



Correcting the trigonometric moments and resultant lengths for grouped circular data

S. Rao Jammalamadaka¹ · György Terdik² · Ali Abuzaid¹

Received: 30 October 2025 / Revised: 10 May 2026 / Accepted: 12 July 2026
© The Author(s) 2026

Abstract

We revisit the problem of correcting the trigonometric moments and the resultant lengths that are obtained from grouped circular data. In circular measurements as elsewhere, it is common in many practical applications to aggregate the data into intervals either for convenience, or because of lack of accuracy in measuring, and then assuming that all the observations in a given interval lie at the midpoint of that interval, for calculating needed statistics. It then becomes important to ask for formulae necessary to adjust for the error introduced by such approximations, and this problem has been considered earlier in Greenwood (1959) and Mardia (1972) focusing on the first two trigonometric moments, because they are the ones used mostly in practice. The proofs provided there also rely on other outside results. We take it considerably further with our principal contributions being twofold: (i) unlike earlier derivations, we provide very simple and straightforward derivations which are complete and hold for trigonometric moments of any order and (ii) make recommendations, based on extensive simulations, as to when and where such corrections make a significant difference. To further support our mathematical derivations, we provide extensive simulations using data that may come from 3 potentially common models, namely, the von Mises, the Generalized von Mises, and the Sine-skewed von Mises, and assess the performance of the correction formulae in these contexts. Our results demonstrate that the corrected trigonometric moments yield more accurate estimates, especially when the underlying distribution that generated the data deviate considerably from uniformity. The methods are illustrated using cross-bedding azimuth data from the Kamthi Formation, showing that grouping resolution can materially affect inference on directional concentration.

Keywords Correction for grouping · Higher-order trigonometric moments · Sheppard's corrections · Testing isotropy

Handling Editor: Luiz Duczmal

Extended author information available on the last page of the article

Published online: 30 July 2026

Springer

1 Introduction

In many applications involving circular measurements, such as animal movement studies, environmental monitoring, and directional sensing, observations are often recorded in grouped or discretized form due to measurement resolution or reporting conventions. Statistical analyses based on such grouped data typically treat all observations within an interval as occurring at the midpoint, which introduces systematic distortion in key circular summaries such as resultant lengths and trigonometric moments. Since these quantities underlie many standard inferential procedures in circular statistics, understanding and correcting the effect of grouping is of both theoretical and practical importance.

In this article, we revisit the classical problem of how the resultant length, on which, e.g., the so-called Rayleigh test for isotropy is based, and more generally the higher-order trigonometric moments are affected when the original data come grouped. We review what is already known, see Greenwood (1959) and Mardia (1972), where the proofs crucially depend on other results, and are mostly concerned with the first two trigonometric moments.

We fill in these important gaps giving simpler proofs and take it a step further as to how these corrections apply to the trigonometric moments of any higher order, as in the case of Sheppard's corrections for data on the real line.

A circular random variable $\Theta \in [0, 2\pi)$ is said to have a circular distribution with a Cumulative Distribution Function (CDF) $F(\vartheta)$ with respect to a chosen origin and a sense of rotation. When appropriate, it has a probability density function $f(\vartheta)$, $\vartheta \in [0, 2\pi)$. Such an $f(\cdot)$ is a periodic function with period 2π and integrates to 1 in the interval $[0, 2\pi)$. See, e.g., Chapter 2 of Jammalamadaka and SenGupta (2001) for some details and examples. We can and shall also represent this random variable Θ as a unit vector $\mathbf{X} = [\cos \Theta, \sin \Theta]^T \in \mathbb{S}_1$, where $\mathbb{S}_1$ denotes the surface (circumference) of the unit circle. Using the Euler formula, a typical unit vector $\mathbf{X} \in \mathbb{S}_1$ can be re-expressed as

$$\mathbf{X} = [\cos \vartheta, \sin \vartheta]^T = [(e^{i\vartheta} + e^{-i\vartheta})/2, (e^{i\vartheta} - e^{-i\vartheta})/2i]^T.$$

For any integer ℓ , the complex-valued coefficients

$$\phi_\ell = \int_0^{2\pi} e^{i\ell\vartheta} f(\vartheta) d\vartheta = E(\cos \ell \vartheta) + i E(\sin \ell \vartheta) = \alpha_\ell + i\beta_\ell = \rho_\ell e^{i\mu_\ell} \quad (1)$$

are referred to as the “trigonometric moments” of order ℓ .

When the density function $f(\vartheta)$ is continuous, it has the Fourier series expansion in terms of coefficients $\{\phi_\ell\}$:

$$f(\vartheta) = \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \phi_\ell e^{-i\ell\vartheta}, \quad (2)$$

where the coefficients satisfy $\overline{\phi_\ell} = \phi_{-\ell}$ with $\overline{\phi_\ell}$ denoting the complex conjugate of ϕ_ℓ and the series (2) is assumed to converge uniformly to f (see Carslaw (1930)). Note that f is real valued with a characteristic function at the integer values,

$\{\phi_\ell = E e^{i\ell\Theta} = \rho_\ell e^{i\mu_\ell}\}$. Thus, the form (2) is indeed the inverse of the characteristic function. We have

$$\begin{aligned}\phi_\ell e^{-i\ell\vartheta} + \overline{\phi_\ell} e^{i\ell\vartheta} &= 2 \operatorname{Re}(\phi_\ell e^{-i\ell\vartheta}) \\ &= 2\alpha_\ell \cos(\ell\vartheta) + 2\beta_\ell \sin(\ell\vartheta) \\ &= 2\rho_\ell \cos(\mu_\ell - \ell\vartheta).\end{aligned}$$

Hence, the density can also be written as

$$f(\vartheta) = \frac{1}{2\pi} \left(1 + 2 \sum_{\ell=1}^{\infty} (\alpha_\ell \cos \ell\vartheta + \beta_\ell \sin \ell\vartheta) \right), \quad (3)$$

where $\{\phi_\ell; \alpha_\ell, \beta_\ell\}$ are the trigonometric moments of order ℓ , mentioned earlier. See Section 2.1 of Jammalamadaka and SenGupta (2001) for more details.

This paper is organized as follows. Section 2 derives the correction formula for estimating trigonometric moments from grouped data. Section 3 discusses the exact grouped trigonometric moments for cardioid and wrapped Cauchy distribution. Section 4 presents simulations comparing corrected and uncorrected moments for some typical circular distributions such as the von Mises, the Generalized von Mises, and the Sine-skewed von Mises distributions. Section 5 applies the proposed correction to classical cross-bedding azimuth data from the Kamthi Formation, examining how different grouping resolutions affect tests of uniformity and directional concentration.

2 Corrections for estimating the trigonometric moments from grouped data

Often, in practice, the observations from which these trigonometric moments are calculated are grouped either for convenience in measurement or for computational simplicity. In such a case, one assumes that all the observations in a given class interval are uniformly distributed in that interval, and in calculating the trigonometric moments, one assumes further that all the observations lie at the midpoint of that interval.

While such an assumption simplifies calculations, it introduces an inherent approximation error by ignoring the original distribution of these observations. The question of practical importance is when one calculates the trigonometric moments for grouped data under these assumptions, how accurate are these, i.e., what correction should one make to recover the values of the trigonometric moments of the original data, prior to such grouping?

Suppose the interval $[0, 2\pi)$ is subdivided into K sub-intervals of equal length $h = 2\pi/K$. For $k = 1, \dots, K$, define the k^{th} sub-interval I_k , to be $I_k = [(k-1)h, kh)$, with the midpoint

$$w_k = (2k-1)h/2.$$

Given the loss of information involved in such grouping, one identifies all the observations $e^{i\Theta}$ that fall in that sub-interval I_k ,

by the midpoint of this interval, viz. e^{iw_k} . We will denote such a grouped variable by $e^{i\Theta^*}$, which then has a discrete distribution with probabilities

$$P(e^{i\Theta^*} = e^{iw_k}) = P(\Theta \in I_k) = p_k = \int_{-h/2}^{h/2} f(w_k + \vartheta) d\vartheta$$

and with the corresponding characteristic function given by

$$E e^{i\Theta^*} = \sum_{k=1}^K p_k e^{iw_k}.$$

For any $\ell \geq 1$, we are interested in the relationship between the original trigonometric moments

$$\phi_\ell = E e^{i\ell\Theta} = \int_0^{2\pi} e^{i\ell\vartheta} f(\vartheta) d\vartheta = \sum_{k=1}^K \int_{I_k} e^{i\ell\vartheta} f(\vartheta) d\vartheta$$

and the corresponding ones obtained from the grouped data, say denoted by ϕ_ℓ^* given by

$$\phi_\ell^* = E e^{i\ell\Theta^*} = \sum_{k=1}^K e^{i\ell w_k} \int_{-h/2}^{h/2} f(w_k + \vartheta) d\vartheta$$

using the integral formula for p_k given above. When h is reasonably small, this expression can be approximated using midpoint averaging, equivalently viewed as a Riemann-sum approximation based on the integral mean-value theorem, by

$$\begin{aligned} & \frac{1}{h} \int_0^{2\pi} e^{i\ell\omega} \int_{-h/2}^{h/2} f(\omega + \vartheta) d\vartheta d\omega \\ &= \frac{1}{h} \int_{-h/2}^{h/2} e^{-i\ell\vartheta} \int_0^{2\pi} f(\omega + \vartheta) e^{i\ell(\omega+\vartheta)} d\omega d\vartheta \\ &= \phi_\ell \frac{1}{h} \int_{-h/2}^{h/2} e^{-i\ell\vartheta} d\vartheta = \phi_\ell \frac{2}{\ell h} \sin(\ell h/2). \end{aligned}$$

Thus, we have an approximation of ϕ_ℓ in terms of the grouped/observed ϕ_ℓ^* as

$$\phi_\ell \sim \phi_\ell^* \frac{\ell h/2}{\sin(\ell h/2)}. \quad (4)$$

The general form of the correction shows that grouped trigonometric moments of arbitrary order can be handled within a unified framework. In practice, lower-order moments are emphasized because they are most commonly used in directional inference, while higher-order corrections follow directly from the same formula.

Remark 1 The approximation (4) above appears to date back to Greenwood (1959) and also appears in Mardia (1972, Sec 3.8) who provides a fairly involved alternate

derivation based on Bernoulli polynomials as well as by borrowing some results that appear in Kendall and Stuart (1969). In our view, those arguments are not as simple or straightforward, and more importantly, do not readily extend to high-order trigonometric moments, as our current results do.

Remark 2 If the original f is uniform, no grouping correction will be needed as one would expect. In this case we know that $\phi_\ell = 0$. And since

$$e^{i\ell w_k} \int_{-h/2}^{h/2} f(w_k + \vartheta) d\vartheta = \frac{h}{2\pi} e^{i\ell w_k},$$

we have

$$\phi_\ell^* = \frac{h}{2\pi} \sum_{k=1}^K e^{i\ell w_k} = 0,$$

whenever ℓ is not a multiple of K , resulting in the same value as ϕ_ℓ . That is to say, if the original pdf is uniform or nearly so in any interval, the grouping and the assumptions we make in calculating the moments from it, are consistent and provide full recovery.

However, in general, the approximation (4) shows that

$$|\phi_\ell| > |\phi_\ell^*|,$$

i.e., grouping reduces the absolute value of trigonometric moments.

Remark 3 Consider the density $f_h(\eta)$ that is a uniform distribution a class interval $(-h/2, h/2]$, i.e., $f_h(\eta) = \mathbf{1}_{(-h/2, h/2]}/h$, then we obtain the Fourier transform of $f_h(\eta)$ as

$$\frac{1}{h} \int_{-h/2}^{h/2} e^{im\eta} d\eta = \frac{1}{imh} (e^{imh/2} - e^{-imh/2}) = \frac{2}{mh} \sin(mh/2).$$

We have the series expansion of $f_h(\eta)$

$$\begin{aligned} f_h(\eta) &= \frac{1}{2\pi} \left(1 + \frac{4}{h} \sum_{m=1}^{\infty} \frac{1}{m} \sin(mh/2) \cos m\eta \right) \\ &= \begin{cases} \frac{1}{h}, & |\eta| < h/2 < \pi \\ \frac{1}{2h}, & |\eta| = h/2 \\ 0, & \eta > h/2. \end{cases} \end{aligned}$$

The extra values $1/2h$ of $f_h(\eta)$ are noticed at the points $\pm h/2$. This is not a continuous density hence the assumptions in approximation (4) are not satisfied.

Remark 4 If we consider a distribution that has a density that is uniform over each sub-interval I_k then we shall have $\Theta = \Theta^* + U_h$ where Θ^* denotes the grouped variable taking the midpoint value w_k whenever $\Theta \in I_k$, and U_h represents the within-interval deviation from that midpoint. Θ^* and U_h are independent (Berkson's case, Schneeweiss et al. (2010)). As a result, we get

$$\phi_\ell = E e^{i\ell(\Theta^* + U_h)} = \phi_\ell^* \frac{1}{\ell h/2} \sin(\ell h/2),$$

quite in contrast to (4).

Starting from a random variate Θ , the rounding (or quantization) process $\Theta \Rightarrow \Theta^*$ generally does not imply independence between either $(\Theta, \Theta - \Theta^*)$ or $(\Theta^*, \Theta - \Theta^*)$. For related discussions in electrical engineering applications, see Vardeman (2005). Nevertheless, approximations based on treating Θ and U_h as independent often lead to correction formulae that are numerically very close to those obtained from the exact grouped data formulation.

3 On exact grouped trigonometric moments for some specific circular models

Let the circle be partitioned into K equal cells of width $h = 2\pi/K$ with midpoints $w_k = (2k - 1)h/2$. If F denotes the cumulative distribution function of Θ , then the grouped trigonometric moment admits the exact representation

$$\phi_\ell^* = \sum_{k=1}^K e^{i\ell w_k} [F(kh) - F((k-1)h)] = \sum_{k=1}^K e^{i\ell w_k} \int_{-h/2}^{h/2} f(w_k + \vartheta) d\vartheta. \quad (5)$$

The approximation derived in (4) follows by replacing the density inside each interval by its midpoint value. This exact identity allows theoretical or numerical comparison between grouped and ungrouped moments.

3.1 Cardioid distribution

The density function for a cardioid circular random variable is

$$f(\theta) = \frac{1}{2\pi} \{1 + 2\rho \cos(\theta - \mu)\}, \quad \forall \theta \in [0, 2\pi), \quad (6)$$

where $0 \leq \mu < 2\pi$ and $-\frac{1}{2} \leq \rho \leq \frac{1}{2}$. Its trigonometric moments are given by $\phi_1 = \rho \cdot e^{i\mu}$ and $\phi_\ell = 0$, for $\ell = \pm 2, \pm 3, \dots$

A direct calculation yields

$$\int_{-h/2}^{h/2} f(w_k + \vartheta) d\vartheta = \frac{h}{2\pi} + \frac{2\rho}{\pi} \cos(w_k - \mu) \sin(h/2).$$

Substituting this expression into (5) produces two terms:

$$\phi_{\ell}^* = \frac{h}{2\pi} \sum_{k=1}^K e^{i\ell w_k} + \frac{2\rho \sin(h/2)}{\pi} \sum_{k=1}^K e^{i\ell w_k} \cos(w_k - \mu).$$

The first sum corresponds to the uniform component of the density and vanishes because the midpoints w_k are equally spaced on the circle, so $\sum_{k=1}^K e^{i\ell w_k} = 0$, for $\ell = 1, \dots, K-1$. For the second term, we write $\cos(w_k - \mu) = \frac{1}{2}(e^{i(w_k - \mu)} + e^{-i(w_k - \mu)})$, giving

$$\sum_{k=1}^K e^{i\ell w_k} \cos(w_k - \mu) = \frac{1}{2} \left[e^{-i\mu} \sum_{k=1}^K e^{i(\ell+1)w_k} + e^{i\mu} \sum_{k=1}^K e^{i(\ell-1)w_k} \right].$$

When $\ell = 1$, the first sum again cancels, while the second reduces to $\sum_{k=1}^K 1 = K$, yielding

$$\phi_1^* = \rho e^{i\mu} \frac{\sin(h/2)}{h/2}.$$

Since $\phi_1 = \rho e^{i\mu}$ for the cardioid distribution, the grouping effect satisfies exactly (4).

Moreover, for $\ell = 2, 3, \dots, K-1$, both exponential sums vanish, so that $\phi_{\ell}^* = 0$. Hence, the grouped cardioid distribution preserves the property that all higher-order trigonometric moments beyond the first are zero.

3.2 Wrapped Cauchy distribution

If Θ follows a wrapped Cauchy distribution with parameters (μ, ρ) , where $0 \leq \rho < 1$, then its density is

$$f(\theta) = \frac{1 - \rho^2}{2\pi \{1 + \rho^2 - 2\rho \cos(\theta - \mu)\}}, \quad 0 \leq \theta < 2\pi.$$

The corresponding trigonometric moments are $\phi_{\ell} = \rho^{|\ell|} e^{i\ell \mu}$ for $\ell \in \mathbb{Z}$.

The wrapped Cauchy distribution is especially convenient because its cumulative distribution function admits a closed form; see Fisher (1993, Eq. 3.27). In particular,

$$F(\theta) = \frac{1}{2\pi} \cos^{-1} \left(\frac{(1 + \rho^2) \cos(\theta - \mu) - 2\rho}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \right), \quad 0 \leq \theta < 2\pi.$$

Equivalently,

$$F(\theta) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left[\frac{1 + \rho}{1 - \rho} \tan \left(\frac{\theta - \mu}{2} \right) \right] \quad 0 \leq \theta < 2\pi.$$

Therefore the grouped trigonometric moment can be evaluated exactly as $\phi_{\ell}^* = \sum_{k=1}^K e^{i\ell w_k} p_k$.

Thus, for fixed K , the grouped moment ϕ_ℓ^* is obtained exactly from a finite sum involving only evaluations of the wrapped Cauchy cumulative distribution function at the class boundaries. Unlike the Fourier-series approach used in the general setting, no infinite expansion or asymptotic approximation is required in this case.

Substituting the above expression for p_k gives an exact finite-sum representation for ϕ_ℓ^* , with no infinite Fourier-series approximation required. Thus, for the wrapped Cauchy family, grouped and ungrouped trigonometric moments can be compared exactly through the closed-form distribution function.

More generally, grouped and ungrouped moments can be compared analytically or evaluated numerically without approximation. Numerical experiments show the same qualitative behavior observed in other families: grouping attenuates the magnitude of trigonometric moments, while the correction compensates for most of this attenuation bias.

The above examples show that, for specific circular models, the effect of grouping on trigonometric moments can be measured exactly and confirms the sinc-type correction obtained in Sect. 2. However for most practical data sets that have been grouped, we either do not know the underlying model, nor may such explicit expressions for grouped trigonometric moments be available, and so the tools we propose here are most relevant, much like the Shappard's correction on the real line. We now therefore turn to a simulation study for some commonly used circular distributions.

4 Simulation results

To get a sense of what the practical implications are, we do several simulations. We consider three commonly used models for circular data namely: the von Mises distribution (vM), the Generalized von Mises distribution (gvM), and the Sine-skewed von Mises distribution (svM). We recall their densities below.

A circular random variable θ follows a von Mises (vM) distribution if its density function is of the form:

$$f(\theta | \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad \forall \theta \in [0, 2\pi), \quad (7)$$

where $0 \leq \mu < 2\pi$ and $\kappa \geq 0$ is the concentration parameter. $I_0(\kappa)$ in the normalizing constant is the modified Bessel function of the first kind and order zero.

The gvM density is given by

$$f(\theta | \mu_1, \mu_2, \kappa_1, \kappa_2) = \frac{1}{2\pi G_0(\delta, \kappa_1, \kappa_2)} \exp\{\kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos 2(\theta - \mu_2)\}, \quad (8)$$

where $\theta \in [0, 2\pi)$, $\mu_1 \in [0, 2\pi)$, $\mu_2 \in [0, \pi)$, $\delta = (\mu_1 - \mu_2) \pmod{\pi}$, and $\kappa_1, \kappa_2 \geq 0$. The $G_0(\delta, \kappa_1, \kappa_2)$ is a normalizing constant, see Gatto and Jammalamadaka (2007).

A general class of skewed circular distributions is proposed and discussed in Umbach and Jammalamadaka (2012). A particular case, called the Sine-skewed

von Mises (svM) distribution is considered in Abe and Pewsey (2011), with density given by

$$f(\theta|\mu, \kappa, \lambda) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)} \{1 + \lambda \sin(\theta - \mu)\}, \quad \forall \theta \in [0, 2\pi), \quad (9)$$

where $0 \leq \mu < 2\pi$, and $\kappa \geq 0$ is the concentration parameter. $\lambda \in [-1, 1]$ is a skewing parameter.

Samples of sizes $n = 30, 50, 70, 100, 150, 200$ are generated from each of these distributions and grouped into K sub-intervals, where $K = 6, 8, 10, 15$. The bias of the uncorrected ρ_ℓ^2 for $\ell = 1$ and 2 was compared to its corrected version, as given in (4), based on 1000 repeated samples. Six different values of the concentration parameter, $\kappa = 0.2, 0.5, 1, 2, 5, 10$, were considered for the von Mises distribution with a fixed mean direction, $\mu = 0$. For the samples from the Generalized von Mises distribution, $\kappa_1 = \kappa_2 = \kappa$ and $\mu_1 = \mu$, while $\mu_2 = \mu + \frac{\pi}{3}$. Finally for the svM, we take $\lambda = 0.5$. The results of the simulation study demonstrate a consistent superiority of the corrected trigonometric moments (i.e., meaning they tend to be closer to what one would get from the original data) compared to the uncorrected ones for the von Mises and skewed von Mises distributions. However, this was not the case for the Generalized von Mises distribution when $\ell = 2$, as illustrated in Fig. 1.

To complement Fig. 1, which illustrates representative scenarios, we also summarize results across several values of n and K for von Mises distribution in Fig. 2, authors have other results which are available to interested readers.

The simulation results indicate that, as the number of classes K increases, the difference between the biases of the corrected and uncorrected estimators diminishes, reflecting the reduced loss of information under finer grouping. For ρ_1^2 , this behavior is consistent across values of K once the sample size exceeds $n = 50$, whereas, for smaller samples, sampling variability may cause the corrected and uncorrected estimators to alternate in bias dominance. In contrast, for ρ_2^2 a stable pattern across values of K is observed for all sample sizes considered, indicating that grouping effects dominate sampling variability for higher-order moments in the scenarios examined. These conclusions were consistent across all combinations of sample sizes and class numbers examined in the simulation study except for gvM. For the Generalized von Mises distribution, the corrected estimator of ρ_2^2 occasionally exhibits larger bias than the uncorrected version. This behavior can be attributed to the strong second-harmonic structure of the distribution, for which grouping effects involve aliasing interactions among multiple Fourier components rather than simple attenuation. Since the correction formula assumes primarily multiplicative attenuation, it may overcompensate in such cases, leading to increased bias.

The degree of concentration of the data and its departure from uniformity play a crucial role in determining the necessity for correction. To assess the need for correction in samples from the three considered circular probability distributions, grouped circular data were generated for various sample sizes and numbers of sub-intervals K . Uniformity of the grouped counts was then evaluated using a Chi-square goodness-of-fit test against the discrete uniform distribution on the class intervals at the 0.05 significance level. The findings of the

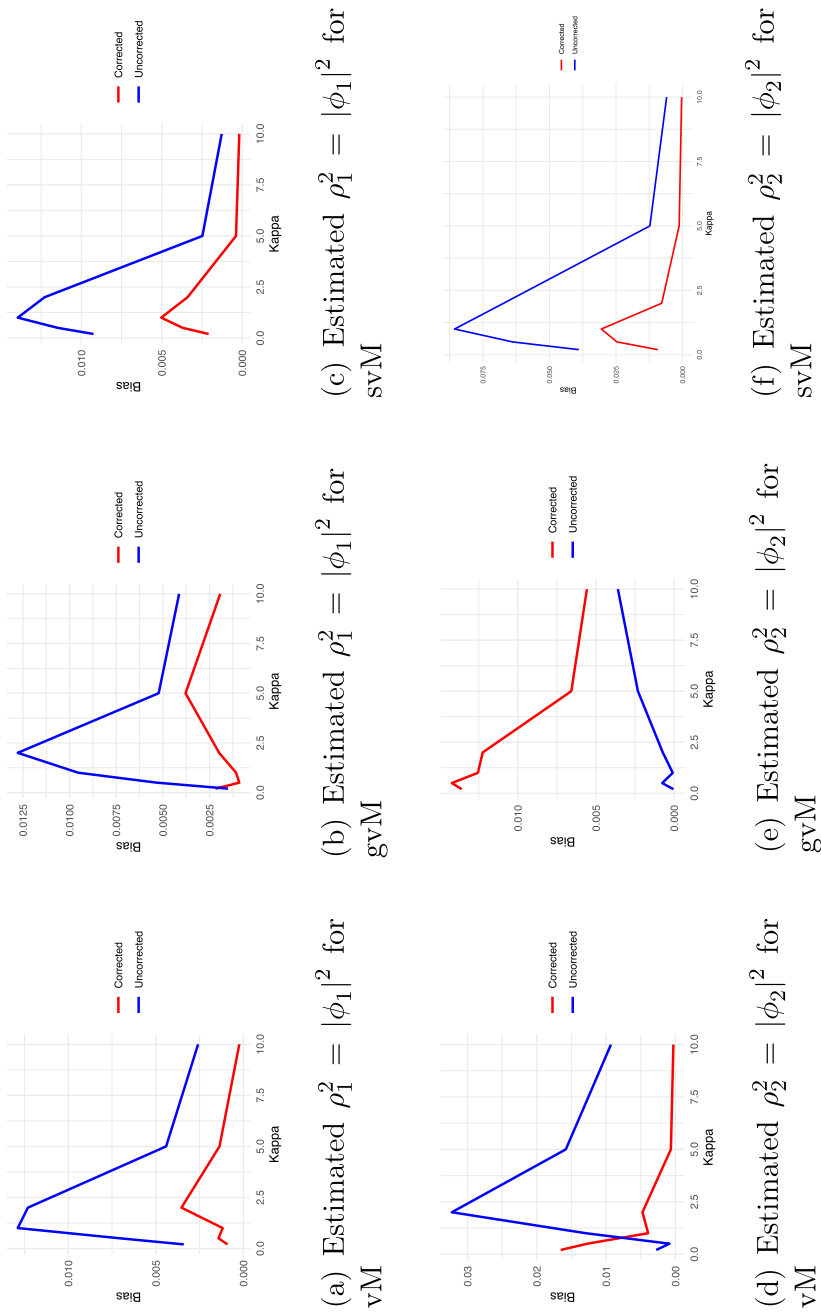


Fig. 1 Bias of estimated $\rho_1^2 = |\phi_1|^2$ and $\rho_2^2 = |\phi_2|^2$ for grouped data, $n = 100$ and $K = 6$

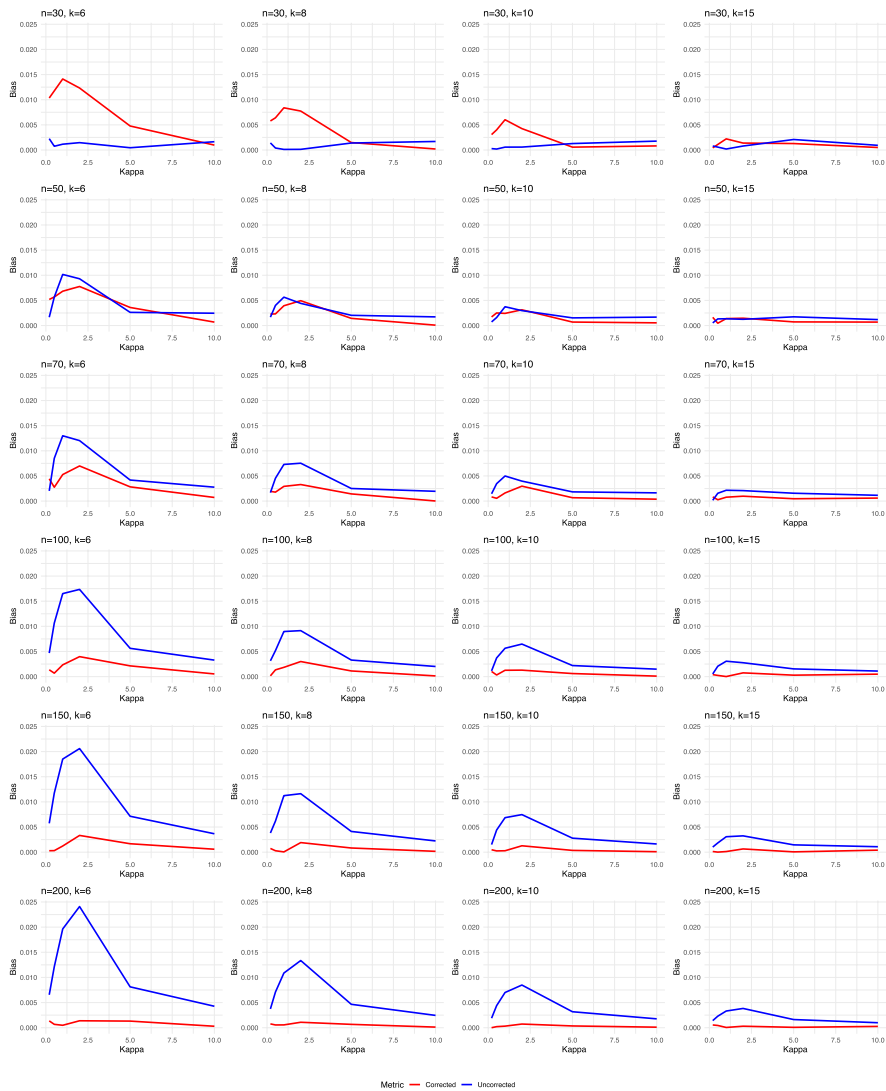


Fig. 2 Bias of estimated ρ_1^2 from von Mises distribution

simulation study indicate that as the concentration parameter (κ) increases, the proportion of uniformly distributed samples decreases. Samples from the svM distribution exhibit significant deviations from uniformity, particularly for larger sample sizes. Additionally, while samples from the gvM distribution with a greater number of modes tend to show increased uniformity, larger sample sizes lead to greater deviations from uniformity.

Consequently, when grouped data appear nearly uniform, grouping correction has little practical effect, whereas for clearly non-uniform distributions applying the correction helps improve recovery of the underlying trigonometric moments.

5 Application: cross-bedding Azimuths from the Kamthi formation

In a path-setting paper on paleo-current analysis, Sengupta and Rao (1966) investigated the change of direction of flow of the Kamthi river based on cross-bedding azimuth measurements. The grouped frequency data, reproduced in Table 1.3 of Jammalamadaka and SenGupta (2001), provide a classical example of circular data recorded in angular intervals. We use these data to illustrate the practical impact of grouping correction under different resolutions.

The original data were recorded in $K = 18$ equal class intervals of width $h = 20^\circ$ as shown in Fig. 3 a. Strong deviation from uniformity is evident from the Chi-square goodness-of-fit test ($\chi^2 = 109.24$, $df = 17$, $p = 1.67 \times 10^{-15}$) and the Rayleigh test of uniformity (test statistic = 0.5018, $p < 0.001$), confirming substantial directional structure.

The uncorrected squared resultant length was $R_g^2 = 146851.70$, while the corrected value was $R_c^2 = 147599.91$, representing an increase of only 0.506%. As predicted by the theoretical attenuation factor $\sin(h/2)/(h/2)$, the correction is negligible when the class width is small relative to the angular concentration.

To assess the effect of coarser resolution, the same data were regrouped into $K = 5$ classes of width $h = 72^\circ$ (Fig. 3, b). Non-uniformity remains highly significant ($\chi^2 = 67.041$, $df = 4$, and $p = 9.56 \times 10^{-14}$; Rayleigh test statistic = 0.6505 and $p < 0.001$).

In this case, the uncorrected squared resultant length $R_g^2 = 142334.5$ increased to $R_c^2 = 152149.8$ after applying the grouping correction, corresponding to an improvement of 6.45%. This change is substantial and directly reflects the attenuation predicted by the sinc-type factor when h is large.

These two cases, derived from the same underlying data, clearly demonstrate that the impact of grouping depends critically on angular resolution. When h is small, attenuation of the trigonometric moments is minimal and the correction has little practical effect. However, when grouping intervals are coarse, the reduction in resultant length is non-negligible and may materially influence inference regarding concentration and directional consistency.

6 Conclusion

This paper considers the very interesting practical problem of estimating the resultant length as well as the trigonometric moments from grouped circular data. We provide a simple but theoretically rigorous derivation for the correction formulae to address the approximation error introduced by grouping, which suggests that

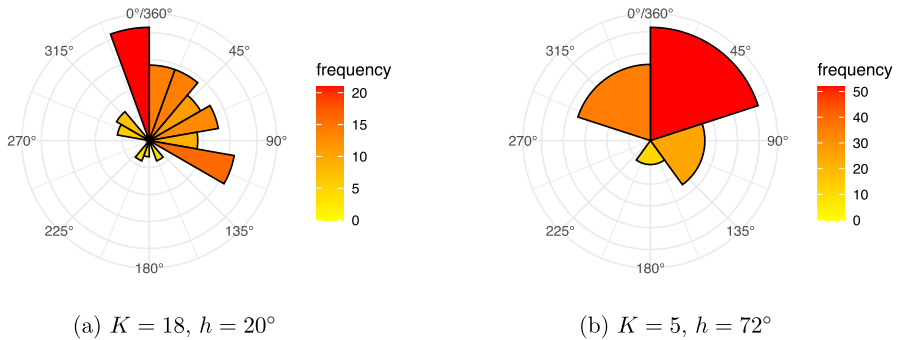


Fig. 3 Rose diagrams of the Lower Kamthi cross-bedding azimuths under two grouping schemes

aggregation generally attenuates the absolute value of trigonometric moments. Our simulation results, based on the von Mises, Generalized von Mises, and Sine-skewed von Mises distributions, show that the corrected moments yield more accurate estimates, particularly for data from non-uniform distributions.

A practical approach that we suggest for deciding whether or not a correction for grouping is desirable is as follows: since the data are often presented already pre-grouped in the form of a frequency table, we first suggest a preliminary test for uniformity using a simple Chi-square test on this frequency table. Since our theory as well as simulations show that a correction for grouping is desirable if this result indicates that the data is non-uniform, one can then decide whether a correction for grouping should be applied to the statistics such as the resultant length and/or the trigonometric moments. The effective correction is not a fixed value but depends on the coarseness of the data grouping. As one might expect, the larger the class width h , the greater the need for a correction. The Kamthi example shows that while grouping may have little impact under fine resolution, coarse angular discretizations can materially attenuate directional concentration, making the proposed correction essential for accurate inference. We believe that our work shed some light on when and how much grouping correction is desirable for researchers working with grouped circular data.

Author contributions S.R.J. and J.T. conceived the study and developed the theoretical framework and carried out the mathematical derivations. A.A. performed numerical simulations and conducted the data analysis and visualization. All authors contributed to interpreting the results, writing the manuscript, and approving the final version of the paper.

Data availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative

Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abe T, Pewsey A (2011) Sine-skewed circular distributions. *Stat Pap* 52(3):683–707
- Carslaw HS (1930) *Introduction to the Theory of Fourier's Series and Integrals*. Dover, (third edition)
- Fisher NI (1993) *Statistical Analysis of Circular Data*. Cambridge University Press
- Gatto R, Jammalamadaka S (2007) The generalized von mises distribution. *Statistical Methodology* 4(3):341–353
- Greenwood JA (1959) Corrections to trigonometric moments for grouping. Statistical Laboratory. Iowa State College, (1.11 Technical Report)
- Jammalamadaka S, SenGupta A (2001) *Topics in Circular Statistics*. World Scientific
- Kendall, M. G. and Stuart, A. (1969). *The advanced theory of statistics*, volume 1. Griffin, London, 3rd edition
- Mardia KV (1972) *Statistics of Directional Data*. Academic Press, New York
- Schneeweiss H, Komlos J, Ahmad AS (2010) Symmetric and asymmetric rounding: a review and some new results. *AStA Adv Stat Anal* 94(3):247–271
- Sengupta S, Rao JS (1966) Statistical analysis of crossbedding azimuths from the Kamthi formation around Bheemaram, Pranhita Godavari valley. *Sankhya* 28:165–174
- Umbach D, Jammalamadaka S (2012) On introducing asymmetry into circular distributions. *Pak J Stat Oper Res* 8(3):531–535
- Vardeman SB (2005) Sheppard's correction for variances and the "quantization noise model". *IEEE Trans Instrum Meas* 54(5):2117–2119

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

S. Rao Jammalamadaka¹ · György Terdik² · Ali Abuzaid¹

✉ Ali Abuzaid
aliabuzaid@ucsb.edu

S. Rao Jammalamadaka
rao@pstat.ucsb.edu

György Terdik
Terdik.Gyorgy@inf.unideb.hu

¹ Department of Statistics and Applied Probability, University of California, Santa Barbara, Santa Barbara, CA, USA

² Faculty of Informatics, University of Debrecen, Debrecen, Hungary