

# Joint regression modelling of intensity and timing of accelerometer counts

Marco Geraci 

MEMOTEF Department, Sapienza  
University of Rome, Rome, Italy

## Correspondence

Marco Geraci, MEMOTEF Department,  
Sapienza University, via del Castro  
Laurenziano 9, 00161 Rome, Italy.  
Email: [marco.geraci@uniroma1.it](mailto:marco.geraci@uniroma1.it)

Accelerometers are commonly used in human medical and public health research to measure physical movement, which is relevant in a wide range of studies, from physical activity and sleep behaviours studies, to identification of movement patterns in people affected by diseases of the locomotor system and prediction of risk of injury in high performance sports. The accelerometer output provides the intensity (activity count) and timing (timestamp) of the movement, which can be used to define bouts of activity (periods of sustained movement of a given intensity). In some contexts, it may be important to include both dimensions to obtain a broader and deeper understanding of the phenomenon under study. Such is the case of a large-scale epidemiological investigation on the daily and weekly physical activity behaviours of school-aged children enrolled in the UK Millennium Cohort Study, which has motivated the present article. I present a statistical approach to joint modelling of intensity and timing of activity bouts that takes advantage of the circular nature of the timing. The model, which accounts for the longitudinal structure of the observations, is remarkably simple to implement using standard statistical software.

## KEYWORDS

circadian rhythm, directional statistics, projected normal, von Mises

## 1 | INTRODUCTION

Accelerometers are portable devices capable of providing an objective measure of the intensity and timing of physical movement. Accelerometer-based measurements can be used to identify activities that involve minimal or no movement and low energy expenditure (sedentary behaviours) or, at the other extreme, energetic physical movements. Medical and epidemiological applications of accelerometers range from the study of physical activity and sleep behaviours,<sup>1,2</sup> to the identification of movement patterns in people affected by diseases of the locomotor system<sup>3,4</sup> and prediction of risk of injury in high performance sports.<sup>5,6</sup>

Nowadays it is not uncommon to design and implement large-scale accelerometer-based studies thanks to the development of devices that are smaller in size and less clumsy to wear, yet more capable in terms of memory and battery, acceleration range, and often at a lower cost than in the past.<sup>7</sup> The Millennium Cohort Study<sup>8</sup> (MCS) and Biobank<sup>9</sup> in the United Kingdom (UK), and the National Health and Nutrition Examination Survey<sup>10</sup> in the United States, among others, provide large-scale accelerometer data sources for hundreds of individuals.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Author. *Statistics in Medicine* published by John Wiley & Sons Ltd.

In the present article, I considered accelerometer data from the MCS that were collected to monitor children's physical activity in free-living conditions. There are a number of studies, both substantive<sup>11-13</sup> and methodological,<sup>14-16</sup> that investigated physical activity behaviours using the MCS accelerometer data. Most of these studies, like others in the field of exercise science, focus on the intensity and duration of physical activities, either regardless of the time when such activities take place or by crudely stratifying the analysis based on broad time intervals. A few exceptions are given by studies where movement is treated as a functional to characterize the intra-day temporal variations in physical activity intensity.<sup>16,17</sup> Notwithstanding the valuable information provided by these studies, here I take a different stance by treating time as *random*, rather than fixed. Arguably, this perspective is more compatible with the characterization of unstructured behaviours of children in free-living conditions. By leveraging on the *circular* nature of time around the clock, I develop a joint regression model for both intensity (linear outcome) and timing (circular outcome) of physical activity via the polar coordinates representation of linear normal models. In the literature of directional statistics,<sup>18,19</sup> a bivariate random variable with a circular and a linear components is known as *cylindrical*.<sup>20</sup>

The representation of linear random variables via polar coordinates (radius and angle) is not new in statistical analysis. Some applications can be found in imaging,<sup>21</sup> genetics<sup>22</sup> and sports science,<sup>23</sup> to mention a few. What makes the use of polar coordinates worth some introductory remarks is the presence of a circular data component (ie, the angle). Circular data are data that can be represented as points on the circumference of the unit circle. Examples of circular variables include movement direction of animals in ecology, protein structure in biology, neuronal activity and circadian rhythm in physiology, human sense of direction and visual perception of space in cognitive psychology, as well as variables related to time of the day or the year. In general, variables that have a periodic nature fall into this category. Statistical methods for data measured using linear metrics (eg, Euclidean) do not apply to circular data. The reason is due to the periodic nature of this type of data: while linear data have a meaningful zero, circular data do not. On the circle, measurements at 0° and 360° represent the same direction whereas on a linear scale they would be located at opposite ends of the scale. Importantly, distances on the circle require trigonometric rather than Euclidean calculations. On the one hand, this means that the analyst must be careful with the calculation of even the most basic statistics such as a sample mean or standard deviation as their formulas differ from those applied to linear random variables.<sup>20</sup> On the other hand, the application of circular statistics is able to provide quite unique information about the distribution of events that cyclically recur over space or time.

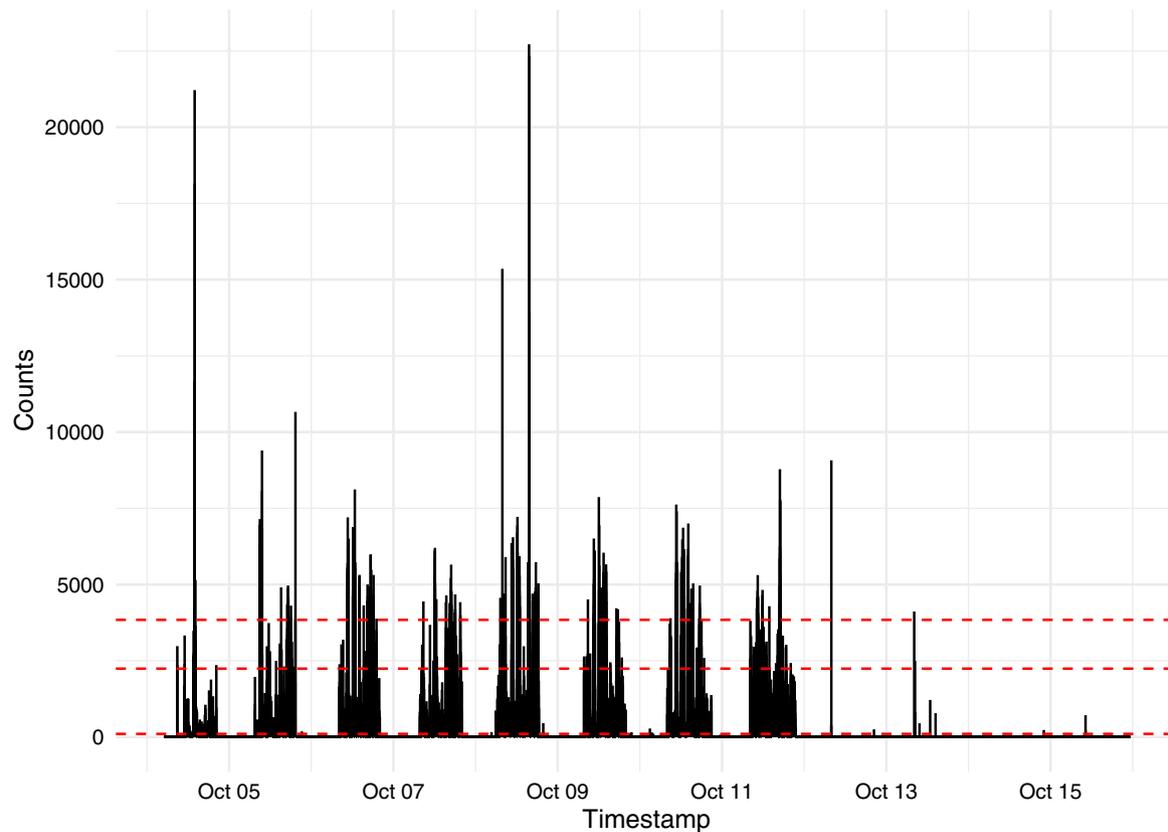
In the next section, I introduce the MCS accelerometer data in detail. In Section 3, I illustrate the models and related inferential issues, while the results of the MCS data analysis are given in Section 4. I conclude with some remarks in Section 5. Supplemental materials for this article, including illustrative R code, are available as an online supplement.

## 2 | THE DATA

The Millennium Cohort Study (MCS) is a UK-wide longitudinal multi-purpose survey that started in 2000–2001 and has been following a large cohort of individuals since their infancy.<sup>24</sup> The first survey was conducted when the children were aged 9 months old, followed by several surveys at an average of 3-year intervals. Physical activity monitoring<sup>8</sup> was carried out at age 7 (more precisely, between ages 6.3 and 8.4). Accelerometers were distributed between May 2008 and August 2009 by post. Children were asked to wear an ActiGraph GT1M accelerometer (ActiGraph, Florida, USA) attached to an elastic belt for seven consecutive days during all waking hours.

The time series plot of accelerometer data collected for one child of the MCS is shown in Figure 1. Accelerometers produce a dimensionless output known as 'acceleration counts' which is proportional to the intensity of the physical movement, along with a timestamp. These counts, aggregated over relatively short consecutive intervals called *epochs*, are typically classified into categories of varying levels of intensity, with sedentary behaviour on one end of the spectrum, followed by light, moderate and vigorous activity. Classification thresholds obtained from a calibration study in 7-year old children<sup>25</sup> are shown in Figure 1.

The MCS accelerometer files were processed following a number of data quality procedures using the package *pawacc*<sup>26</sup> for the statistical environment R.<sup>27</sup> Since full details are given elsewhere,<sup>15,28</sup> here I report selected processing criteria. Non-wear time status was defined as 20 minutes or more of consecutive zero-counts on an epoch-by-epoch basis, while accelerometer counts were classified using the cut-offs reported in Figure 1. In addition, a small number of extreme counts were removed before data aggregation.<sup>29</sup> Only participants contributing with at least seven valid unique days of the week were retained in the dataset, where 'valid day' here is defined as a day including at least

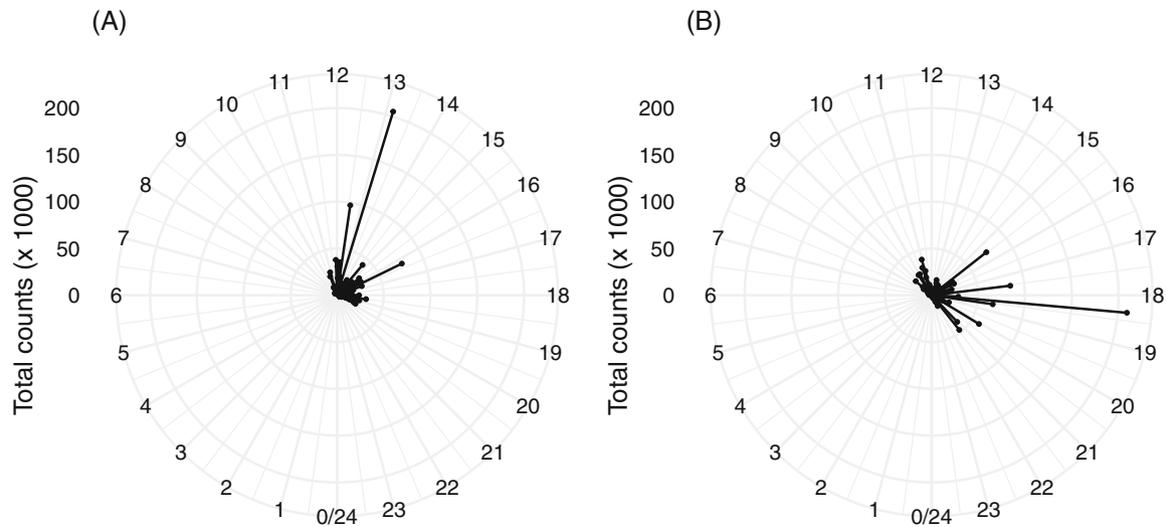


**FIGURE 1** Plot of accelerometer counts aggregated in 1-min epochs for one child of the UK Millennium Cohort Study. Red dashed lines mark thresholds for sedentary behaviour (0–100 counts), light (100–2241 counts), moderate (2241–3841 counts) and vigorous activity (> 3841 counts).

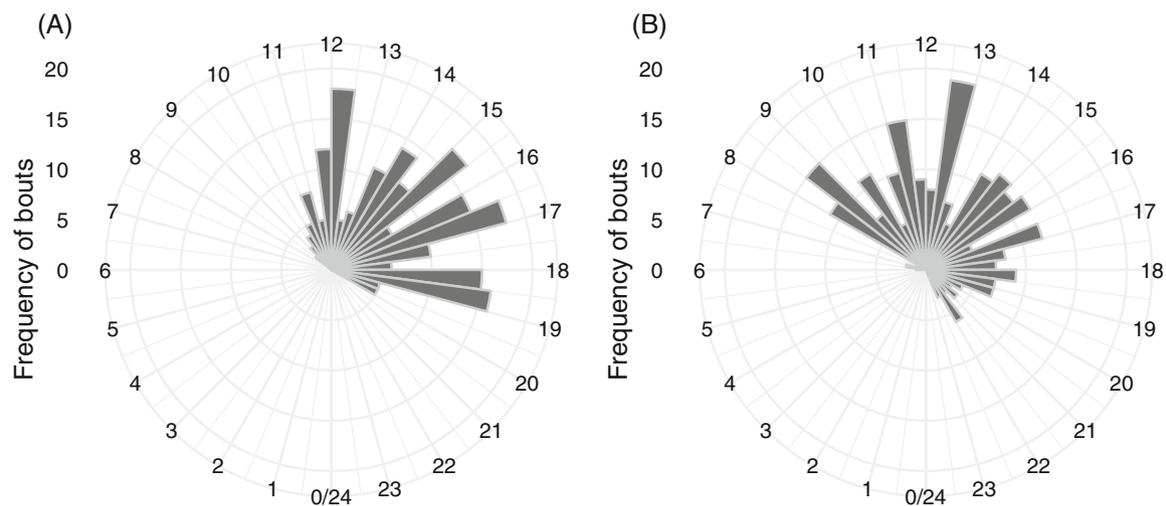
ten hours of wear time. This reliability threshold was determined in a previous study using the same data.<sup>30</sup> The analyses presented in this article are restricted to moderate-to-vigorous physical activity (MVPA) bouts from 1873 singleton children who provided valid data for seven consecutive days in the time period going from May 2008 through June 2009.

More formally, let  $r(t_h)$  denote the variable ‘activity counts’ that is measured by the accelerometer at time  $t_h$  over a 24-hour clock,  $h = 1, \dots, H$ . For the purposes of this analysis, I consider epochs of 1 minute (while raw data are available at 15-second epochs), therefore each child provides at most  $H = 24 \times 60 = 1440$  measurements daily. Measurements are subsequently classified into MVPA if  $r(t_h) \geq 2241$ . Let  $t^{(1)}$  and  $t^{(2)}$  be the start and end of a sequence of MVPA measurements and let  $r = \sum_{h:t^{(1)} \leq t_h \leq t^{(2)}} r(t_h)$  be the total activity counts for such a bout. It should be stressed that, before observation,  $r$ ,  $t^{(1)}$  and  $t^{(2)}$  are random variables. Moreover,  $t^{(1)}$  and  $t^{(2)}$  are circular. This represents a key distinctive aspect as compared to other approaches like, for example, functional data analysis<sup>16</sup> where  $r(t_h)$  is treated as an observation from a functional defined over the *fixed* and *linear* temporal dimension. Given that, in general, MVPA bouts have short durations (for the MCS data, the median was 1 minute while the 99th centile was 11 minutes), the timing of a bout is from now on defined to be an intermediate time point, say  $t$ , between  $t^{(1)}$  and  $t^{(2)}$ . This is illustrated in Figure 2 where bouts of MVPA of two children in the dataset are plotted on the circular scale. Basically, the length of each segment is proportional to the intensity of the bout, while its direction, determined by the angle  $\theta = t \cdot \pi/12$ , defines the timing of the bout. The distribution of these bouts over the clock is depicted in Figure 3 using a circular histogram.

Overall, the resulting dataset consists of 347,430 bouts of MVPA, corresponding to an average of 185 bouts per child over the 7-day long monitoring period. It should be noted that the pre-processing of the MCS accelerometer data<sup>8</sup> indicated that some children wore the accelerometer not only during waking hours (as instructed) but also during night-time. An inspection of the distribution of MVPA bouts by hour of the day (Supplemental Figure S1) revealed a very small number (552) of bouts between 23:00 and 6:00. Since truncating the day to any time interval would be arbitrary and given that



**FIGURE 2** Bouts of moderate-to-vigorous activity for (A) a girl and (B) a boy of the Millennium Cohort Study over 7 days. The length of each segment is proportional to the intensity of the bout, while its direction is determined by the angle that corresponds to the midpoint of the bout interval.



**FIGURE 3** Circular histograms of bouts of moderate-to-vigorous activity for (A) a girl and (B) a boy of the Millennium Cohort Study over 7 days.

the proportion (0.16%) of these unexpected nocturnal measurements was limited anyway, no temporal restrictions were made to the dataset.

Some descriptive statistics are given in Table 1, namely the average number of bouts per child-hour, average bout duration and average counts per bout, stratified by sex, time of the day (AM, 0:00–12:00; PM, 12:00–24:00), and day of the week. There is clear evidence that on average boys engage in more MVPA than girls in terms of frequency, duration, and intensity of the bouts, regardless of time of the day and day of the week. When stratified by time of the day, the picture is more complex. MVPA bouts are more frequent, tend to last longer, and are more intense (ie, with higher average counts) in the afternoons, for either boys or girls. Such an ‘afternoon effect’ does not seem to differ substantially by day of the week. Of course, there are potentially several other variables that may explain differences in these outcomes. For example, the data indicate that average MVPA intensity is notably stronger in the spring. Other covariates obtained from the main MCS survey and included in the present analysis are UK country (England, Wales, Scotland and Northern Ireland) in which the child was living, age (years) at the time of physical activity monitoring, ethnicity (White or other than White), mode of transport to and from school (active or passive), waist circumference (cm), and body fat (%).

**TABLE 1** Average frequency (number of bouts per child-hour), average bout duration (minutes), and average intensity (activity counts per bout) by sex, time of the day (AM, 0:00–12:00; PM, 12:00–24:00), and day of the week (weekday, Monday–Friday; weekend, Saturday–Sunday) for moderate-to-vigorous physical activity (MVPA) in the Millennium Cohort Study accelerometer data

Sex	Time of the day	Day of the week	MVPA bouts		
			Frequency	Duration	Intensity
Male	AM	Weekday	0.61	2.05	6756.00
Female			0.44	1.86	6176.03
Male	PM		1.05	2.14	7265.27
Female			0.81	1.81	6249.34
Male	AM	Weekend	0.65	2.01	6765.03
Female			0.42	1.68	5768.09
Male	PM		1.01	2.07	7114.12
Female			0.76	1.79	6251.01

### 3 | METHODS

In this section, I will first lay the groundwork by introducing the probability distribution of interest. Then, I will propose an approach to the joint modelling of intensity and timing of activity bouts, followed by details on inference. It is helpful to introduce some notation that will be used throughout. The  $n$ -dimensional vectors of zeros and ones will be denoted by  $\mathbf{0}_n$  and  $\mathbf{1}_n$ , respectively;  $\|\mathbf{u}\|$  will denote the Euclidean norm of the vector  $\mathbf{u}$  and  $\mathbf{u}^a = (u_1^a, \dots, u_n^a)^\top$ , for  $a \in \mathbb{R}$ , will represent the element-wise exponentiation of a vector;  $\mathbf{I}_n$  will denote the  $n \times n$  identity matrix,  $\mathbf{J}_n = \mathbf{1}_n \mathbf{1}_n^\top$  the  $n \times n$  matrix of ones, and  $\text{diag}\{\mathbf{u}\}$  a diagonal matrix whose diagonal elements are the elements of the vector  $\mathbf{u}$ ; and  $\Pi = (-\pi, \pi]$  will denote the circle. The function  $\text{atan2} : \mathbb{R}_{\neq 0} \times \mathbb{R}_{\neq 0} \rightarrow \Pi$  will denote the arctangent function restricted to the interval  $\Pi$  and  $I(\cdot)$  will denote the indicator function, which is equal to 1 if its argument is true or 0 otherwise. The sample circular mean and standard deviation for an angular variable will be calculated using standard formulas.<sup>20</sup> Finally, the symbol  $\otimes$  will denote the Kronecker product while  $\text{vec}(\mathbf{A})$  the vectorization of the matrix  $\mathbf{A}$  which stacks its columns into a single vector.

#### 3.1 | Directional data modelling

Let  $\mathbf{y} \in \mathbb{R}^2$  be a bivariate normal random variable with mean  $\mathbf{m}$  and covariance  $\Sigma$ ,  $\mathbf{y} \sim \mathcal{N}_2(\mathbf{m}, \Sigma)$ . Then the unit vector  $\mathbf{z} = \mathbf{y}/\|\mathbf{y}\|$  is a random point on the circle. Let  $(r, \theta) \in \mathbb{R}_{\geq 0} \times \Pi$  denote the polar coordinates of  $\mathbf{y}$ . The joint density of the length (or radius)  $r$  and angle  $\theta$  is given by

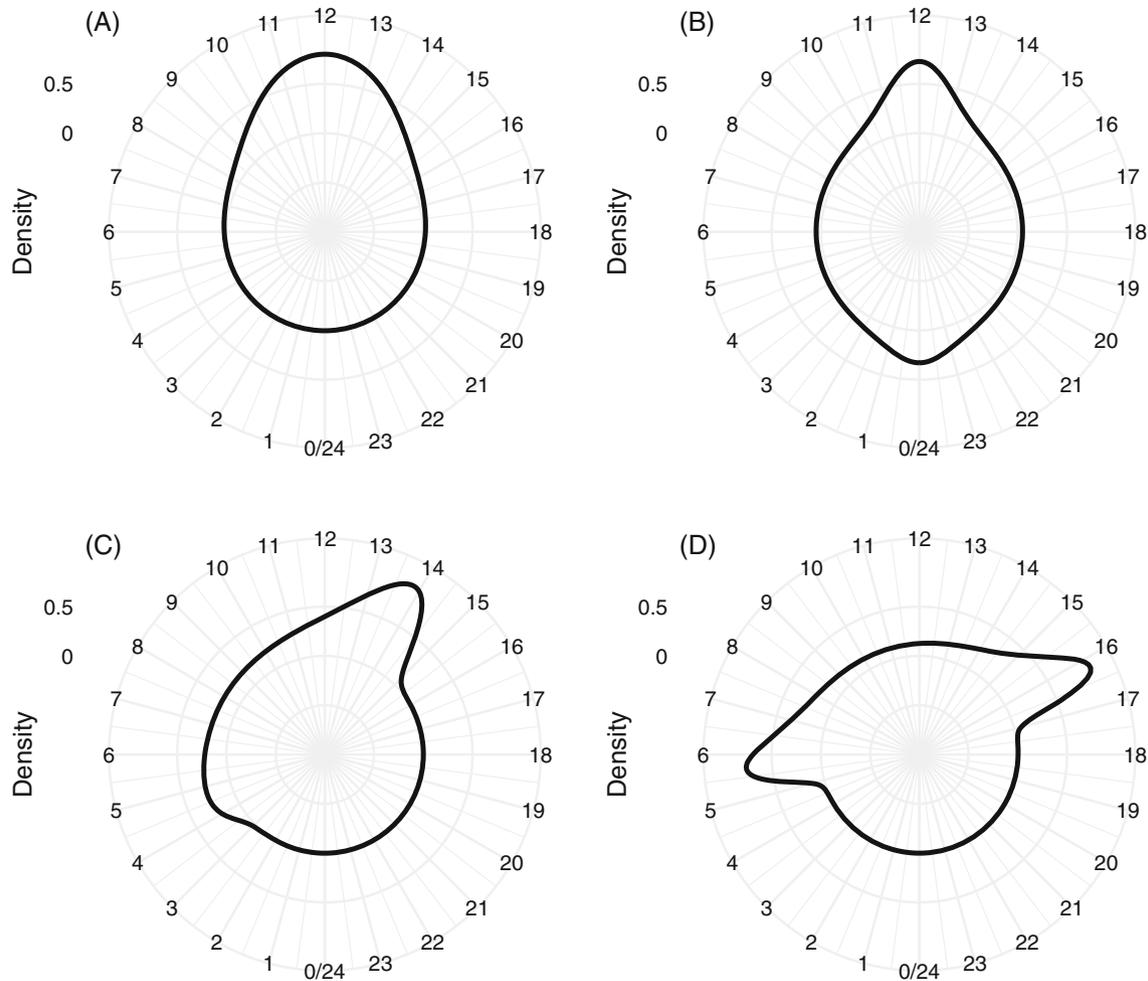
$$f(r, \theta) = \frac{r}{2\pi} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (r\mathbf{z} - \mathbf{m})^\top \Sigma^{-1} (r\mathbf{z} - \mathbf{m}) \right\}, \quad (1)$$

where  $\mathbf{z} = (\cos \theta, \sin \theta)^\top$  is the direction associated with  $\theta$ .

The marginal distribution  $f(\theta) = \int_0^\infty f(r, \theta) dr$  that results from (1) is of particular interest in directional data modelling. The distribution of the angle  $\theta$ , which is called general projected normal (GPN) distribution, has the analytical form

$$f(\theta) = \frac{\exp\left(-\frac{1}{2} \mathbf{m}^\top \Sigma^{-1} \mathbf{m}\right)}{2\pi (\mathbf{z}^\top \Sigma^{-1} \mathbf{z}) |\Sigma|^{\frac{1}{2}}} \left[ 1 + \frac{q\Phi(q)}{\phi(q)} \right], \quad (2)$$

where  $\Phi(u) = \int_{-\infty}^u \phi(s) ds$  is the cumulative distribution function of the standard normal and  $q = \mathbf{z}^\top \Sigma^{-1} \mathbf{m} / \sqrt{\mathbf{z}^\top \Sigma^{-1} \mathbf{z}}$ . Applications of the GPN in directional statistics with either cross-sectional or longitudinal data span numerous fields including biology (eg, protein structure), zoology (eg, animal migration), environmental epidemiology (eg, wind direction), and radiology (eg, imaging).<sup>31–35</sup>



**FIGURE 4** Density of the general projected normal with (A)  $\mathbf{m} = (-2, 0)^\top$  and isotropic, diagonal  $\Sigma$ ; (B)  $\mathbf{m} = (-1, 0)^\top$  and anisotropic, diagonal  $\Sigma$ ; (C)  $\mathbf{m} = (-1, 0)^\top$  and isotropic, non-diagonal  $\Sigma$ ; and (D)  $\mathbf{m} = (-1, 0)^\top$  and anisotropic, non-diagonal  $\Sigma$ .

It may be instructive to see how the GPN is relevant in directional statistics by considering the *conditional* distribution  $f(\theta|r=1)$ . Set  $\mathbf{m} = (\cos \mu, \sin \mu)^\top$  and  $\Sigma = \sigma^2 \mathbf{I}_2$ , where  $\sigma > 0$ . Then the density in (1) is given by

$$\begin{aligned} f(r, \theta) &= \frac{r}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2\sigma^2} (r^2 \mathbf{z}^\top \mathbf{z} - 2r\mathbf{z}^\top \mathbf{m} + \mathbf{m}^\top \mathbf{m}) \right\} \\ &= \frac{r}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2\sigma^2} (r^2 + 1) \right\} \exp \left\{ \frac{r \cos(\theta - \mu)}{\sigma^2} \right\}, \end{aligned}$$

since  $\mathbf{z}^\top \mathbf{z} = \mathbf{m}^\top \mathbf{m} = 1$  and  $\mathbf{z}^\top \mathbf{m} = \cos(\theta - \mu)$ . It follows that

$$f(\theta|r=1) = \frac{(2\pi\sigma^2)^{-1} \exp\{-\sigma^{-2}\} \exp\{\sigma^{-2} \cos(\theta - \mu)\}}{\sigma^{-2} \exp\{-\sigma^{-2}\} I_0(\sigma^{-2})} = \frac{\exp\{\sigma^{-2} \cos(\theta - \mu)\}}{2\pi I_0(\sigma^{-2})}. \quad (3)$$

where  $I_0$  denotes the modified Bessel function of order 0. That is, the angle  $\theta$  obtained from the polar coordinates of a bivariate isotropic normal distribution, conditional on the unit length, follows a von Mises distribution (with location parameter  $\mu$  and concentration parameter  $\sigma^{-2}$ ). The von Mises distribution is very popular in the literature of directional statistics,<sup>18-20</sup> though it presents some limitations when it comes to its flexibility. While it encompasses shapes ranging from the uniform ( $\sigma \rightarrow \infty$ ) to the normal ( $\sigma \rightarrow 0$ ), the von Mises distribution is always unimodal and symmetric. In contrast, the GPN distribution allows for asymmetry and bimodality.<sup>31,34</sup> The modelling flexibility of the GPN comes at a cost, however, since its shape depends on both its parameters  $\mathbf{m}$  and  $\Sigma$  (see Figure 4), making difficult their interpretation when taken individually. This may not be seen as a limitation if the main goal of the analysis is prediction. In any case, ad hoc solutions to overcome this drawback have been proposed.<sup>36</sup>

While the distribution of angles is naturally the main focus of analysis in directional statistics, it is worth noting that one can also obtain the marginal density  $f(r) = \int_{-\pi}^{\pi} f(r, \theta) d\theta$  from the joint distribution (1). Generally the radial distance of a data point is not of direct interest and thus it is not measured, or it may not even be conceptually easy to define (eg, as in the modelling of wind direction). Therefore, estimation algorithms for fitting (1) assume that  $r$  is latent,<sup>37,38</sup> which brings about an issue of identifiability of the scale of the GPN as a side-effect.<sup>34</sup> In the present article, the activity counts represent the magnitude associated with the vector  $\mathbf{z}$ . Thus it is perfectly reasonable to derive  $f(r)$ , whose analytical form may be available in specific cases (eg, see the Rayleigh, Rice and Hoyt distributions). For example, for  $\mathbf{m} = \mathbf{0}_2$  and  $\Sigma = \sigma^2 \mathbf{I}_2$ , the marginal distribution of  $r$  is Rayleigh (or Rice if  $\mathbf{m} \neq \mathbf{0}$ ) with scale parameter  $\sigma$  (which governs the mean and variance of Rayleigh and Rice random variables). The Rayleigh and Rice distributions are right-skewed on the positive real line, which is apposite to the application of the present article (see Supplemental Figure S2 for the histogram depicting the distribution of activity counts in the MCS physical activity data). The gamma distribution, which in turn is tied to the Rayleigh and Rice distributions, has been proposed<sup>39</sup> to deal with the skewness of physical activity outcomes. The case for  $\mathbf{m} \neq \mathbf{0}$  and general  $\Sigma$  does not seem to lead to a closed form for  $f(r)$ , though the latter can be obtained by means of analytical approximations<sup>40</sup> and, more in general, numerical integration.

I conclude this section by noting that in physical activity studies, it may also be of interest to contrast the intensity at different times of the day by examining the conditional distribution

$$f(r|\theta) = \frac{\mathbf{z}^T \Sigma^{-1} \mathbf{z}}{\sqrt{2\pi} \{\phi(q) + q\Phi(q)\}} r \exp\left(-\frac{1}{2} \mathbf{z}^T \Sigma^{-1} \mathbf{z}\right) \left(r - \frac{\mathbf{z}^T \Sigma^{-1} \mathbf{m}}{\mathbf{z}^T \Sigma^{-1} \mathbf{z}}\right)^2,$$

where  $q$  has been introduced in (2), or, more summarily, the expected value

$$E(r|\theta) = \frac{\mathbf{z}^T \Sigma^{-1} \mathbf{m}}{\mathbf{z}^T \Sigma^{-1} \mathbf{z}} + \frac{\Phi(q)}{\sqrt{\mathbf{z}^T \Sigma^{-1} \mathbf{z}} \{\phi(q) + q\Phi(q)\}}. \quad (4)$$

The second moment  $E(r^2|\theta)$ , provided elsewhere,<sup>38</sup> can be used to derive variance and standard deviation.

### 3.2 | Accelerometer counts

In my application,  $\theta$  measures the timing of an activity bout, while  $r$  represents the intensity of the bout acting as a (random) “scale” that multiplies the unit vector  $\mathbf{z}$ . Let  $r$  denote the total activity counts accumulated during a bout of activity. Also, let the angles  $\theta^{(1)} = t^{(1)} \cdot \pi/12$  and  $\theta^{(2)} = t^{(2)} \cdot \pi/12$  measure the time (in radians) of the start and end, respectively, of such a bout. The timing of the activity bout is represented by the circular mean of these angles, that is,  $\theta = \text{atan2}(\bar{s}, \bar{c})$ , with  $\bar{c} = (\cos \theta^{(1)} + \cos \theta^{(2)})/2$  and  $\bar{s} = (\sin \theta^{(1)} + \sin \theta^{(2)})/2$ . I assume that the response  $(r, \theta)$  has density as specified in (1), with parameters  $\mathbf{m}$  and  $\Sigma$ . Furthermore, I assume that  $\mathbf{m} \equiv \mathbf{m}(\mathbf{x})$  and  $\Sigma \equiv \Sigma(\mathbf{w})$  are functions of two vectors of covariates  $\mathbf{x}$  and  $\mathbf{w}$  of lengths  $p$  and  $q$ , respectively (only the vector  $\mathbf{x}$  may contain 1 as first element to account for the intercept). These vectors may include subject characteristics (eg, sex and ethnicity) as well as variables associated with the bouts (eg, morning or afternoon, day of the week, season). It is worth stressing that  $\mathbf{w}$  is not necessarily a subset of  $\mathbf{x}$ . The dependence of  $\Sigma$  on  $\mathbf{w}$  is operationalized via a (strictly positive) variance function  $g(\cdot; \delta) : \mathbb{R}^q \rightarrow \mathbb{R}_{>0}^2$  that depends on a parameter vector  $\delta$  and returns a vector of diagonal elements. The off-diagonal elements of  $\Sigma$  may be defined by a correlation function, say  $h(\rho)$  that depends on a scalar parameter  $\rho$ . For more details about possible models for  $g$  and  $h$ , the reader is referred to Section 3.3.

Since observations are clustered within individuals, I consider two alternative model specifications: marginal and conditional. In the former, the covariance (hence correlation) between subject-specific activity bouts is modelled via the error distribution’s parameters. In the latter, random effects induce such a correlation. There are advantages and disadvantages of one specification over the other and the reader is referred to the longitudinal data analysis literature for in-depth discussions.<sup>41-43</sup> Here, the main argument for adopting a conditional approach is the ensuing availability of subject-specific estimates. In contrast, the marginal approach is computationally less demanding.

#### 3.2.1 | Marginal approach

The model for a single bout can be written in terms of  $\mathbf{y} \equiv r\mathbf{z} = (r \cos \theta, r \sin \theta)^\top$  as

$$\mathbf{y} = \mathbf{B}^\top \mathbf{x} + \mathbf{e}, \quad \mathbf{e} \sim \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Sigma}) \quad (5)$$

where  $\mathbf{B}$  is a  $p \times 2$  matrix of regression coefficients, and  $\boldsymbol{\Sigma}$  is  $2 \times 2$ . I assume that  $\boldsymbol{\Sigma}$  is a covariance matrix with diagonal elements  $\sigma^2 \mathbf{g}^2$ , where  $\sigma > 0$  is the error scale and  $\mathbf{g} \equiv g(\mathbf{w}, \boldsymbol{\delta})$ . In this heteroskedastic model, the variance function  $g$  links  $r$  to the covariates  $\mathbf{w}$ . This is tantamount to a regression model (on the multiplicative scale) for the activity counts with regression parameter  $\boldsymbol{\delta}$ . The off-diagonal elements of  $\boldsymbol{\Sigma}$  control the correlation between the rectangular coordinates in  $\mathbf{y}$  (see discussion further below).

Now, suppose that for an individual we have multiple, say  $n$ , bouts of activity, with responses  $\mathbf{y}_j = (r_j \cos \theta_j, r_j \sin \theta_j)^\top$  or, equivalently,  $(r_j, \theta_j)$ ,  $j = 1, \dots, n$ . The vectors of covariates  $\mathbf{x}_j^\top$  and  $\mathbf{w}_j^\top$  are defined to be the rows of the design matrices  $\mathbf{X}$  and  $\mathbf{W}$ , respectively. For each individual, the observables are then given by the  $2n$ -long vector of responses  $\tilde{\mathbf{y}} = (r_1 \cos \theta_1, \dots, r_n \cos \theta_n, r_1 \sin \theta_1, \dots, r_n \sin \theta_n)^\top$ , the  $2n \times 2p$  design matrix  $\tilde{\mathbf{X}} = \mathbf{I}_2 \otimes \mathbf{X}$  and the  $n \times q$  design matrix  $\mathbf{W}$ . The proposed model becomes

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\mathbf{b} + \tilde{\mathbf{e}}, \quad \tilde{\mathbf{e}} \sim \mathcal{N}_{2n}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}) \quad (6)$$

where  $\mathbf{b} = \text{vec}(\mathbf{B})$ . The diagonal of the  $2n \times 2n$  covariance matrix  $\tilde{\boldsymbol{\Sigma}}$  is given by  $\sigma^2 \mathbf{1}_2 \otimes \mathbf{g}^2$ , where  $\mathbf{g} \equiv g(\text{vec}(\mathbf{W}), \boldsymbol{\delta})$  and  $g(\cdot; \boldsymbol{\delta}) : \mathbb{R}^{nq} \rightarrow \mathbb{R}_+^n$ . The presence of repeated measurements adds complexity to the model in that  $\tilde{\boldsymbol{\Sigma}}$  now must accommodate for the correlation within subjects. Depending on the application at hand, one can evaluate alternative modelling choices for  $\tilde{\boldsymbol{\Sigma}}$ . Since  $\tilde{\mathbf{y}}$  is a vector with  $2n$  elements, the first  $n$  of which are the cosines of the  $\theta_j$ 's and the latter  $n$  are the sines of the  $\theta_j$ 's, it follows that the covariance matrix can be partitioned into

$$\text{cov}(\tilde{\mathbf{y}}) = \tilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{cc} & \tilde{\boldsymbol{\Sigma}}_{cs} \\ \tilde{\boldsymbol{\Sigma}}_{cs} & \tilde{\boldsymbol{\Sigma}}_{ss} \end{pmatrix}. \quad (7)$$

The diagonal block  $\tilde{\boldsymbol{\Sigma}}_{cc}$  (or  $\tilde{\boldsymbol{\Sigma}}_{ss}$ ) controls the variances and covariances of the cosines (or sines) across bouts: the variances are the reflection of the variability in the timings of the activity bouts, while the covariances reflect the correlation between repeated measurements within subjects.

The diagonal elements of the block  $\tilde{\boldsymbol{\Sigma}}_{cs}$  represent within-bout covariances between sines and cosines. These covariances do not have an immediate interpretation although they do have a role in determining the shape of the marginal distribution of  $\theta$  as discussed in Section 3.1. On the other hand, the off-diagonal elements of  $\tilde{\boldsymbol{\Sigma}}_{cs}$  (cross-correlations) may have a lesser role after the within-bout correlation has been taken into account.

For example, consider (7). One could start with a naïve structure of the type  $\tilde{\boldsymbol{\Sigma}}_{cc} = \tilde{\boldsymbol{\Sigma}}_{ss} = \sigma^2 \mathbf{I}_n$  and  $\tilde{\boldsymbol{\Sigma}}_{cs} = \sigma_{cs} \mathbf{I}_n$ , where  $\sigma_{cs} = \rho_{cs} \sigma^2$  and  $\rho_{cs} \in (-1, 1)$  is the cosine-sine correlation (constant across individuals and activity bouts). Among other things, this simple structure implies that the within-bout correlation between cosines (or sines) is null. On the other hand, fitting a general covariance matrix for  $\tilde{\boldsymbol{\Sigma}}_{cc}$  and  $\tilde{\boldsymbol{\Sigma}}_{ss}$  would be unreasonable if we consider that  $n$  can be rather large. In the MCS physical activity data analysis, an autoregressive model proved to be effective.

### 3.2.2 | Conditional approach

The conditional specification of the proposed model involves the inclusion of subject-specific random effects  $\mathbf{u}$  in the location parameter, namely  $\mathbf{m}_u = \mathbf{B}^\top \mathbf{x} + \mathbf{u}$ . The vector  $\mathbf{u}$  is assumed to be  $\mathcal{N}_2(\mathbf{0}, \boldsymbol{\Phi})$ , where  $\boldsymbol{\Phi}$  is the  $2 \times 2$  covariance matrix that controls the unobserved heterogeneity among subjects. The conditional model is then given by

$$\mathbf{y} = \mathbf{B}^\top \mathbf{x} + \mathbf{u} + \boldsymbol{\eta}, \quad (8)$$

where  $\boldsymbol{\eta} \sim \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Psi})$  is independent of  $\mathbf{u}$ . Model (8) can be cast into the marginal form (5) if we let  $\mathbf{e} = \mathbf{u} + \boldsymbol{\eta}$  and  $\boldsymbol{\Sigma} = \boldsymbol{\Phi} + \boldsymbol{\Psi}$ .

The advantage of (8) is that inference (see next section) can be done at the individual level. The extension of (8) to multiple activity bouts is analogous to (6), that is,

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\mathbf{b} + \mathbf{D}\mathbf{u} + \tilde{\boldsymbol{\eta}}, \quad \tilde{\boldsymbol{\eta}} \sim \mathcal{N}_{2n}(\mathbf{0}, \tilde{\boldsymbol{\Psi}}), \quad (9)$$

where  $\mathbf{D} = \mathbf{I}_2 \otimes \mathbf{1}_n$ , while  $\tilde{\mathbf{y}}$ ,  $\tilde{\mathbf{X}}$  and  $\mathbf{b}$  have been defined in the previous section.

Of course, one must be careful not to slip into identifiability issues<sup>41</sup> by over-parameterizing  $\Phi$  and  $\Psi$ . As Pinheiro and Bates<sup>41</sup> (p. 398) put it, “these two components may compete with each other in the model specification, in the sense that similar  $\Sigma$  matrices may result from a more complex random-effects component being added to a simpler within-group component [ ... ], or a simpler random-effects component [ ... ] being added to a more complex within-group component. There will generally be a trade-off between the complexity of the two components [ ... ] and some care must be exercised to prevent nonidentifiability, or near nonidentifiability, of the parameters in the model”. In model (9), the error’s variance matrix  $\Psi$  may contain the variance function  $g$  through which one can introduce the covariates  $\mathbf{w}$  associated with the intensity  $r$ . If this is the case, then the random effects’ structure should be kept to a minimum, for example, by using random intercepts only. If, on the other hand, we were to increase the complexity of the random effects’ structure by adding random ‘slopes’, we might risk losing the interpretability of the parameter  $\delta$  as covariates’ effects on  $r$ . This is briefly exemplified at the end of this section.

A reasonable starting point, which I considered in the MCS physical activity data analysis, is to use a general covariance matrix for  $\Phi$  and a diagonal matrix for  $\Psi$ , which leads to

$$\text{cov}(\tilde{\mathbf{y}}) \equiv \tilde{\Sigma} = \mathbf{D}\Phi\mathbf{D}^\top + \Psi = \begin{pmatrix} \phi_c^2 \mathbf{J}_n & \phi_{cs} \mathbf{J}_n \\ \phi_{cs} \mathbf{J}_n & \phi_s^2 \mathbf{J}_n \end{pmatrix} + \mathbf{I}_2 \otimes \text{diag}\{\psi^2 \mathbf{g}^2\}. \quad (10)$$

From this model, it follows that the marginal covariance matrix as partitioned in (7) has blocks  $\tilde{\Sigma}_{cc} = \phi_c^2 \mathbf{J}_n + \text{diag}\{\psi^2 \mathbf{g}^2\}$ ,  $\tilde{\Sigma}_{ss} = \phi_s^2 \mathbf{J}_n + \text{diag}\{\psi^2 \mathbf{g}^2\}$ , and  $\tilde{\Sigma}_{cs} = \phi_{cs} \mathbf{J}_n$ . Moreover, for two bouts  $j$  and  $l$ ,  $j \neq l$ ,  $1 \leq j, l \leq n$ ,

$$\rho_c = \frac{\phi_c^2}{\sqrt{(\phi_c^2 + \psi^2 g_j^2)(\phi_c^2 + \psi^2 g_l^2)}},$$

is the intra-cluster correlation between  $\cos \theta_j$  and  $\cos \theta_l$ ,

$$\rho_s = \frac{\phi_s^2}{\sqrt{(\phi_s^2 + \psi^2 g_j^2)(\phi_s^2 + \psi^2 g_l^2)}},$$

is the intra-cluster correlation between  $\sin \theta_j$  and  $\sin \theta_l$ , and

$$\rho_{cs} = \frac{\phi_{cs}}{\sqrt{(\phi_c^2 + \psi^2 g_j^2)(\phi_s^2 + \psi^2 g_l^2)}},$$

is the intra-cluster correlation between  $\cos \theta_j$  and  $\sin \theta_l$ . The between-cosines and between-sines correlations can be positive only, while the cross-correlations can be positive or negative.

In the example above, the matrix  $\tilde{\Sigma}$  depends on the covariates only through its diagonal elements. To broaden the dependency to its off-diagonal elements as well, one could consider a correlation function  $h(\rho)$  as discussed in Section 3.3. However, correlation functions typically implemented in statistical software are restricted to model temporal and spatial correlation (eg, see `corClasses` in R). Alternatively, one could extend model (9) to include random slopes so that the term  $\mathbf{D}\Phi\mathbf{D}^\top$  in  $\tilde{\Sigma}$  reflects the desired covariate-dependent correlation structure. However, as mentioned above, if  $\mathbf{D}$  includes covariates also included in  $g$ , the parameter  $\delta$  alone would only partially reflect the association between  $\mathbf{w}$  and  $r$ .

### 3.3 | Inference

Consider accelerometer observations, possibly over multiple days, from a sample of  $M$  individuals that have been processed into activity bouts according to, for example, the procedures described in Section 2. Let  $r_{ij}$  be the total counts associated with the  $j$  th bout of activity for the  $i$  th individual,  $j = 1, \dots, n_i$  and  $i = 1, \dots, M$ , and let  $\theta_{ij}$  be the timing of such a bout, with  $N = \sum_{i=1}^M n_i$ . Let  $\mathbf{x}_{ij}^\top$  and  $\mathbf{w}_{ij}^\top$  be the  $j$  th rows of the design matrices  $\mathbf{X}_i$  and  $\mathbf{W}_i$ , respectively. The responses  $\mathbf{y}_{ij} = (r_{ij} \cos \theta_{ij}, r_{ij} \sin \theta_{ij})^\top$ ,  $j = 1, \dots, n_i$ , are arranged in the  $2n_i$ -long vector  $\tilde{\mathbf{y}}_i = (r_{i1} \cos \theta_{i1}, \dots, r_{in_i} \cos \theta_{in_i}, \dots, r_{i1} \sin \theta_{i1}, \dots, r_{in_i} \sin \theta_{in_i})^\top$ . Estimation of models (6) and (9) for  $M$  subjects can be carried

out using generalized least squares (GLS) and maximum likelihood (ML) via standard software for linear models (eg, `gls` in  $\mathbb{R}^{27}$ ) and linear mixed effects models (eg, `lme` in  $\mathbb{R}$ ). An example with simulated data is provided in Supplemental Material.

To accommodate for the activity counts' covariates, one can resort to several types of variance functions (in  $\mathbb{R}$ , these are called `varClasses` and are nicely discussed in Pinheiro and Bates<sup>41</sup>). For example, if  $w \in \{1, 2, \dots, S\}$  is a stratification variable taking  $S$  distinct values, the function

$$g_{Ident}(w, \delta) = \delta_w$$

represents a variance model with different variances for each level of  $w$  (`varIdent` in  $\mathbb{R}$ ). The parameter  $\delta = (\delta_1, \delta_2, \dots, \delta_S)^\top \in \mathbb{R}_{>0}^S$  consists of multipliers of the *baseline* scale  $\sigma$  (hence, for identifiability one of the elements is set equal to 1). Therefore, the element  $\delta_h$  of  $\delta$ ,  $h = 2, \dots, S$ , can be interpreted as the *ratio* between the standard deviations of strata  $h$  and 1, the latter being taken as baseline. If, say,  $\delta_S = 1.5$ , then it means that the residual standard deviation for those in stratum  $S$  is 50% larger than  $\sigma$ . Other models, such as power and exponential, can be used for continuous covariates, and different variance functions for individual covariates may be combined into a single variance function (eg, via `varComb` in  $\mathbb{R}$ ).

The off-diagonal elements of  $\tilde{\Sigma}$  (marginal specification) or  $\tilde{\Psi}$  (conditional specification) of the error can be modelled using a correlation function (eg, see `corClasses` in  $\mathbb{R}$ ). In the data analysis of Section 4, I considered only two alternatives, although a few others are available,<sup>41</sup> starting from the compound symmetric (CSym) correlation function

$$h_{CSym}(\rho) = \rho,$$

where the intra-class correlation parameter  $\rho \in \mathbb{R}$  is constant among all within-group errors pertaining to the same group. Although the CSym model is extremely parsimonious, it may oversimplify the correlation structure of time series data. Alternatively, an equally parsimonious but potentially more realistic model is given by an autoregressive correlation function of order 1. If the spacings between the events are irregular, as is the case in the type of applications that motivated this article, one must consider the *continuous time* autoregressive (CAR) correlation function

$$h_{CAR}(\rho) = \rho^t,$$

where  $\rho \geq 0$  and  $t \geq 0$ .

One caveat about fitting models (6) and (9) is that the size of the problem can quickly grow with the number of bouts. For example, in the application of the present study, the size of the dataset to be fed to `gls` and `lme` was about 668 thousand rows. While this size is not particularly challenging to be handled by even standard computers, speeding up computation may be desirable for scalability.

Once fitted, models can be evaluated overall by their likelihood and related test and criteria (eg, Akaike Information Criterion). Note, however, that residuals from models (6) and (9) cannot be used directly to assess the goodness of fit for individual response distributions. For the intensity  $r$ , I propose the mean absolute deviations  $D_r = N^{-1} \sum_i \sum_j |r_{ij} - \hat{r}_{ij}|$ , while for the timing  $\theta$  the mean circular distance<sup>18</sup>  $D_\theta = N^{-1} \sum \{1 - \cos(\theta_{ij} - \hat{\theta}_{ij})\}$ . The predicted values can be worked out from the marginal distributions of  $r$  and  $\theta$ .

### 3.4 | Comparison with alternative approaches and simulation study

To recap, the proposed model starts from the joint distribution of  $r$  and  $\theta$ , which are the polar coordinates of the vector  $\mathbf{y} \equiv r\mathbf{z} = (r \cos \theta, r \sin \theta)^\top$ , it allows for repeated measurements, and it can be fitted using standard software. An alternative approach is the one that goes under the label of cylindrical data modelling, where the circular component  $\theta$  and the linear component, say  $s$ , are modelled separately. That is, the joint model  $f(s, \theta)$  is obtained via the product  $f(s|\theta)f(\theta)$  or by means of copulas on the marginals  $f(s)$  and  $f(\theta)$ . Note that the role of the intensity in cylindrical models is played by  $s$ . The radius  $r$  either does not appear in the definition of the model or, if does, it is latent. For example, in a recently published work,<sup>44</sup> the distribution of interest is obtained via  $f(s, \theta) = \int_0^\infty f(s, \theta, r) dr = f(s|\theta) \int_0^\infty f(r, \theta) dr$ , where  $f(r, \theta)$  follows the distribution in (1) and, as a result,  $f(\theta)$  is the GPN defined in (2). (And, as already mentioned, identifiability of the GPN requires imposing a constraint on  $\Sigma$ .) Another similar distribution, which dates back to 1970s,<sup>45</sup> makes use of

$f(s, \theta|r = 1) = f(s|\theta)f(\theta|r = 1)$ , where  $f(\theta|r = 1)$  follows the distribution in (3). The advantage of building the distribution  $f(s, \theta)$  from two separate distributions (one linear and one circular) is the flexibility of exploring different models for one or the other. For example,  $f(s|\theta)$  in the above-referenced works can be normal<sup>45</sup> or skewed-normal,<sup>44</sup> which leads to interpret the location parameter for the intensity on the additive scale (as opposed to the multiplicative scale associated with the scale parameter of  $r$  in (1)). However, because of this flexibility, the resulting joint, marginal, or conditional distributions may be analytically difficult to handle.<sup>46</sup> In any case, there seem to be limited methodological work and software for cylindrical regression and this is confined to the Bayesian framework,<sup>47</sup> while proposals for cylindrical regression with repeated measures seem to be lacking.

Disjoint modelling of  $r$  and  $\theta$  is, of course, an option. Besides the disadvantage of precluding multivariate calculations (eg. conditional distributions and product moments), disjoint modelling poses some specific issues vis-à-vis the models and software commonly available. As far as  $r$  is concerned, one may use a linear model with general covariance for the errors or a mixed-effects model. While these models handle the intra-cluster correlation, the distributional assumptions do not bode well with the skewed distribution of the errors. Skewed models for repeated measurements that can compete computationally with those for normal responses are still lacking. As far as  $\theta$  is concerned, circular regression<sup>19,20</sup> for independent observations revolves around the von Mises distribution, whose unimodal shape may be a limitation as already stressed in Section 4.1. There are also some proposals for longitudinal circular responses,<sup>33,48-51</sup> although they have limited or no software implementations. It is noteworthy to point out that the conditional specification in (8) starts from the same model considered by others.<sup>33</sup> However, their ultimate goal is to estimate the marginal distribution of  $\theta$  by assuming a latent radial distance. This requires not only addressing an issue of identifiability of the scale, but also the application of an expectation-maximization algorithm. The latter is notoriously time-consuming and practically unusable when the sample size is hundreds of thousands observations as in the application of the present article.

In a simulation study (Supplemental Material), I assessed the performance of the proposed models (6) and (9). The first aim of the simulation (part A) was to gauge the ability of the models to provide a satisfactory description of the data generating process when this is not completely known a priori, using distributions that are not exactly the same distributions underlying models (5) and (8). The data were generated from three scenarios, the first of which mimicked aspects one would encounter in the analysis of accelerometer data, while the other two scenarios provided gross deviations from the first one. Overall, there was excellent agreement between estimated summary statistics and their sample equivalents in the first scenario (Supplemental Table S1). The performance deteriorated somewhat when the distribution of  $\theta$  was more peaked (scenario 2, Supplemental Table S2) and when the distribution of  $r$  was symmetric (scenario 3, Supplemental Table S3), in particular with respect to the estimated conditional standard deviations.

The second aim of the simulation (part B) was to compare the proposed models to existing models capable of achieving the goals set out in Sections 3.2.1 and 3.2.2. However, given the paucity of alternative approaches and of readily-available software, the comparison was limited to disjoint modelling of intensity and timing as described above. In summary, the disjoint model was generally less efficient and unable to deal with heteroskedastic timings (Supplemental Table S4). Also, the proposed model gave on average a higher likelihood.

In summary, the proposed approach is particularly apt for the analysis of accelerometer data as it is inexpensive (computationally and software-wise), reasonably accurate and flexible enough to accommodate covariate effects on the mean and the variance of both outcomes, as well as the intra-cluster correlation due to repeated measurements. On the other hand, the simulation has shown that the proposed models present some limitations when the distribution of the timing is peaked or the distribution of the intensity is symmetric, neither of which, however, is a concern for the analysis of the MCS data (Section 4). Currently there are no alternative approaches that have all these characteristics at once, although one may expect that further developments in cylindrical regression will bring about competitive models that are well-developed in all aspects.

## 4 | RESULTS

In this section, I analyze the MCS physical activity data introduced in Section 2 using the models described in Section 3. I start with the marginal models discussed in Section 3.2.1 and then consider the conditional approach discussed in Section 3.2.2. The response variable  $\tilde{y}$  was centred at 2240 and scaled by 1000 to avoid numerical instability. After a small number of children (71) with incomplete observations were removed, there were  $M = 1802$  children and a total number of MVPA bouts equal to 333, 977 available for the analyses. The latter were conducted in R version 4.1.2 on a 64-bit operating system machine with 16 GB of RAM and quad-core processor at 2.80 GHz. The main R packages used were `n1me`<sup>52</sup> for

GLS and ML estimation, `mvtnorm`<sup>53</sup> for multivariate normal calculations, and `SphericalCubature`<sup>54</sup> for numerical integration and transformations to or from polar coordinates.

## 4.1 | Marginal approach

Let  $\mathbf{x}$  denote the vector that contains 3 dummy variables for country (reference: England), age (centred at 7 years), a dummy variable for sex (reference: male), a dummy for ethnicity (reference: White), bodyfat (centred at 20%), waist circumference (centred at 56 cm), 2 dummies for mode of transport to and from school (reference: active), a dummy for weekend (reference: weekday, ie, Monday to Friday), and a dummy for spring (reference: not spring) (hence,  $p = 13$ ). The vector  $\mathbf{w}$  includes the interaction between the dummies for sex and *dichotomized* time of the day (reference: male and AM), a dummy for ethnicity, and a dummy for spring (hence,  $q = 6$ ). Moreover, *continuous* time of the day, denoted by  $t$ , is expressed in minutes. In increasing order of complexity, I considered the following four marginal models:

1. *Model 1*, where neither the location parameter  $\mathbf{m}$  nor the covariance matrix  $\tilde{\Sigma}$  depends on the covariates, the error's variance is constant and equal to  $\sigma^2$ , and the correlation function is CSymm with parameter  $\rho_{cs}$  (constant across individuals and activity bouts). This random structure has been referred to as naïve at the end of Section 3.2.1. The parameters to be estimated are  $\mathbf{b} = (b_{0,1}, b_{0,2})^\top$ ,  $\rho_{cs}$  and  $\sigma$ .
2. *Model 2*, where the parameters are specified as in Model 1, except the location parameter depends on  $\mathbf{x} = (1, x_1, \dots, x_{12})^\top$ , thus  $\mathbf{b} = (b_{0,1}, b_{1,1}, \dots, b_{12,1}, b_{0,2}, b_{1,2}, \dots, b_{12,2})^\top$
3. *Model 3*, where the parameters are specified as in Model 2, except the variance depends on  $\mathbf{w} = (w_1, \dots, w_6)^\top$  via the variance function  $g(\mathbf{w}, \delta) = \prod_{r=1}^6 g_{Ident}(w_r, \delta_r)$ , thus  $\delta = (1, \delta_2, \dots, \delta_6)^\top$
4. *Model 4*, where the parameters are specified as in Model 3, except that the CAR correlation function  $h_{CAR}(\rho) = \rho^t$ , nested within child and date, is used in place of the CSymm correlation function.

The estimates of the marginal models described above, along with 95% confidence intervals, are reported in Table 2. The likelihood ratio test (LRT) statistic comparing Model 2 to Model 1 was 1656 ( $p$ -value  $< 0.0001$ ) indicating that the fit improves when the covariates are included in the location parameter. Similarly, the LRT statistic comparing Model 3 to Model 2 was 20425 ( $p$ -value  $< 0.0001$ ), which strongly supports the inclusion of a covariate-dependent variance function. However, Model 4 was clearly superior to the others in terms of both goodness of fit and Akaike Information Criterion (AIC). The improvement of Model 4 compared to Model 3 was due merely to the correlation functional form, that is, with no increase in complexity.

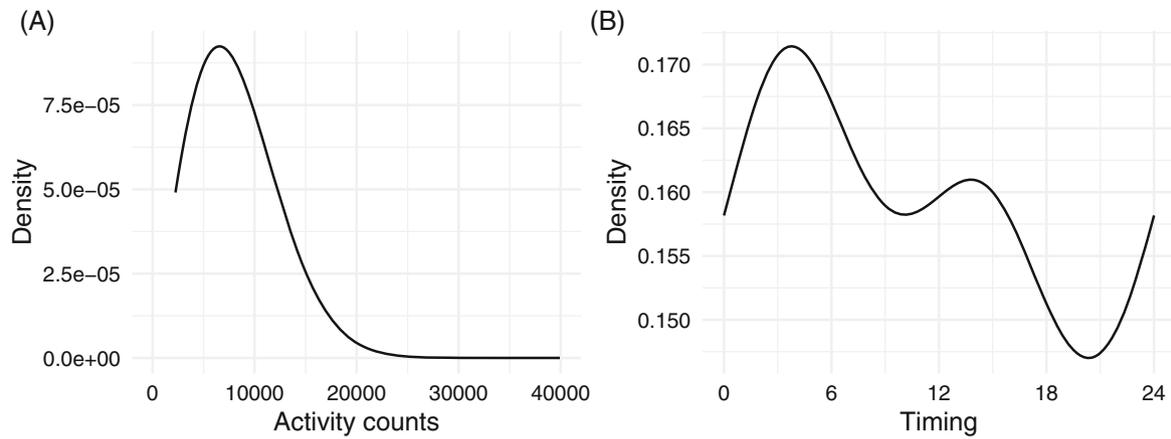
A particular strength of jointly modelling intensity and timing of activity bouts is that, as a byproduct, marginal and conditional distributions can be obtained. For example, let's start with the simplest of the four models described above. The marginal distributions of  $r$  and  $\theta$  obtained from the fitted Model 1 are given in Figure 5. The former was calculated via numerical integration, while the latter was obtained analytically using (2). As remarked in Section 3.1, the distributions of physical activity outcomes are notoriously skewed. This aspect is nicely captured by the estimated  $\hat{f}(r)$ , which gives an overall mean and standard deviation equal to 8226 and 4302 (very close to mean and standard deviation of a Rayleigh distribution with parameter  $\sigma = 6.56 \cdot 1000$ ). On the other hand, the estimated  $\hat{f}(\theta)$  captures the bimodality of the timing distribution with a peak in the early morning and one in the afternoon. This aspect is ignored by the von Mises distribution as it is unimodal. Moreover, the application of Equation (4) to calculate the mean of  $r$  conditional on  $\theta$ , gave 8459 (standard deviation 4403) for the morning peak and 8305 (standard deviation 4350) for the afternoon peak indicating that these two peak times do not differ greatly in terms of mean activity counts.

While Model 1 is a convenient starting point, it clearly does not explain the relationships among the variables. As Model 4 was the preferred model, I elaborate on a few of its estimated parameters. It should be borne in mind that how ultimately any of these parameters impact the (marginal) distributions of  $r$  or  $\theta$  cannot be appreciated fully until these distributions and related statistics are worked out. Still, one can get a feel for the direction and magnitude of the effects with some back-of-the-envelope calculations. For example, consider country. The estimates of the intercepts indicate that, for children in the reference category, the mean timing of the activity bouts for those living in England was sometime in the afternoon. More precisely, one gets  $\text{atan2}(\hat{b}_{0,2}, \hat{b}_{0,1}) = \text{atan2}(-1.80, -2.96) = 3.69$  in radians, corresponding to 14:05 (while the sample circular mean for the reference category with arbitrary age, bodyfat and waist circumference was 14:23). On the other hand, the estimated coefficients for the other British countries imply a counter-clockwise angle rotation, that is, a *later* mean timing compared to England. More precisely,  $\text{atan2}(\hat{b}_{0,2} + \hat{b}_{1,2}, \hat{b}_{0,1} + \hat{b}_{1,1}) = 3.73$  (corresponding to 14:14) for Wales,  $\text{atan2}(\hat{b}_{0,2} + \hat{b}_{2,2}, \hat{b}_{0,1} + \hat{b}_{2,1}) = 3.74$  (corresponding to 14:17) for Scotland, and  $\text{atan2}(\hat{b}_{0,2} + \hat{b}_{3,2}, \hat{b}_{0,1} + \hat{b}_{3,1}) = 3.80$

**TABLE 2** Estimates (95% confidence intervals) of the parameters from four different marginal models for the Millennium Cohort Study accelerometer data

Parameter	Model 1	Model 2	Model 3	Model 4
<i>Fixed (cos <math>\theta</math>)</i>				
Intercept	- 2.25 (- 2.28, - 2.23)	- 2.76 (- 2.82, - 2.71)	- 2.78 (- 2.83, - 2.73)	- 2.96 (- 3.04, - 2.89)
Wales		0.11 (0.04, 0.19)	0.11 (0.04, 0.18)	0.06 (- 0.05, 0.16)
Scotland		0.18 (0.11, 0.25)	0.16 (0.09, 0.23)	0.19 (0.09, 0.30)
Northern Ireland		0.55 (0.47, 0.63)	0.49 (0.41, 0.57)	0.43 (0.31, 0.54)
Age (years)		- 0.07 (- 0.16, 0.02)	- 0.09 (- 0.17, 0.00)	- 0.09 (- 0.22, 0.04)
Female		0.51 (0.46, 0.55)	0.57 (0.53, 0.62)	0.69 (0.62, 0.75)
Other than White		0.62 (0.56, 0.69)	0.61 (0.55, 0.67)	0.63 (0.54, 0.72)
Bodyfat (%)		0.00 (- 0.01, 0.00)	0.00 (- 0.01, 0.00)	0.00 (- 0.01, 0.01)
Waist (cm)		0.01 (0.00, 0.01)	0.01 (0.00, 0.01)	0.01 (0.00, 0.02)
Passive transport to school		0.18 (0.14, 0.23)	0.18 (0.14, 0.23)	0.31 (0.24, 0.37)
Passive transport from school		0.02 (- 0.09, 0.14)	0.03 (- 0.08, 0.14)	0.06 (- 0.10, 0.23)
Weekend		0.23 (0.18, 0.28)	0.24 (0.19, 0.29)	0.29 (0.22, 0.36)
Spring		- 0.16 (- 0.23, - 0.09)	- 0.19 (- 0.26, - 0.12)	- 0.21 (- 0.32, - 0.11)
<i>Fixed (sin <math>\theta</math>)</i>				
Intercept	- 1.98 (- 2.00, - 1.96)	- 2.06 (- 2.11, - 2.01)	- 1.99 (- 2.05, - 1.94)	- 1.80 (- 1.87, - 1.72)
Wales		- 0.15 (- 0.22, - 0.08)	- 0.13 (- 0.20, - 0.05)	- 0.13 (- 0.23, - 0.02)
Scotland		- 0.02 (- 0.10, 0.05)	- 0.05 (- 0.12, 0.02)	- 0.09 (- 0.19, 0.02)
Northern Ireland		- 0.17 (- 0.25, - 0.09)	- 0.13 (- 0.21, - 0.05)	- 0.14 (- 0.26, - 0.03)
Age (years)		0.04 (- 0.05, 0.13)	0.04 (- 0.05, 0.12)	0.06 (- 0.07, 0.19)
Female		0.37 (0.33, 0.42)	0.21 (0.17, 0.26)	0.32 (0.26, 0.39)
Other than White		0.13 (0.06, 0.19)	0.14 (0.08, 0.20)	0.12 (0.03, 0.21)
Bodyfat (%)		- 0.01 (- 0.02, 0.00)	- 0.01 (- 0.01, 0.00)	- 0.01 (- 0.01, 0.00)
Waist (cm)		0.01 (0.00, 0.01)	0.01 (0.00, 0.01)	0.00 (- 0.01, 0.01)
Passive transport to school		0.07 (0.02, 0.11)	0.06 (0.01, 0.10)	0.03 (- 0.04, 0.09)
Passive transport from school		0.17 (0.05, 0.28)	0.20 (0.09, 0.31)	0.20 (0.04, 0.37)
Weekend		- 0.22 (- 0.27, - 0.17)	- 0.23 (- 0.27, - 0.18)	- 0.26 (- 0.33, - 0.19)
Spring		- 0.40 (- 0.47, - 0.33)	- 0.44 (- 0.50, - 0.37)	- 0.43 (- 0.53, - 0.33)
<i>Random</i>				
Male $\times$ PM			1.03 (1.03, 1.04)	1.08 (1.08, 1.09)
Female $\times$ AM			0.85 (0.85, 0.86)	0.85 (0.84, 0.85)
Female $\times$ PM			0.80 (0.79, 0.80)	0.86 (0.85, 0.86)
Other than White			0.90 (0.89, 0.90)	0.90 (0.89, 0.90)
Spring			1.11 (1.10, 1.12)	1.11 (1.10, 1.12)
Residual correlation (CSymm)	0.04 (0.03, 0.04)	0.04 (0.03, 0.04)	0.04 (0.03, 0.04)	
Residual correlation (CAR)				0.90 (0.90, 0.90)
Residual standard deviation	6.56 (6.55, 6.57)	6.55 (6.54, 6.56)	6.97 (6.94, 7.00)	7.18 (7.15, 7.22)
<i>Log-likelihood (df)</i>	- 2204122 (4)	- 2203294 (28)	- 2193081 (33)	- 2137776 (33)
<i>AIC</i>	4408251	4406643	4386228	4275618

Note: The response is centered at 2240 and scaled by 1000.



**FIGURE 5** Estimated marginal distributions of (A) activity counts and (B) timing obtained from Model 1 for the Millennium Cohort Study accelerometer data.

(corresponding to 14:29) for Northern Ireland. Sex, ethnic group, mode of transport to school, day of the week and season also had non-negligible effects on timing, with female sex, ethnicity other than White, passive home-to-school commuting, weekend and spring associated with counter-clockwise angle rotations.

The parameters of the random structure of the model have a direct relationship with the covariates' effects on the intensity of the activity bouts (this is supported by the simulation study in Supplemental Section S1). The 'afternoon effect', first noted in Table 1, is particularly evident for males since, as estimated by Model 4, the intensity of their activity was 8% higher in the postmeridian hours of the day. On the other hand, females were notably less active than boys by about 15%, while children of non-White ethnicity were less active than their counterparts by about 10%. These results are consistent with previous findings.<sup>11</sup> Spring had a strong positive effect equal to 11% more activity counts compared to other seasons. Finally, the temporal correlation of the MVPA bouts was very high (0.90) as one would expect.

The mean absolute deviations and circular distance for Model 4 were  $D_r = 4.303$  and  $D_\theta = 0.356$ , respectively.

## 4.2 | Conditional approach

In this section, I present the analysis of the MCS physical activity data using the conditional approach. Although several models were fitted, for the sake of brevity and in light of the the results discussed in the previous section, I restrict the attention to a model that had the same location parameter and variance function as those of Model 3, but no residual correlation function. For this conditional model a general covariance matrix  $\Phi$  with diagonals  $\phi_c^2$  and  $\phi_s^2$ , and off-diagonal  $\phi_{cs} = \xi_{cs}\phi_s\phi_c$ , was used for the distribution of the random intercepts. Thus, the marginal covariance matrix resulting from this conditional model had the form (10).

The estimates of the location parameter and of the variance function were similar to those from marginal Model 3 (see Table 2). As for the random-effects parameters, the estimates were relatively modest in magnitude compared to the residual standard deviation (Table 3). However, to appreciate how these estimates translate into heterogeneity among children, subject-specific estimates for all children and activity bouts are given in Figure 6. The left-hand plot shows the estimated subject-specific distributions of  $r$  obtained from the fitted model (8), that is, the distributions  $\mathcal{N}_2(\hat{\mathbf{B}}^\top \mathbf{x}_{ij} + \hat{\mathbf{u}}_i, \hat{\Psi}_{ij})$  marginalized with respect to  $\theta$ . These describe the individual children's MVPA activity counts distributions. The right-hand plot shows the estimated circular means calculated from the subject-specific distributions of  $\theta$ . These describe the average individual children's timings of activity bouts. The red dashed lines mark the 'population' level marginals.

As compared to the marginal approach, there was a slight improvement in terms of mean absolute deviations and circular distance from this conditional model, namely  $D_r = 4.295$  and  $D_\theta = 0.344$ , respectively.

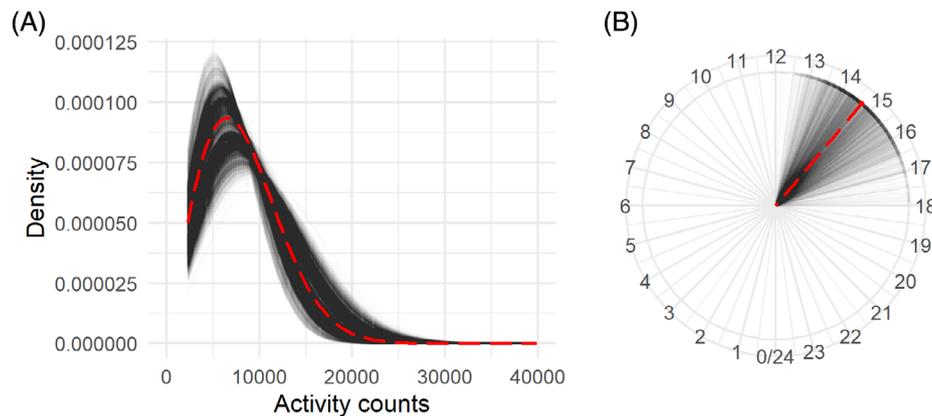
## 4.3 | Disjoint modelling

Supplemental Table S5 shows the estimates from a linear mixed-effects regression<sup>52</sup> for the activity counts ( $r$ ) and, fitted separately, a circular-linear regression<sup>55</sup> for the timing of the activity bouts ( $\theta$ ). The parameters of the latter are reported

**TABLE 3** Estimates (95% confidence intervals) of the random-effects parameters and residual standard deviation from the conditional model for the Millennium Cohort Study accelerometer data

Parameter	
$\phi_c^2$	1.07 (0.99, 1.15)
$\phi_s^2$	0.91 (0.86, 0.97)
$\xi_{cs}$	0.13 (− 0.05, 0.31)
Residual standard deviation	6.81 (6.75, 6.88)

Note: The response is centered at 2240 and scaled by 1000.



**FIGURE 6** Estimated subject-specific distributions of (A) activity counts and (B) timing obtained from the conditional model for the Millennium Cohort Study accelerometer data. The red dashed lines mark the ‘population’ level marginals.

as provided by the R function `lm.circular`, except for the estimate of the concentration  $\hat{\kappa} = 1.7$ , which corresponds to a circular standard deviation equal to 0.94.

Notably, the circular-linear regression coefficients were unexpectedly small, suggesting that according to this model only country and ethnicity had non-negligible effects on timing. On the other hand, the goodness of fit statistics for intensity and timing were somewhat worse than those reported in Sections 4.1 and 4.2, namely,  $D_r = 4.940$  and  $D_\theta = 0.354$ .

## 5 | FINAL REMARKS

Joint regression modelling of intensity and timing of accelerometer counts can be done easily by exploiting the polar coordinates representation of the response. In this instance, I considered normal regression models, but nothing prevents this approach from being implemented by means of other distributions.<sup>56</sup> It should be stressed, though, that the normality assumption on the polar coordinates does not necessarily imply unimodality and symmetry of the resultant marginal distributions. On the contrary, as seen in the analysis of the MCS accelerometer data, skewness in the distribution of the activity counts and bimodality in the distribution of the timings arose naturally from the marginalized normal models. This represents an advantage over disjoint modelling of  $r$  and  $\theta$  with, respectively, normal and von Mises assumptions. But an even greater advantage of joint modelling over separate univariate models is the ability of the former to allow, at relatively small computational costs, multivariate calculations.

It might be worth considering an extension of the present approach to the modelling of both start and end times of the bouts, rather than just their intermediate time points. While it is reasonable to use a single time point for accelerometer outcomes that have relatively short durations (such as those of MVPA bouts), it may be preferable or even necessary to consider both onset and offset of the events with extended durations such as those that occur during sedentary and sleep behaviours.

## FUNDING INFORMATION

None reported.

## CONFLICT OF INTEREST

The author declares no potential conflict of interests.

## DATA AVAILABILITY STATEMENT

The Millennium Cohort Study data are available from the UK Data Service to its registered users. The R code implementing the proposed methods is available in supplemental materials for this article.

## ORCID

Marco Geraci  <https://orcid.org/0000-0002-6311-8685>

## REFERENCES

1. Ward DS, Evenson KR, Vaughn A, Rodgers AB, Troiano RP. Accelerometer use in physical activity: Best practices and research recommendations. *Med Sci Sports Exerc.* 2005;37(11):S582-S588.
2. Jean-Louis G, Kripke DF, Cole RJ, Assmus JD, Langer RD. Sleep detection with an accelerometer actigraph: comparisons with polysomnography. *Physiol Behav.* 2001;72(1):21-28.
3. Duque R, Nieto-Reyes A, Martínez C, Montaña JL. Detecting human movement patterns through data provided by accelerometers. A case study regarding Alzheimer's disease. In: García CR, Caballero-Gil P, Burmester M, Quesada-Arencibia A, eds. *Ubiquitous Computing and Ambient Intelligence* Springer. Cham, Switzerland: Springer; 2016: 56-66.
4. Bringas S, Salomón S, Duque R, Lage C, Montaña JL. Alzheimer's disease stage identification using deep learning models. *J Biomed Inform.* 2020;109:103514.
5. Brennan JH, Mitra B, Synnot A, et al. Accelerometers for the assessment of concussion in male athletes: a systematic review and meta-analysis. *Sports Med.* 2017;47(3):469-478.
6. Kiernan D, Hawkins DA, Manoukian MAC, et al. Accelerometer-based prediction of running injury in national collegiate athletic association track athletes. *J Biomech.* 2018;73:201-209.
7. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med.* 2014;48(13):1019.
8. Griffiths LJ, Rich C, Geraci M, et al. *Technical Report on the Enhancement of Millennium Cohort Study Data with Accelerometer-Derived Measures of Physical Activity and Sedentary Behaviour in Seven Year Olds.* London, UK: Tech Rep Centre for Longitudinal Studies; 2013.
9. Doherty A, Jackson D, Hammerla N, et al. Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study. *PLOS One.* 2017;12(2):1-14.
10. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc.* 2008;40(1):181-188.
11. Griffiths LJ, Cortina-Borja M, Sera F, et al. How active are our children? Findings from the millennium cohort study. *BMJ Open.* 2013;3(8):e002893.
12. Griffiths LJ, Geraci M, Cortina-Borja M, et al. Associations between children's behavioural and emotional development and objectively measured physical activity and sedentary time: findings from the UK millennium cohort study. *Longitud Life Course Studies.* 2016;7(2):124-143.
13. Noonan RJ, Fairclough SJ. Is there a deprivation and maternal education gradient to child obesity and moderate-to-vigorous physical activity? Findings from the millennium cohort study. *Pediatr Obes.* 2018;13(7):458-464.
14. Geraci M, Farcomeni A. Probabilistic principal component analysis to identify profiles of physical activity behaviours in the presence of nonignorable missing data. *J R Stat Soc C.* 2016;65(1):51-75.
15. Geraci M. Additive quantile regression for clustered data with an application to children's physical activity. *J R Stat Soc C.* 2018;68(4):1071-1089.
16. Sera F, Griffiths LJ, Dezateux C, Geraci M, Cortina-Borja M. Using functional data analysis to understand daily activity levels and patterns in primary school-aged children: cross-sectional analysis of a UK-wide study. *PLOS One.* 2017;12(11):e0187677.
17. Morris JS, Arroyo C, Coull BA, Ryan LM, Herrick R, Gortmaker SL. Using wavelet-based functional mixed models to characterize population heterogeneity in accelerometer profiles: a case study. *J Am Stat Assoc.* 2006;101(476):1352-1364.
18. Mardia KV. *Statistics of Directional Data.* Orlando, FL: Academic Press; 1972.
19. Mardia KV, Jupp PE. *Directional Statistics.* Chichester, UK: Wiley; 2000.
20. Pewsey A, Neuhauser M, Ruxton GD. *Circular Statistics in R.* Oxford, UK: Oxford University Press; 2013.
21. Spasojević G, Malobabić S, Mikić D, Vujnović S, Pilipović Spasojević O, Maliković A. Sex differences of human corpus callosum revealed by polar coordinate system: magnetic resonance imaging study. *Folia Morphol.* 2015;74(4):414-420.
22. DaRe JT, Kouri DP, Zimmerman PA, Thomas PJ. Differentiating plasmodium falciparum alleles by transforming Cartesian X,Y data to polar coordinates. *BMC Genet.* 2010;11(1):1-13.
23. Navarro A, Morillo JP, Reigal RE, Hernández-Mendo A. Polar coordinate analysis in the study of positional attacks in beach handball. *Int J Perform Anal Sport.* 2018;18(1):151-167.
24. Smith K, Joshi H. The millennium cohort study. *Popul Trends.* 2002;107:30-34.
25. Pulsford RM, Cortina-Borja M, Rich C, Kinnafick FE, Dezateux C, Griffiths LJ. Actigraph accelerometer-defined boundaries for sedentary behaviour and physical activity intensities in 7 year old children. *PLOS One.* 2011;6(8):e21822.

26. Geraci M. pawacc: physical activity with accelerometers, R package version 1.2.2.; 2017. <https://CRAN.R-project.org/package=pawacc>.
27. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2021. <https://www.R-project.org/>.
28. Geraci M, Rich C, Sera F, Cortina-Borja M, Griffiths LJ, Dezateux C. *Technical report on accelerometry data processing in the Millennium Cohort Study*. London, UK: Tech Rep University College London; 2012.
29. Rich C, Geraci M, Griffiths LJ, Sera F, Dezateux C, Cortina-Borja M. Quality control methods in accelerometer data processing: identifying extreme counts. *PLOS One*. 2014;9(1):e85134.
30. Rich C, Geraci M, Griffiths LJ, Sera F, Dezateux C, Cortina-Borja M. Quality control methods in accelerometer data processing: defining minimum wear time. *PLOS One*. 2013;8(6):e67206.
31. Wang F, Gelfand AE. Directional data analysis under the general projected normal distribution. *Stat Methodol*. 2013;10(1):113-127.
32. Nuñez-Antonio G, Gutiérrez-Peña E. A Bayesian model for longitudinal circular data based on the projected normal distribution. *Comput Stat Data Anal*. 2014;71:506-519.
33. Maruotti A. Analyzing longitudinal circular data by projected normal models: a semi-parametric approach based on finite mixture models. *Environ Ecol Stat*. 2016;23(2):257-277.
34. Hernandez-Stumpfhauser D, Breidt MJ. The general projected normal distribution of arbitrary dimension: modeling and Bayesian inference. *Bayesian Anal*. 2017;12(1):113-133.
35. Gutiérrez L, Gutiérrez-Peña E, Mena RH. A Bayesian approach to statistical shape analysis via the projected normal distribution. *Bayesian Anal*. 2019;14(2):427-447.
36. Cremers J, Mulder KT, Klugkist I. Circular interpretation of regression coefficients. *Br J Math Stat Psychol*. 2018;71(1):75-95.
37. Presnell B, Morrison SP, Littell RC. Projected multivariate linear models for directional data. *J Am Stat Assoc*. 1998;93(443):1068-1077.
38. Ahmed AA, Zahran AR, Ismail MA, Saad AEN. On maximum likelihood estimation of the general projected normal distribution. *J Stat Comput Simul*. 2021;91(16):3453-3472.
39. Baldwin SA, Fellingham GW, Baldwin AS. Statistical models for multilevel skewed physical activity data in health research and behavioral medicine. *Health Psychol*. 2016;35(6):552-562.
40. Weil H. The distribution of radial error. *Ann Math Stat*. 1954;25(1):168-170.
41. Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. New York, NY: Springer; 2000.
42. Diggle PJ, Heagerty PJ, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. Second ed. Oxford, UK: Oxford University Press; 2002.
43. Lee Y, Nelder JA. Conditional and marginal models: another view (with discussion). *Stat Sci*. 2004;19(2):219-228.
44. Mastrantonio G. The joint projected normal and skew-normal: a distribution for poly-cylindrical data. *J Multivar Anal*. 2018;165:14-26.
45. Mardia KV, Sutton TW. A model for cylindrical variables with applications. *J R Stat Soc B*. 1978;40(2):229-233.
46. Abe T, Ley C. A tractable, parsimonious and flexible model for cylindrical data, with applications. *Econ Stat*. 2017;4:91-104.
47. Cremers J, Pennings HJM, Ley C. Regression models for cylindrical data in psychology. *Multivar Behav Res*. 2020;55(6):910-925.
48. Lagona F. Regression analysis of correlated circular data based on the multivariate von Mises distribution. *Environ Ecol Stat*. 2016;23(1):89-113.
49. Maruotti A, Punzo A, Mastrantonio G, Lagona F. A time-dependent extension of the projected normal regression model for longitudinal circular data based on a hidden Markov heterogeneity structure. *Stoch Environ Res Risk Assess*. 2016;30(6):1725-1740.
50. Lagona F. A copula-based hidden Markov model for toroidal time series. In: Petrucci A, Racioppi F, Verde R, eds. *New Statistical Developments in Data Science*. Cham, CH: Springer; 2019:435-446.
51. Rivest LP, Kato S. A random-effects model for clustered circular data. *Can J Stat*. 2019;47:712-728.
52. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: linear and nonlinear mixed effects models, R package version 3; 2021:1-153. <https://CRAN.R-project.org/package=nlme>.
53. Genz A, Bretz F, Miwa T, et al. mvtnorm: Multivariate normal and *t* distributions, R package version 1; 2021:1-3. <https://CRAN.R-project.org/package=mvtnorm>.
54. Nolan J. SphericalCubature: numerical integration over spheres and balls in *n*-dimensions; multivariate polar coordinates, R package version 1.5; 2021. <https://CRAN.R-project.org/package=SphericalCubature>.
55. Agostinelli C, Lund U. R package circular: circular statistics, R package version 0.4-93; 2017. <https://CRAN.R-project.org/package=circular>.
56. Geraci M, Farcomeni A. A family of linear mixed-effects models using the generalized Laplace distribution. *Stat Methods Med Res*. 2020;29(9):2665-2682.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Geraci M. Joint regression modelling of intensity and timing of accelerometer counts. *Statistics in Medicine*. 2023;42(4):579-595. doi: 10.1002/sim.9633