent (explanatory) variable was denoted as u, while the dependent variable (peak systolic blood pressure) was denoted as v. The relationship between these variables was examined based on Mardia's theory, which assumes that conditions must remain similar at the time of measurement. The day was divided into two periods of 12 hours each. The first period ran from midnight (12:00 AM) to noon (12:00 PM), and the second period from noon (12:00 PM) to midnight (12:00 AM). The measurement time at midnight was considered the starting point, corresponding to an angle of 0°, while noon corresponded to 180°. The second period began at noon, with the angle progressing back to 0° by midnight. The 180° span was divided by the number of hours (12 hours), giving an angle increment of 15° per hour. Data were recorded for two groups, labelled S1 and S2.

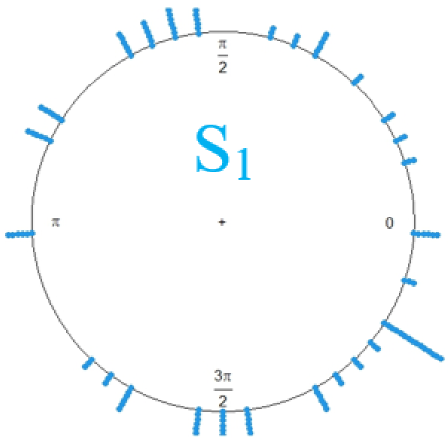

Figure 1: Real data distribution (S1)

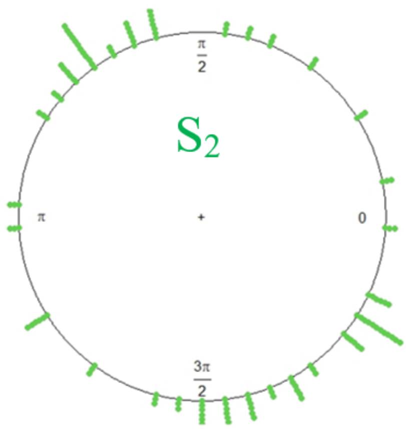

Figure 2: Distribution of real data (S2)

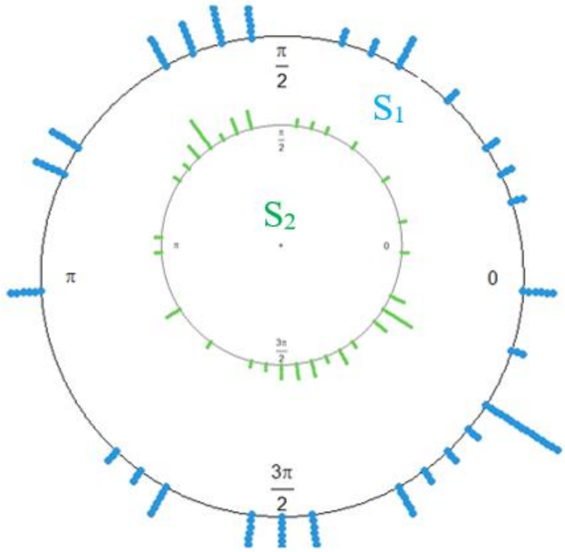

Figure 3: Distribution of real data within the unit circle (S1, S2)

Table (1) Real data for patients with systolic blood pressure

t	1	2	3	4	5	6	7	8	9	10
S1	45	55	-95	-45	-50	-80	-156	-144	-17	-114
S2	15	-138	-32	-34	-91	-102	-120	-100	-154	1

t	11	12	13	14	15	16	17	18	19	20
S1	2	-44	-77	-102	30	-17	-23	23	-65	-135
S2	-14	-13	-32	-48	-76	-98	-39	-102	-17	2

t	21	22	23	24	25	26	27	28	29	30
S1	-126	-120	-84	-137	-124	-46	62	-14	-76	-13
S2	-44	-23	-17	-45	-77	-65	-95	-180	-121	-180

t	31	32	33	34	35	36	37	38	39	40
S1	-32	-98	-47	2	-102	-33	-150	-120	90	12
S2	-162	-30	-165	-136	-102	-17	-48	-45	-13	-95

t	41	42	43	44	45	46	47	48	49	50
S1	36	-61	-162	165	-98	-167	-45	-35	-61	-132
S2	-39	-17	-1	-136	2	-140	-166	-93	-108	-77

To evaluate and compare the efficiency of the estimated parametric and nonparametric models, we will use the Mean Circular Error (MCE) as the standard metric, as previously defined in Equation (35). The results of this comparison are summarized in the table below:

Table (2): Results of comparison between models for real data

Model	MCE
SCR	0.2089918
NW	0.1104407

P-ISSN 2518-5764
E-ISSN 2227-703X

The second column of the table above represents the mean circular error. It is evident that the lowest value corresponds to the Nadaraya-Watson circular model (NW), demonstrating its superior performance in practical applications. This model outperformed the circular least squares (SCR) model, likely due to the nonlinear nature of the data. Based on these results, the Nadaraya-Watson circular model (NW) is identified as the most effective model. In other words, circular nonparametric models are shown to be the most suitable for analyzing circular data.

5. Conclusion:

Based on our findings, the following conclusions were drawn:

i. The applied experiment demonstrated the superior performance of the Nadaraya-Watson circular model (NW) compared to the circular least squares model (SCR) in fitting the circular regression model.

ii. The sample size significantly influences the statistical testing of parametric models. Larger sample sizes enhance the likelihood of obtaining statistically significant models.

Acknowledgements:

The authors express their sincere gratitude to the referees for their valuable comments and suggestions, which have greatly enhanced the quality of this paper.

Authors Declaration:

Conflicts of Interest: None

-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved by The Local Ethical Committee in The University.

References:

Abbas, H. H., & Abood, S. N. (2022). Comparison of Robust Circular S and Circular Least Squares Estimators for Circular Regression Model using Simulation. *Journal of Economics and Administrative Sciences*, *28*(134), 168–185.

Abe, T., & Pewsey, A. (2011). Symmetric circular models through duplication and cosine perturbation. *Computational Statistics & Data Analysis*, *55*(12), 3271–3282.

Abuzaid, A. H., & Allahham, N. R. (2015). SIMPLE CIRCULAR REGRESSION MODEL ASSUMING WRAPPED CAUCHY ERROR. *Pakistan Journal of Statistics*, *31*(4).

Abuzaid, A. H., Hussin, A. G., & Mohamed, I. B. (2008). Identifying single outlier in linear circular regression model based on circular distance. *Journal of Applied Probability and Statistics*, *3*(1), 107–117.

Abuzaid, A. H., Hussin, A. G., & Mohamed, I. B. (2013). Detection of outliers in simple circular regression models using the mean circular error statistic. *Journal of Statistical Computation and Simulation*, *83*(2), 269–277.

Abuzaid, A., Mohamed, I., Hussin, A. G., & Rambli, A. (2011). COVRATIO statistic for simple circular regression model. *Chiang Mai Journal of Science*, *38*(3), 321–330.

Ahmed, A. D., Abdulwahhab, B. I., & Abdulah, E. K. (2020). A comparison among Different Methods for Estimating Regression Parameters with Autocorrelation Problem under Exponentially Distributed Error. *Baghdad Science Journal*, *17*(3 (Suppl.)), 980.

Bai, Z. D., Rao, C. R., & Zhao, L. C. (1989). Kernel estimators of density function of directional data. In *Multivariate statistics and probability* (pp. 24–39). Elsevier.

Choi, K., & Lim, Y. (2000). Inverse monoids of Möbius type. *Journal of Algebra*, *223*(1), 283–294.

Di Marzio, M., Panzera, A., & Taylor, C. C. (2013). Non-parametric regression for circular responses. *Scandinavian Journal of Statistics*, *40*(2), 238–255.

Downs, T. D., & Mardia, K. V. (2002). Circular regression. *Biometrika*, *89*(3), 683–698.

Fisher, N. I., & Lee, A. J. (1992). Regression models for an angular response. *Biometrics*, 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*, 683–700.

Hitchman, M. P. (2018). Geometry with An Introduction to Cosmic Topology. *Whitman College, USA*.

Hornik, K., & Grün, B. (2014). movMF: An R package for fitting mixtures of von Mises-Fisher distributions. *Journal of Statistical Software*, *58*(10), 1–31.

Jupp, P. E., & Mardia, K. V. (1989). A unified view of the theory of directional statistics, 1975-1988. *International Statistical Review/Revue Internationale de Statistique*, 261–294.

Kato, S., Shimizu, K., & Shieh, G. S. (2008). A circular–circular regression model. *Statistica Sinica*, 633–645.

Kazem, A. H., & Al-Dulaimi, M. M. (1988). *Introduction to Linear Regression Analysis*. Dar Al-Kutub for Printing and Publishing, Iraq

Lund, U. (1999). Least circular distance regression for directional data. *Journal of Applied Statistics*, *26*(6), 723–733.

McCullagh, P. (1996). Möbius transformation and Cauchy parameter estimation. *The Annals of Statistics*, *24*(2), 787–808.

Meilán-Vila, A., Crujeiras, R. M., & Francisco-Fernández, M. (2024). Nonparametric estimation for a functional-circular regression model. *Statistical Papers*, *65*(2), 945–974.

Mahmood, E.A., Midi, H., & Hussin, A.G. (2019). The Comparison Between Maximum Weighted and Trimmed Likelihood Estimator of the Simple Circular Regression Model. *Journal of Modern Applied Statistical Methods, 18(2), 1-14.*

Rasheed, D.H., Hmood, M.Y., & Hamza, S. K. (2012). Nadaraya-Watson Estimator a Smoothing Technique for Estimating Regression Function. *Journal of Economics and Administrative Sciences, 18(65), 283–291*.

Nazer, R. S. & Ali, O. A. (2024). Some of Parametric and Non parametric Estimations for Circular Regression Model via Simulation. *Misan Journal for Academic Studies, 23(49), 331-340.*