

## A NONLINEAR MIXED EFFECTS DIRECTIONAL MODEL FOR THE ESTIMATION OF THE ROTATION AXES OF THE HUMAN ANKLE

BY MOHAMMED HADDOU, LOUIS-PAUL RIVEST<sup>1</sup> AND  
MICHAEL PIERRYNOWSKI

*Université Laval, Université Laval and McMaster University*

This paper suggests a nonlinear mixed effects model for data points in  $SO(3)$ , the set of  $3 \times 3$  rotation matrices, collected according to a repeated measure design. Each sample individual contributes a sequence of rotation matrices giving the relative orientations of the right foot with respect to the right lower leg as its ankle moves. The random effects are the five angles characterizing the orientation of the two rotation axes of a subject's right ankle. The fixed parameters are the average value of these angles and their variances within the population. The algorithms to fit nonlinear mixed effects models presented in Pinheiro and Bates (2000) are adapted to the new directional model. The estimation of the random effects are of interest since they give predictions of the rotation axes of an individual ankle. The performance of these algorithms is investigated in a Monte Carlo study. The analysis of two data sets is presented. In the biomechanical literature, there is no consensus on an in vivo method to estimate the two rotation axes of the ankle. The new model is promising. The estimates obtained from a sample of volunteers are shown to be in agreement with the clinically accepted results of Inman (1976), obtained by manipulating cadavers. The repeated measure directional model presented in this paper is developed for a particular application. The approach is, however, general and might be applied to other models provided that the random directional effects are clustered around their mean values.

**1. Introduction.** The human ankle joint complex has been modeled as a two fixed axis mechanism. It is the primary joint involved in the motion of the rearfoot with respect to the lower leg. The characterization of walking disorders associated with cerebral palsy, clubfoot or flatfoot deformities uses altered external moments (torques) about the two rotation axes of the ankle. An accurate and reliable determination of the orientation of these two axes is important to successfully evaluate and treat patients with these conditions. There is no consensus, in the biomechanical literature, on a noninvasive method for estimating the location and orientation of these ankle axes in a live individual.

---

Received September 2009; revised January 2010.

<sup>1</sup>Supported in part by the Natural Sciences and Engineering Research Council of Canada and of the Canada Research Chair in Statistical Sampling and Data Analysis.

*Key words and phrases.* Mixed effects model, penalized likelihood, Bayesian analysis, directional data, rotation matrices, joint kinematics.

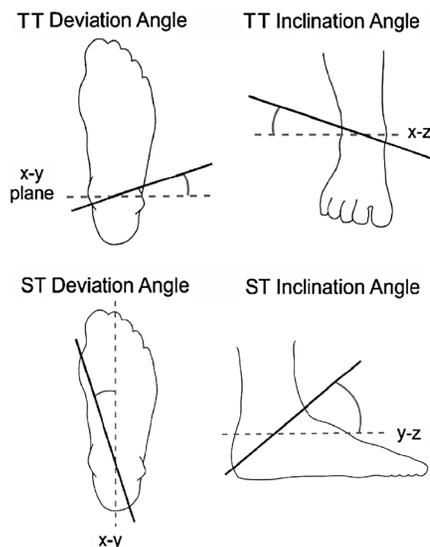


FIG. 1. The deviation and inclination angles of the tibiotarsal (TT) and subtalar (ST) human ankle rotation axes.

The two rotation axes of the ankle can be recorded in an RFU coordinate system where the  $x$ -axis points Right, the  $y$ -axis goes Forward and the  $z$ -axis goes Up. Anatomically, plantarflexion–dorsiflexion occurs about the tibiotarsal, or  $tt$  axis, which is attached to the lower leg, while the subtalar, or  $st$  axis, attached to the calcaneus, is used for the supination–pronation motion of the foot. These two axes are presented in Figure 1. Their orientations are determined by four anatomical angles ( $ttinc$ ,  $ttdev$ ,  $stinc$ ,  $stdev$ ) giving the inclinations and the deviations of these two axes referenced to the RFU coordinate system.

Using an average generic orientation for each axis results in substantial errors [Lewis et al. (2009)] because of an important between subject variation in axis locations as characterized in Section 7.1 below. Several in vivo estimation methods have been proposed in biomechanical journals [see, for instance, van den Bogert, Smith and Nigg (1994) and Lewis et al. (2009)], but none was completely successful at estimating the angles in Figure 1. The poor numerical results obtained by Lewis, Sommer and Piazza (2006) led them to question the validity of the two-axis model for the ankle.

The in vivo estimation of the orientation of the two axes of the ankle is a statistical problem. The data set for one individual is a sequence of  $3 \times 3$  rotation matrices giving the relative orientation of the foot of the subject with respect to its lower leg as its ankle moves up and down, right and left, as much as possible. Rivest, Baillargeon and Pierrynowski (2008) and van den Bogert, Smith and Nigg (1994) provide more details about data collection.

Following van den Bogert, Smith and Nigg (1994), Rivest, Baillargeon and Pierrynowski (2008) developed a statistical model for analyzing the rotation data col-

lected on a single ankle. Its parameters are the four angles defined in Figure 1 and a fifth one for the relative position of the two rotation axes. This model fits well; the residual standard deviations reported in Rivest, Baillargeon and Pierrynowski (2008) are around one degree. For a subject whose ankle has an average range of motion, the estimates are, however, not repeatable. Two data sets collected on the same ankle in similar conditions can give different estimated angles. This occurs because the likelihood function does not have a clear maximum; it has a plateau and some angles cannot be estimated independently of the others. Indeed, Rivest, Baillargeon and Pierrynowski (2008) demonstrate that only three parameters can be reliably estimated in a subject with an average ankle range of motion. Considering the small range of the residual angles, a failure of the two-axis model is an unlikely cause for the poor repeatability of the results.

This paper suggests methods to improve the numerical stability of the estimates derived from the ankle model. A penalized likelihood is proposed for estimating the parameters. The penalty is obtained by assuming a prior multivariate normal distribution for the five angle parameters. When the mean and the variance covariance matrix of the prior distribution are not known, one is confronted with a nonlinear mixed effects directional model whose parameters can be estimated using ankle's data collected on a sample of volunteers. Two numerical algorithms to fit this model are proposed. Their performances are evaluated in a Monte Carlo experiment; two data sets are then analyzed using the new nonlinear mixed effects directional model.

Nonlinear mixed effects vary with the parametrization of the random effects. Thus, the first step of the analysis is to parameterize the model of Rivest, Baillargeon and Pierrynowski (2008) in terms of the angles presented in Figure 1 and to derive inference techniques using this parametrization. These procedures are then generalized to a Bayesian setting obtained by multiplying the likelihood by the prior distribution of the model parameters. The algorithms of Lindstrom and Bates (1990) [see also Pinheiro and Bates (2000)] for fitting nonlinear mixed effects models are then adapted to the new directional model.

Nonlinear mixed effects and Bayesian models with concentrated prior distributions could potentially be used in many problems of directional statistics. These techniques could, for instance, be applied to the spherical regression model considered by Kim (1991) and Bingham, Chang and Richards (1992) and to the directional one way ANOVA model of Rancourt, Rivest and Asselin (2000) to characterize the between subject variability of the mean rotation.

This new statistical methodology has important applications in biomechanics. By fitting a nonlinear mixed effects directional model to data collected on a sample of volunteers, one is able to estimate the mean values and the between subject variances of the five anatomical angles in the population. These estimates are found to be close to the clinically accepted results obtained by Inman (1976) who used direct measurements from cadaveric feet. This provides an empirical validation of the two-axis model of van den Bogert, Smith and Nigg (1994) for the ankle. The

new model also allows the analysis of the rotation data collected by Pierrynowski et al. (2003) which compared the orientations of the subtalar axis of two groups of individuals classified according to their lower extremity injuries. Finally, the penalized predictions associated to the mixed effect model are shown to be more stable than the estimates obtained by maximizing the standard unpenalized likelihood for one subject.

**2. Parameterization of unit vectors and of rotation matrices using anatomical angles.** Let  $\mathbf{A}_1$  and  $\mathbf{B}_2$  be  $3 \times 1$  unit vectors giving the tibiotarsal and the subtalar rotation axes in a coordinate system defined according to the RFU convention. These unit vectors are first expressed in terms of the anatomical angles. Then the Cardan angle decomposition for a  $3 \times 3$  rotation matrix is briefly reviewed. This section uses the arctan function with two arguments, such that  $\arctan(a, b)$  is the angle whose sine and cosine are given by  $a/\sqrt{a^2 + b^2}$  and  $b/\sqrt{a^2 + b^2}$ , respectively.

First consider the tibiotarsal axis,  $\mathbf{A}_1 = (A_{11}, A_{21}, A_{31})^\top$ , where “ $\top$ ” denotes a matrix transpose. Formally, the anatomical angles are defined as  $ttinc = t_1 = -\arctan(A_{31}, A_{11})$  and  $tdev = t_2 = \arctan(A_{21}, A_{11})$ . Without loss of generality, we assume that the first coordinate of  $\mathbf{A}_1$  is positive. A general expression for unit vectors in the half unit sphere is

$$(2.1) \quad \mathbf{A}_1 = \frac{1}{D_t} \begin{pmatrix} \cos(t_1) \\ \cos(t_1) \tan(t_2) \\ -\sin(t_1) \end{pmatrix}, \quad t_1, t_2 \in [-\pi/2, \pi/2),$$

where  $D_t = \sqrt{1 + \cos^2(t_1) \tan^2(t_2)}$ . In a similar manner, one can parameterize the subtalar axis in terms of the anatomical angles  $stinc = s_1 = \arctan(B_{32}, B_{22})$  and  $stdev = s_2 = -\arctan(B_{12}, B_{22})$  as follows,

$$(2.2) \quad \mathbf{B}_2 = \frac{1}{D_s} \begin{pmatrix} -\cos(s_1) \tan(s_2) \\ \cos(s_1) \\ \sin(s_1) \end{pmatrix}, \quad s_1, s_2 \in [-\pi/2, \pi/2),$$

where  $D_s = \sqrt{1 + \cos^2(s_1) \tan^2(s_2)}$ .

The set  $SO(3)$  of  $3 \times 3$  rotation matrices is a three-dimensional manifold whose properties are reviewed in McCarthy (1990), Chirikjian and Kyatkin (2001) and León, Massé and Rivest (2006). This paper uses the Cardan angle parametrization with the  $X - Z - Y$  convention. It expresses an element of  $SO(3)$  as a function of the Cardan angles  $\alpha \in [-\pi, \pi)$ ,  $\gamma \in [-\pi/2, \pi/2)$  and  $\phi \in [-\pi, \pi)$  as

$$(2.3) \quad \mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{pmatrix} \\ = \begin{pmatrix} \cos \gamma \cos \phi & -\sin \gamma & \cos \gamma \sin \phi \\ \dots & \cos \alpha \cos \gamma & \dots \\ \dots & \sin \alpha \cos \gamma & \dots \end{pmatrix},$$

where  $\dots$  stands for complex trigonometric expressions that are not used in the sequel. We also write  $\mathbf{R} = \mathbf{R}(\alpha, x) \times \mathbf{R}(\gamma, z) \times \mathbf{R}(\phi, y)$ , where the arguments of  $\mathbf{R}(\cdot, \cdot)$  are the angle and the axis of the rotation respectively.

The model presented in the next section uses rotation matrices,  $\mathbf{A}(t_1, t_2)$  and  $\mathbf{B}(s_1, s_2)$ , whose first and second columns are respectively equal to  $\mathbf{A}_1$  and  $\mathbf{B}_2$ . These matrices are given by

$$\begin{aligned} \mathbf{A}(t_1, t_2) &= \mathbf{R}(t_1, y)\mathbf{R}[\arctan\{\cos(t_1) \tan(t_2), 1\}, z] \\ &= \frac{1}{D_t} \begin{pmatrix} \cos(t_1) & -\cos^2(t_1) \tan(t_2) & \sin(t_1)D_t \\ \cos(t_1) \tan(t_2) & 1 & 0 \\ -\sin(t_1) & \sin(t_1) \cos(t_1) \tan(t_2) & \cos(t_1)D_t \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} \mathbf{B}(s_1, s_2) &= \mathbf{R}(s_1, x)\mathbf{R}[\arctan\{\cos(s_1) \tan(s_2), 1\}, z] \\ &= \frac{1}{D_s} \begin{pmatrix} 1 & -\cos(s_1) \tan(s_2) & 0 \\ \cos^2(s_1) \tan(s_2) & \cos(s_1) & -\sin(s_1)D_s \\ \sin(s_1) \cos(s_1) \tan(s_2) & \sin(s_1) & \cos(s_1)D_s \end{pmatrix}, \end{aligned}$$

where  $D_t$  and  $D_s$  are defined in (2.1) and (2.2), respectively.

**3. The model for estimating the rotation axes of a single ankle.** This section expresses the model of Rivest, Baillargeon and Pierrynowski (2008) for the estimation of the anatomical angles for a single subject in terms of the rotation matrices  $\mathbf{A}(t_1, t_2)$  and  $\mathbf{B}(s_1, s_2)$ . The data set is a sequence of time ordered  $3 \times 3$  rotation matrices  $\{\mathbf{R}_i : i = 1, \dots, n\}$ . The model for  $\mathbf{R}_i$  involves the four anatomical angles, rotation angles  $\{\alpha_i : i = 1, \dots, n\}$  and  $\{\phi_i : i = 1, \dots, n\}$  in  $[-\pi, \pi)$  about the two rotation axes and the angle  $\gamma_0 \in (-\pi/2, \pi/2)$ , related to the relative position of the two axes. The predicted value  $\Psi_i$  for  $\mathbf{R}_i$  is given by

$$(3.1) \quad \Psi_i = \mathbf{A}(t_1, t_2)\mathbf{R}(\alpha_i, x)\mathbf{R}(\gamma_0, z)\mathbf{R}(\phi_i, y)\mathbf{B}(s_1, s_2)^\top.$$

The errors are assumed to have a symmetric Fisher–von Mises matrix distribution with density  $f(\mathbf{E}) = \exp\{\kappa \text{tr}(\mathbf{E})\}/c_\kappa$ ,  $\mathbf{E} \in SO(3)$ , where  $c_\kappa$  is a normalizing constant; see Mardia and Jupp (2000). If the parameter  $\kappa$  is assumed to be large so that the error rotations are clustered around the identity matrix  $\mathbf{I}_3$ , one has

$$(3.2) \quad \mathbf{E} = \mathbf{I}_3 + \begin{pmatrix} 0 & -\varepsilon_3 & \varepsilon_2 \\ \varepsilon_3 & 0 & -\varepsilon_1 \\ -\varepsilon_2 & \varepsilon_1 & 0 \end{pmatrix} + O_p\left(\frac{1}{\kappa}\right),$$

where the entries  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  of the skew-symmetric matrix have independent  $\mathcal{N}\{0, 1/(2\kappa)\}$  distributions;  $1/(2\kappa)$  is called the residual variance. The model postulates that  $\mathbf{R}_i = \Psi_i \mathbf{E}_i$ , for  $i = 1, \dots, n$ . The likelihood is  $L[t_1, t_2, s_1, s_2, \gamma_0, \kappa,$

$\{\alpha_i\}, \{\phi_i\} = \prod_i f(\Psi_i^\top \mathbf{R}_i)$ . Rivest, Baillargeon and Pierrynowski (2008) show that the angles  $\{\alpha_i\}$  and  $\{\phi_i\}$  can be profiled out of the likelihood. Indeed,

$$(3.3) \quad \begin{aligned} L[t_1, t_2, s_1, s_2, \gamma_0, \kappa, \{\alpha_i\}, \{\phi_i\}] &\leq L_p(t_1, t_2, s_1, s_2, \gamma_0, \kappa) \\ &= \frac{1}{c_\kappa^n} \exp \left[ \kappa \sum_{i=1}^n \{2 \cos(\theta_i^z - \gamma_0) + 1\} \right], \end{aligned}$$

where  $\theta_i^z = -\arcsin(\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)$  is the  $Z$ -Cardan angle of  $\mathbf{A}(t_1, t_2)^\top \mathbf{R}_i \mathbf{B}(s_1, s_2)$  in the  $X - Z - Y$  convention; see (2.3).

Several methods are available to maximize (3.3). However, for the implementation of the mixed effects directional model, a closed form expression for the score vector for  $\boldsymbol{\beta} = (t_1, t_2, s_1, s_2, \gamma_0)^\top$  is needed. This is derived now. Observe that  $\cos(\theta) = 1 - 2 \sin^2(\theta/2)$ , where  $\theta/2$  can be assumed to lie in the interval  $[-\pi/2, \pi/2)$ . The log profile likelihood is equal to

$$\log L_p(t_1, t_2, s_1, s_2, \gamma_0, \kappa) = -4\kappa \sum_{i=1}^n \sin^2\left(\frac{\theta_i^z - \gamma_0}{2}\right) - n \log c_\kappa + 3n\kappa.$$

The score for  $\gamma_0$  is easily evaluated, viz.

$$\frac{\partial}{\partial \gamma_0} \log L_p(t_1, t_2, s_1, s_2, \gamma_0, \kappa) = 4\kappa \sum_{i=1}^n \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \cos\left(\frac{\theta_i^z - \gamma_0}{2}\right).$$

The score for the four remaining parameters involve the following partial derivatives that are evaluated using elementary methods. The property that the partial derivatives of a unit vector and the unit vector itself are orthogonal was used to get the following results:

$$\begin{aligned} \frac{\partial}{\partial t_1} \mathbf{A}_1 &= -\frac{\sin(t_1) \tan(t_2)}{D_t^2} \mathbf{A}_2 - \frac{1}{D_t} \mathbf{A}_3, \\ \frac{\partial}{\partial t_2} \mathbf{A}_1 &= \frac{\cos(t_1) \{1 + \tan^2(t_2)\}}{D_t^2} \mathbf{A}_2, \\ \frac{\partial}{\partial s_1} \mathbf{B}_2 &= \frac{\sin(s_1) \tan(s_2)}{D_s^2} \mathbf{B}_1 + \frac{1}{D_s} \mathbf{B}_3, \\ \frac{\partial}{\partial s_2} \mathbf{B}_2 &= -\frac{\cos s_1 \{1 + \tan^2(s_2)\}}{D_s^2} \mathbf{B}_1. \end{aligned}$$

Since  $\theta_i^z = -\arcsin(\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)$ , the score for  $t_1$  is given by

$$\begin{aligned} \frac{\partial}{\partial t_1} \log L_p(t_1, t_2, s_1, s_2, \gamma_0, \kappa) &= -4\kappa \sum_{i=1}^n \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \cos\left(\frac{\theta_i^z - \gamma_0}{2}\right) \\ &\quad \times \frac{-1}{\sqrt{1 - (\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2}} \frac{\partial}{\partial t_1} \mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2. \end{aligned}$$

This can be evaluated using the previous expressions for the partial derivatives. Repeating this for the other anatomical angles leads to

$$\begin{aligned}
 \frac{\partial}{\partial \boldsymbol{\beta}} \log L_p(\boldsymbol{\beta}, \kappa) &= -4\kappa \sum_{i=1}^n \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \cos\left(\frac{\theta_i^z - \gamma_0}{2}\right) \frac{\partial}{\partial \boldsymbol{\beta}} (\theta_i^z - \gamma_0) \\
 (3.4) \qquad \qquad \qquad &= -4\kappa \sum_{i=1}^n \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \mathbf{X}_i,
 \end{aligned}$$

where

$$\begin{aligned}
 \mathbf{X}_i &= -\frac{\cos\{(\theta_i^z - \gamma_0)/2\}}{\sqrt{1 - (\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2}} \\
 (3.5) \qquad \qquad \qquad &\times \begin{pmatrix} -\frac{\sin(t_1) \tan(t_2)}{D_t^2} \mathbf{A}_2^\top \mathbf{R}_i \mathbf{B}_2 - \frac{1}{D_t} \mathbf{A}_3^\top \mathbf{R}_i \mathbf{B}_2 \\ \frac{\cos(t_1)\{1 + \tan^2(t_2)\}}{D_t^2} \mathbf{A}_2^\top \mathbf{R}_i \mathbf{B}_2 \\ \frac{\sin(s_1) \tan(s_2)}{D_s^2} \mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_1 + \frac{1}{D_s} \mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_3 \\ -\frac{\cos(s_1)\{1 + \tan^2(s_2)\}}{D_s^2} \mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_1 \\ \sqrt{1 - (\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2} \end{pmatrix}.
 \end{aligned}$$

Evaluated at  $\boldsymbol{\beta} + \delta(\boldsymbol{\beta})$ , where  $\delta(\boldsymbol{\beta})$  is a  $5 \times 1$  vector with entries close to 0, the score vector (3.4) is

$$\begin{aligned}
 &-4\kappa \sum_{i=1}^n \left\{ \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \mathbf{X}_i + \frac{1}{2} \mathbf{X}_i \mathbf{X}_i^\top \delta(\boldsymbol{\beta}) \right\} \\
 &\quad + O(\|\delta(\boldsymbol{\beta})\|^2) + O(\max[\|\sin\{(\theta_i^z - \gamma_0)/2\}\| \|\delta(\boldsymbol{\beta})\|]).
 \end{aligned}$$

One can consider that the last two terms are negligible since the residuals  $(\theta_i^z - \gamma_0)$  are small. This is standard in the large  $\kappa$  asymptotics used to approximate the sampling distributions of estimators in a directional model: both  $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}$  and the errors  $\varepsilon_j$  in (3.2) are assumed to be  $O(1/\sqrt{\kappa})$ ; see, for instance, Rivest and Chang (2006). Equating this to 0 yields a simple updating formula. Given its current value  $\boldsymbol{\beta}$ , the updated value is  $\boldsymbol{\beta} + \delta(\boldsymbol{\beta})$ , where

$$\delta(\boldsymbol{\beta}) = -\left(\sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top\right)^{-1} \sum_{i=1}^n \left\{ 2 \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \mathbf{X}_i \right\}.$$

This calculation can be carried out by regressing the residual vector  $[2 \sin\{(\theta_i^z - \gamma_0)/2\}]$  on the explanatory variables  $\{\mathbf{X}_i\}$ . Theorem 1 of Rivest, Baillargeon and

Pierrynowski (2008) holds and, as  $\kappa$  goes to  $\infty$ , the maximum likelihood estimator  $\hat{\boldsymbol{\beta}}$  is approximately normally distributed. Once the model is fitted, the residual variance  $1/(2\kappa)$  can be estimated using the sum of the squared residuals,

$$\frac{1}{2\hat{\kappa}} = \frac{1}{n} \sum_{i=1}^n 4 \sin^2\{(\hat{\theta}_i^z - \hat{\gamma}_0)/2\},$$

where the residual angle  $\hat{\theta}_i^z = -\arcsin(\hat{\mathbf{A}}_1^\top \mathbf{R}_i \hat{\mathbf{B}}_2)$  is the Z-Cardan angle of  $\hat{\mathbf{A}}_1^\top \mathbf{R}_i \hat{\mathbf{B}}$  in the X – Z – Y convention. Using Grood and Suntay (1983) clinical interpretation, rotations of angle  $\{\hat{\theta}_i^z\}$  occur about a floating axis that is orthogonal to both the  $tt$  and the  $st$  axes. Plots of these angles appear in Figure 2 of Rivest, Baillargeon and Pierrynowski (2008); their residual standard deviation,  $\sqrt{1/2\hat{\kappa}}$ , is about one degree. The distribution of the residuals  $\{[2 \sin\{(\theta_i^z - \gamma_0)/2\}]\}$  is usually approximately normal; the normality assumption in (3.2) is not violated for most of the data sets investigated. Thus, the proposed model fits well to the ankle data.

Many individuals have a limited ankle range of motion; the domains for angles  $\{\alpha_i\}$  and  $\{\phi_i\}$  in (3.1) are therefore limited. This makes the estimates of the anatomical angles numerically unstable. There can be important differences between the estimates calculated on two data sets collected in succession on the same ankle; see Rivest, Baillargeon and Pierrynowski (2008). Thus, individual measurements do not allow the estimation of a complete set of anatomical angles. This suggests to borrow strength from other individuals and to consider a Bayesian model whose prior distribution penalizes extreme parameter values.

**4. A Bayesian ankle model.** Assume, for now, that the residual variance  $1/(2\kappa)$  is known and that the  $5 \times 1$  vector of anatomical angles  $\boldsymbol{\beta}$  is random and has a  $\mathcal{N}_5(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ , where  $\boldsymbol{\beta}_0$  is the average vector of anatomical angles within the population and the  $5 \times 5$  variance covariance matrix  $\boldsymbol{\Sigma}_0$  characterizes the variability of the anatomical angles within the population; both are assumed to be known. Elements of  $\boldsymbol{\beta}_0$  and  $\boldsymbol{\Sigma}_0$  could be set equal to the values of Inman (1976), who studied the variability of these angles. We assume that  $\boldsymbol{\Sigma}_0$  is  $O(1/(2\kappa))$ , thus, there is a fixed  $5 \times 5$  upper triangular matrix  $\boldsymbol{\Delta}_0$  such that  $\boldsymbol{\Sigma}_0 = \boldsymbol{\Delta}_0^{-1}(\boldsymbol{\Delta}_0^{-\top})/(2\kappa)$ , where  $\boldsymbol{\Delta}_0^{-\top}$  is the inverse of  $\boldsymbol{\Delta}_0^\top$ . This section presents an algorithm to derive the mode of the posterior distribution of  $\boldsymbol{\beta}$  and suggests an approximation for its posterior distribution.

The posterior distribution of  $\boldsymbol{\beta}$  is proportional to

$$\begin{aligned} &\exp\left[-\kappa\left\{\sum_{i=1}^n 4 \sin^2\left(\frac{\theta_i^z - \gamma_0}{2}\right) + (\boldsymbol{\beta} - \boldsymbol{\beta}_0)^\top \boldsymbol{\Delta}_0^\top \boldsymbol{\Delta}_0 (\boldsymbol{\beta} - \boldsymbol{\beta}_0)\right\}\right] \\ &= \exp\{-\kappa SSE(\boldsymbol{\beta})\}. \end{aligned}$$

The posterior mode  $\hat{\boldsymbol{\beta}}$  is the value of  $\boldsymbol{\beta}$  that minimizes  $SSE$ . It can be evaluated by adapting the regression algorithm of Section 3 to the Bayesian framework. The

vector of partial derivatives of  $SSE$  with respect to  $\boldsymbol{\beta}$  is easily derived,

$$(4.1) \quad \frac{\partial}{\partial \boldsymbol{\beta}} SSE(\boldsymbol{\beta}) = 4 \sum_{i=1}^n \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \mathbf{X}_i + 2\Delta_0^\top \Delta_0(\boldsymbol{\beta} - \boldsymbol{\beta}_0),$$

where the vector of partial derivatives  $\mathbf{X}_i$  is defined by (3.5).

Proceeding as in Section 3, one constructs an algorithm for minimizing  $SSE$ . The current value  $\boldsymbol{\beta}$  is updated to  $\boldsymbol{\beta} + \delta(\boldsymbol{\beta})$ , where

$$\delta(\boldsymbol{\beta}) = -\left(\sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top + \Delta_0^\top \Delta_0\right)^{-1} \left[ \sum_{i=1}^n \left\{ 2 \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) \mathbf{X}_i \right\} + \Delta_0^\top \Delta_0(\boldsymbol{\beta} - \boldsymbol{\beta}_0) \right].$$

An alternative expression for the updated value is

$$\begin{aligned} \boldsymbol{\beta} + \delta(\boldsymbol{\beta}) &= \left(\sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^\top + \Delta_0^\top \Delta_0\right)^{-1} \\ &\quad \times \left[ \sum_{i=1}^n \left\{ -2 \sin\left(\frac{\theta_i^z - \gamma_0}{2}\right) + \mathbf{X}_i^\top \boldsymbol{\beta} \right\} \mathbf{X}_i + \Delta_0^\top \Delta_0 \boldsymbol{\beta}_0 \right]. \end{aligned}$$

An approximation to the posterior distribution of the anatomical angles is given next.

**PROPOSITION 1.** *As  $\kappa \rightarrow \infty$ , the posterior distribution for  $\boldsymbol{\beta}$  satisfies*

$$\sqrt{2\kappa}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \sim \mathcal{N}_5 \left\{ 0, \left( \sum_{i=1}^n \hat{\mathbf{X}}_i \hat{\mathbf{X}}_i^\top + \Delta_0^\top \Delta_0 \right)^{-1} \right\},$$

where  $\hat{\mathbf{X}}_i$  denotes the  $5 \times 1$  vector of partial derivatives defined by (3.5) and evaluated at  $\hat{\boldsymbol{\beta}}$ .

The posterior density of  $\delta = \sqrt{2\kappa}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})$  is proportional to  $\exp\{-\kappa SSE(\hat{\boldsymbol{\beta}} + \delta/\sqrt{2\kappa})\}$ . The result is proved by taking a second-order Taylor series expansion around  $\delta = 0$ . The first order derivatives are null and the variance covariance matrix of Proposition 1 is obtained by dropping the  $o\{1/(2\kappa)\}$  terms in the matrix of second-order derivatives.

We now study some frequentist properties of  $\hat{\boldsymbol{\beta}}$ . Let  $\boldsymbol{\beta}^{(t)}$  be the true values of the anatomical angles for the individual under consideration. Thus,  $\boldsymbol{\beta}^{(t)}$  is a realization of the  $\mathcal{N}_5(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$  prior distribution, such that  $\|\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}_0\|$  is  $O(1/\sqrt{2\kappa})$ . The difference  $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^{(t)}$  is  $O(1/\sqrt{2\kappa})$ . The leading term of this difference consists of a linear combination of individual experimental errors and of the penalty associated

to the prior distribution. To get a closed form expression, one can proceed as in Appendix B of Rivest, Baillargeon and Pierrynowski (2008). It suffices to carry out a first-order Taylor series expansion of (4.1) in terms of the difference  $\delta(\boldsymbol{\beta}) = \boldsymbol{\beta} - \boldsymbol{\beta}^{(t)}$  and of the experimental errors. This yields

$$\begin{aligned}
 & \frac{\partial}{\partial \boldsymbol{\beta}} SSE(\boldsymbol{\beta}^{(t)} + \delta(\boldsymbol{\beta})) \\
 (4.2) \quad &= 2 \left\{ \sum_{i=1}^n \epsilon_i \mathbf{X}_i^{(t)} + \sum_{i=1}^n \mathbf{X}_i^{(t)} \mathbf{X}_i^{(t)T} \delta(\boldsymbol{\beta}) + \Delta_0^\top \Delta_0 (\delta(\boldsymbol{\beta}) + \boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}_0) \right\} \\
 &+ O(1/\kappa),
 \end{aligned}$$

where  $\epsilon_i$  is a  $\mathcal{N}(0, 1/(2\kappa))$  random variable that depends on the error matrix  $\mathbf{E}_i$ , and  $\mathbf{X}_i^{(t)}$  denotes the vector of partial derivatives  $\mathbf{X}_i$  evaluated at the true value  $\boldsymbol{\beta}^{(t)}$ , with  $R_i$  set equal to  $\Psi_i$  in (3.5). Now  $\hat{\boldsymbol{\beta}}$  corresponds to the value of  $\delta(\boldsymbol{\beta})$  for which (4.2) is null, thus,

$$\begin{aligned}
 \hat{\boldsymbol{\beta}} &= \boldsymbol{\beta}^{(t)} - \left( \sum_{i=1}^n \mathbf{X}_i^{(t)} \mathbf{X}_i^{(t)T} + \Delta_0^\top \Delta_0 \right)^{-1} \\
 &\quad \times \left\{ \sum_{i=1}^n \epsilon_i \mathbf{X}_i^{(t)} + \Delta_0^\top \Delta_0 (\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}_0) \right\} + O(1/\kappa) \\
 &= \left( \sum_{i=1}^n \mathbf{X}_i^{(t)} \mathbf{X}_i^{(t)T} + \Delta_0^\top \Delta_0 \right)^{-1} \\
 &\quad \times \left\{ \sum_{i=1}^n (\mathbf{X}_i^{(t)T} \boldsymbol{\beta}^{(t)} - \epsilon_i) \mathbf{X}_i^{(t)} + \Delta_0^\top \Delta_0 \boldsymbol{\beta}_0 \right\} + O(1/\kappa).
 \end{aligned}$$

This expansion provides an approximation for the prediction error of  $\hat{\boldsymbol{\beta}}$  as a predictor of  $\boldsymbol{\beta}^{(t)}$ . The posterior variance covariance matrix of  $\boldsymbol{\beta}^{(t)}$  given in Proposition 1 is an estimate of the variance covariance matrix of the approximate prediction error.

**5. A mixed model for the simultaneous estimation of several sets of anatomical angles.** We now have  $M$  subjects and  $n_i, 1 \leq i \leq M$ , observed rotation matrices on each subject. The data set consists of the  $3 \times 3$  rotation matrices  $\{\mathbf{R}_{ij} : i = 1, \dots, M; j = 1, \dots, n_i\}$ . Let  $\boldsymbol{\beta}_i = (t_{1i}, t_{2i}, s_{1i}, s_{2i}, \gamma_{0i})^\top$  be the anatomical angles for the  $i$ th ankle. As in Section 4, the angles  $\boldsymbol{\beta}_i$  are assumed to be random deviates with a five-dimensional normal distribution,  $\mathcal{N}_5(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ . The fixed parameters  $\kappa, \boldsymbol{\beta}_0$  and  $\boldsymbol{\Sigma}_0$  are assumed to be unknown.

In a mixed effects model, the fixed regression parameters and the variance components are estimated using a marginal likelihood. To estimate  $\boldsymbol{\beta}_0$  and  $\boldsymbol{\Sigma}_0$ , we

construct such a likelihood using the profile likelihood defined in (3.3), rather than the full likelihood for the ankle model. This is acceptable since this profile likelihood is also a likelihood, constructed by assuming that the angles  $\theta_{ij}^z - \gamma_{0i}$  defined in (3.3) have a centered angular von Mises distribution with shape parameter  $2\kappa$ . Indeed, since  $\kappa$  is large, the distribution of  $2 \sin\{(\theta_{ij}^z - \gamma_{0i})/2\}$  is approximately  $\mathcal{N}\{0, 1/(2\kappa)\}$ . Using this approximation in the evaluation of the marginal likelihood  $L_{1p}$  gives

$$(5.1) \quad \prod_{i=1}^M \left(\frac{\kappa}{\pi}\right)^{(n_i+5)/2} |\mathbf{\Delta}_0| \times \int_{\mathbb{R}^5} \exp\left\{-\kappa \sum_{j=1}^{n_i} 4 \sin^2\left(\frac{\theta_{ij}^z - \gamma_{0i}}{2}\right) - \kappa \|\mathbf{\Delta}_0(\boldsymbol{\beta}_i - \boldsymbol{\beta}_0)\|^2\right\} d\boldsymbol{\beta}_i.$$

Because of a highly nonlinear integrand, the above expression cannot be evaluated explicitly. To find the maximum value numerically, we adapt the algorithm of Lindstrom and Bates (1990) to the directional ankle model.

For each  $i$ , and for fixed  $(\boldsymbol{\beta}_0, \mathbf{\Delta}_0)$ , the integrand of (5.1) is maximized using the method presented in Section 3. Let  $\hat{\boldsymbol{\beta}}_i$  be the maximum value for the  $i$ th sample. Using the first order Taylor series expansion derived in Section 3 leads to the following approximation:

$$2 \sin\left(\frac{\theta_{ij}^z - \gamma_{0i}}{2}\right) \approx 2 \sin\left(\frac{\hat{\theta}_{ij}^z - \hat{\gamma}_{0i}}{2}\right) + \mathbf{X}_{ij}^\top (\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i),$$

where the angles  $\hat{\theta}_{ij}^z$  and  $\hat{\gamma}_{0i}$  and the  $5 \times 1$  vector of partial derivatives  $\mathbf{X}_{ij}$  are evaluated at  $\hat{\boldsymbol{\beta}}_i$ . Using this approximation, and changing variables  $\boldsymbol{\beta}_i - \boldsymbol{\beta}_0 = \mathbf{z}$  in the integral, (5.1) becomes

$$\begin{aligned} L_{1p} &\approx \prod_{i=1}^M \left(\frac{\kappa}{\pi}\right)^{(n_i+5)/2} |\mathbf{\Delta}_0| \\ &\quad \times \int_{\mathbb{R}^5} \exp\left\{-\kappa \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \mathbf{X}_{ij}^\top \mathbf{z} - \mathbf{X}_{ij}^\top \boldsymbol{\beta}_0)^2 - \kappa \|\mathbf{\Delta}_0 \mathbf{z}\|^2\right\} d\mathbf{z}, \\ &= \prod_{i=1}^M \frac{(\kappa/\pi)^{n_i/2}}{|\mathbf{I} + \mathbf{X}_i \mathbf{\Delta}_0^{-1} \mathbf{\Delta}_0^{-\top} \mathbf{X}_i^\top|^{1/2}} \\ &\quad \times \exp\{-\kappa (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}_0)^\top (\mathbf{I} + \mathbf{X}_i \mathbf{\Delta}_0^{-1} \mathbf{\Delta}_0^{-\top} \mathbf{X}_i^\top)^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}_0)\}, \end{aligned}$$

where  $\mathbf{X}_i$  is the  $n_i \times 5$  matrix of partial derivatives for the  $i$ th unit, and  $\mathbf{y}_i$  is the  $n_i \times 1$  vector whose  $j$  entry is given by  $\mathbf{X}_{ij}^\top \hat{\boldsymbol{\beta}}_i - 2 \sin\{(\hat{\theta}_{ij}^z - \hat{\gamma}_{0i})/2\}$ . This evaluation of the integral with respect to  $\mathbf{z}$  follows the argument presented in Pinheiro

and Bates (2000), Section 7.2. It involves the five-dimensional normal density with mean vector  $(\mathbf{X}_i^\top \mathbf{X}_i + \mathbf{\Delta}_0^\top \mathbf{\Delta}_0)^{-1} \mathbf{X}_i^\top (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}_0)$  and variance covariance matrix equal to  $(\mathbf{X}_i^\top \mathbf{X}_i + \mathbf{\Delta}_0^\top \mathbf{\Delta}_0)^{-1} / (2\kappa)$ . Thus,  $L_{1p}$  is approximately equal to the likelihood function for the following linear mixed effects model:

$$(5.2) \quad \mathbf{y} = \begin{pmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_M \end{pmatrix} \boldsymbol{\beta}_0 + \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_M \end{pmatrix} \mathbf{b} + \boldsymbol{\epsilon},$$

where  $\mathbf{y}$  is the  $(\sum n_i) \times 1$  vector of the dependent variable,  $\boldsymbol{\beta}_0$  is the mean vector of the anatomical angles in the population, and  $\mathbf{b}$  is the  $5M \times 1$  vector of the individual random effects  $\mathbf{b}_i = \boldsymbol{\beta}_i - \boldsymbol{\beta}_0$  that are assumed to be independent random vectors with a  $\mathcal{N}_5(0, \boldsymbol{\Sigma}_0)$  distribution. Finally,  $\boldsymbol{\epsilon}$  is the  $(\sum n_i) \times 1$  vector of the experimental errors of the directional model containing independent  $\mathcal{N}(0, 1/(2\kappa))$  random deviates. The second step of the algorithm estimates  $\boldsymbol{\beta}_0$ ,  $1/(2\kappa)$  and  $\boldsymbol{\Sigma}_0$  by fitting the linear mixed effects model (5.2). This gives updated values for the parameters of the prior distribution that are used to get a new set of penalized estimates  $\{\hat{\boldsymbol{\beta}}_i\}$ ; these new estimates are used to get a new approximation to (5.1) and to update the marginal parameter values by fitting (5.2). This two-step algorithm typically converges after a few iterations. It is called the PLME algorithm since it uses a Penalized least squares and an algorithm for fitting Linear Mixed Effects models. Lindstrom and Bates (1990) showed that the first step can be bypassed by using  $\hat{\boldsymbol{\beta}}_i^k = \boldsymbol{\beta}_0^k + \mathbf{b}_i^k$  as the values around which the linearization of (5.1) is carried out at iteration  $k + 1$ , where  $\mathbf{b}_i^k$  is the estimate of the random effect for the  $i$ th individual at the  $k$ th iteration. This one-step algorithm is labeled LME.

This section has considered the one-sample problem, where all the individuals share the same fixed effect vector  $\boldsymbol{\beta}_0$ . Section 7 considers a two-sample model where the direction of the subtalar axis is allowed to vary between samples. The directional mixed effects model for this problem is easily constructed. The local linear mixed effects model at step 2 of the Lindstrom and Bates (1990) algorithm has a  $7 \times 1$  vector of fixed regression parameters featuring 5 entries for the mean angles in the first sample and two parameters for the between group differences of the two subtalar angles. The two algorithms proposed in this section can be used to estimate the parameters of this enlarged model.

**6. A simulation study.** This section reports the results of a Monte Carlo experiment to investigate the sampling properties of the estimators of  $\boldsymbol{\beta}_0$  and  $\boldsymbol{\Sigma}_0$  obtained by maximizing (5.1) with the two versions of the Lindstrom and Bates (1990) algorithm. First, the method used to simulate the data is reviewed, then some results will be presented. In the simulations the calculations are carried out with angles expressed in radians; for convenience the results are presented in degrees.

Simulations were carried out for the one-sample model only. Values of  $M = 30, 50, 100$  and  $n = 50, 100, 200$  were considered. The simulation used 500 Monte Carlo samples. The following parameter values were used:

$$(6.1) \quad \begin{aligned} \boldsymbol{\beta}_0 &= (8, -6, 42, 23, 17)^\top, & \boldsymbol{\Sigma}_0 &= \text{diag}(7, 4, 9, 11, 11)^2, \\ 1/(2\kappa) &= 1. \end{aligned}$$

The mean and standard deviations for the first four anatomical angles are as given by Inman (1976). The residual standard error  $1/\sqrt{2\kappa}$  of one degree was similar to estimates found in the numerical examples of Rivest, Baillargeon and Pierryowski (2008).

For each individual, the five anatomical angles were first simulated from a  $\mathcal{N}_5(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$  and the rotation matrices  $\mathbf{A}(t_1, t_2)$  and  $\mathbf{B}(s_1, s_2)$  were evaluated. To construct the predicted values  $\Psi_{ij}$  given in (3.1), angles  $(\alpha_{ij}, \phi_{ij})$  obtained by fitting the one subject model to some real data were used. The average values for  $(\alpha_{ij}, \phi_{ij})$  were (38, 14) with standard deviations of (12, 10.5). Thus, the motion about the  $st$  axis has a smaller range than that about the  $tt$  axis. The rotation errors  $\mathbf{E}_{ij}$  were generated from  $\mathbf{z} = (z_1, z_2, z_3)^\top$  three independent  $\mathcal{N}(0, 0.017^2)$  random variables (a standard deviation of 0.017 radian is 1 when expressed in degrees). Its rotation axis was set to  $\mathbf{z}/\|\mathbf{z}\|$ , while its rotation angle was equal to  $\|\mathbf{z}\|$ . To understand the numerical challenges associated to the maximization of the likelihood for the model of Section 3, it is convenient to evaluate the vector of partial derivatives  $\mathbf{X}_{ij}$  at  $\boldsymbol{\beta} = \boldsymbol{\beta}_0$  in an error free model. One gets  $\mathbf{X}_{ij} = (0.01 \cos(\alpha_{ij}) - 0.99 \sin(\alpha_{ij}), 0.99 \cos(\alpha_{ij}), 0.26 \cos(\phi_{ij}) + 0.95 \sin(\phi_{ij}), -0.80 \cos(\phi_{ij}), 1)^\top$ . The matrix  $\mathbf{X}_i$  of the vectors of partial derivatives for one subject has a condition number larger than 100. This multi-collinearity affects especially  $ttdev$ ,  $stdev$  and  $\gamma_0$ , three rotation angles about different  $z$ -axes that are not well differentiated when the ankle exhibits a small range of motion.

The simulations compared the two algorithms proposed by Lindstrom and Bates (1990), PLME and LME, as described in Section 5. The R-function `lme` was used to fit a linear mixed effects model at step 2 of the PLME algorithm. This function provides estimates of the sampling variances for the fixed effects. The biases of these variance estimators were also investigated in the Monte Carlo study. The two algorithms gave almost the same results. Only those obtained with PLME are presented.

Tables 1 and 2 report findings where the fitted model has a diagonal  $\boldsymbol{\Sigma}_0$ . In Table 1, the estimates for  $t_1$ ,  $t_2$  and  $s_1$  have small biases, not significantly different from 0, and small root mean squared errors. The estimates for  $s_2$  and  $\gamma_0$  are less precise. This is caused by the multi-collinearity problem mentioned above. The `lme` variance estimates underestimate the true variances; this underestimation is more severe for the angles  $s_2$ ,  $t_2$  and  $\gamma_0$ . Increasing  $n$ , the number of data points by subject reduces this bias. Table 2 is concerned with the estimation of the standard deviations. It shows small negative biases for all the variances. The parameters

TABLE 1

Bias and root mean squared error, in parenthesis, of the estimator of  $\beta_0$  and the relative bias of the 1me variance estimator, in parenthesis, when the fitted model assumes that  $\Sigma_0$  is diagonal

$n$	$M$	$s_1$		$s_2$		$\gamma_0$		$t_1$		$t_2$	
50	30	-0.09	(1.70, -2)	0.67	(3.10, -16)	0.55	(3.27, -14)	0.10	(1.56, -12)	-0.09	(1.26, -14)
50	50	-0.17	(1.34, -4)	0.74	(2.44, -13)	0.54	(2.74, -19)	0.03	(1.14, -2)	-0.06	(0.99, -12)
50	100	-0.13	(0.97, -9)	0.85	(1.91, -33)	0.68	(1.96, -23)	0.05	(0.85, -12)	-0.12	(0.73, -17)
100	30	-0.09	(1.58, 11)	0.26	(2.69, -10)	0.22	(2.92, -13)	0.02	(1.37, 0)	0.04	(1.09, -7)
100	50	-0.08	(1.28, 4)	0.52	(2.13, -11)	0.24	(2.17, -4)	0.06	(1.06, 2)	-0.03	(0.79, 0)
100	100	-0.07	(0.90, 4)	0.42	(1.52, -14)	0.30	(1.63, -17)	-0.09	(0.79, -8)	-0.02	(0.58, -9)
200	30	-0.01	(1.70, -5)	0.16	(2.34, -3)	0.06	(2.50, -9)	0.06	(1.38, -7)	0.05	(0.86, 9)
200	50	-0.09	(1.32, -5)	0.39	(1.87, -6)	0.27	(1.96, -8)	-0.02	(1.03, 0)	0.00	(0.65, 13)
200	100	-0.06	(0.89, 4)	0.21	(1.34, -9)	0.06	(1.44, -15)	-0.01	(0.75, -5)	0.05	(0.53, -13)

TABLE 2

Biases and root mean squared errors, in parenthesis, of the estimators of the standard deviations,  $\text{diag}(\Sigma_0^{1/2})$ , when the fitted model assumes that  $\Sigma_0$  is diagonal

$n$	$M$	$\sigma_{t_1}$	$\sigma_{t_2}$	$\sigma_{s_1}$	$\sigma_{s_2}$	$\sigma_{\gamma_0}$
50	30	-0.05 (1.10)	-0.28 (1.28)	-0.19 (1.25)	-0.04 (2.34)	-0.46 (2.50)
50	50	-0.06 (0.78)	-0.16 (0.89)	-0.12 (0.99)	-0.02 (1.74)	-0.18 (1.94)
50	100	0.00 (0.53)	-0.07 (0.63)	-0.05 (0.66)	0.03 (1.27)	-0.10 (1.26)
100	30	-0.11 (0.98)	-0.11 (0.97)	-0.07 (1.21)	-0.14 (2.09)	-0.16 (2.27)
100	50	0.00 (0.75)	-0.01 (0.67)	-0.06 (0.92)	0.13 (1.45)	-0.02 (1.66)
100	100	-0.04 (0.54)	-0.03 (0.49)	-0.05 (0.70)	-0.02 (1.12)	-0.15 (1.20)
200	30	-0.09 (0.96)	-0.06 (0.76)	-0.09 (1.14)	-0.18 (1.82)	-0.29 (1.95)
200	50	-0.05 (0.75)	-0.06 (0.58)	-0.07 (0.92)	-0.02 (1.40)	-0.12 (1.51)
200	100	-0.04 (0.54)	-0.06 (0.41)	0 (0.68)	0.03 (0.97)	-0.05 (1.07)

$\sigma_{s_2}$  and  $\sigma_{\gamma_0}$  have the largest root mean squared errors. Still, Tables 1 and 2 show that the directional mixed effects model gives reliable estimates when the random effects are assumed to be independent.

It is likely for the anatomical angles to be correlated so that the true value of  $\Sigma_0$  might not be diagonal. Inman (1976) did not consider this question; he did not report the correlations between anatomical angles. One could investigate this problem by fitting a model with an unstructured  $\Sigma_0$  with 15 parameters (that is the variances of the 5 angles and the 10 covariances between pairs of angles). Unfortunately, the results obtained with such a model are not reliable. In simulations, not reported here, biased estimates of the off-diagonal elements of  $\Sigma_0$  were obtained, especially for the covariances involving  $t_2$ ,  $s_2$  and  $\gamma_0$ . Apparently, the nonlinear mixed effects model cannot distinguish a true correlation between random effects in the population from a correlation caused by an ill-conditioned likelihood for the estimation of the random effects. Additional investigations of this problem could consider models where  $\Sigma_0$  has a small number of nonnull covariances.

**7. Numerical examples.** This section presents the analysis of two data sets collected in the Human Movement Laboratory of the School of Rehabilitation Science at McMaster University, using an OptoTrak camera system at a frequency of 50 Hz; see Rivest, Baillargeon and Pierrynowski (2008) for a detailed description of the data collection protocol. The successive rotation matrices in a data set are not independent measurements; the residual autocorrelation when fitting the model of Section 3 on the data collected on one subject is larger than 0.80. In order to satisfy, at least approximately, the assumption of independence underlying the construction of the penalized likelihood in Section 4, the model was fitted to a subsample of the data obtained by keeping one observation out of 30. A sampling frequency of 1.67 Hz yielded smaller residual autocorrelations, in the range  $-0.3$  to  $0.5$ ; the assumption of independence was approximately satisfied. Subsampling the data was also used by van den Bogert, Smith and Nigg (1994) to get stable estimates of the anatomical angles.

*7.1. An empirical validation of the estimates of Inman (1976).* The mean values and the population standard deviations of the four anatomical angles characterizing the direction of the two rotation axes of cadaver ankles were presented by Inman (1976), by manipulating unloaded cadavers feet. Inman's estimates are given in Table 3. The model of Section 5 provides an in vivo method for estimating the same angles. In this section, we use right foot data collected on  $M = 65$  volunteers with sample size  $n = 50$  rotation matrices. The estimates obtained by fitting the model of Section 5 to this data set are also presented in Table 3. This model has an estimated residual standard error,  $1/\sqrt{2\hat{k}}$  of 0.021 radians, or 1.21 degrees.

Although two of four of Inman's values are outside of the 95% confidence interval, the agreement between the two sets of estimates is reasonably good (see

TABLE 3  
*The population estimates obtained by fitting the basic model to the volunteer data set compared with the values presented by Inman (1976)*

		$t_1$	$t_2$	$s_1$	$s_2$	$\gamma_0$
Data	$\hat{\beta}_{0j}$	3.32	-8.05	38.27	20.92	22.42
	s.e.	0.89	1.52	0.89	2.50	2.37
	$\sqrt{\hat{\Sigma}_{0jj}}$	5.34	10.25	6.46	14.77	8.87
Inman	$\beta_{0j}$	8	-6	42	23	NA
	$\sqrt{\Sigma_{0jj}}$	7	4	9	11	NA

Table 3). The standard deviations are remarkably close, except possibly for the angle  $t_2 = ttdev$ . Overall, Inman’s estimates and the one derived using the directional mixed effects model with unloaded ankle motion data are similar.

To investigate the Bayesian model of Section 4, two sets of estimates of the anatomical angles of the right ankle of the 65 experimental subjects were calculated. The first set used  $n = 50$  observations per subject and the Bayesian penalty using (6.1) as the parameters for the prior distribution. The second set was obtained by fitting the unpenalized model of Rivest, Baillargeon and Pierrynowski (2008) to the complete data sets of  $n = 1500$  frames per subject. For the two sets of estimates, the mean values for the five angles were similar to the estimate  $\hat{\beta}_0$  in Table 3. There were important differences in the between subject standard deviations: for the first set they were (5.15, 6.62, 6.61, 12.84, 8.25), respectively, for  $(t_1, t_2, s_1, s_2, \gamma_0)$ , while for the second it was (14.72, 18.70, 15.94, 36.56, 28.54). Thus, the  $n = 1500$  estimates are much more variable than their penalized alternatives. The added variability is caused by the numerical problems in maximizing the likelihood of the unpenalized model. Similar numerical problems were encountered by van den Bogert, Smith and Nigg (1994) when they fitted the ankle model using an ad hoc loss function. They proposed setting  $s_2 = 0$  to get stable estimates. The penalty of the Bayesian model allows to get meaningful values for the five angles.

Several studies have found that  $s_1 = stinc$  is the only angle that can be reliably estimated using either the model of Rivest, Baillargeon and Pierrynowski (2008) or the approach of van den Bogert, Smith and Nigg (1994). Figure 2 gives the boxplots of the two sets of estimates for  $s_1$ ; an outlier with a negative estimate for  $s_1$  has been removed. The penalty protects against values of  $s_1 = stinc$  larger than 70 degrees that are not possible anatomically.

7.2. *Data analysis for the location of lower extremity injuries.* Pierrynowski et al. (2003) collected ankle motion data from 31 participants who had experienced either knee ( $n_1 = 15$ ) or foot ( $n_2 = 16$ ) injuries during weight-bearing activities.

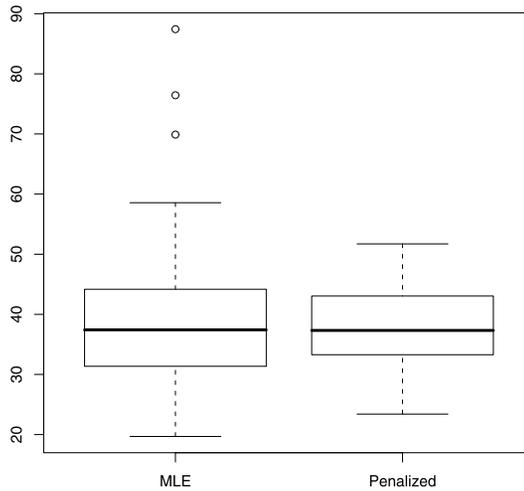


FIG. 2. *Boxplots for two sets of estimates for stinc expressed in degrees.*

The experiment’s goal was to determine whether there were significant differences in the orientation of the subtalar axis between the two groups. This section uses the data collected on the right ankle; the individual sample size is  $n_i = 50$  for each subject.

Table 4 presents the estimates obtained by fitting three models to this data; all the models have five independent random effects, one for each component of  $\beta$ . The first one has 7 fixed parameters; in addition to  $\beta$ , it features parameters  $ds_1$  and  $ds_2$  for the foot-knee differences of the two subtalar angles. Model 1 postulates

TABLE 4  
*The estimates obtained by fitting three models to the lower extremity injury data*

Model		$t_1$	$t_2$	$s_1$	$ds_1$	$s_2$	$ds_2$	$\gamma_0$
1	$\hat{\beta}_j$	-4.14	7.09	45.93	-7.41	6.69	-2.89	-7.33
1	s.e.	1.49	1.99	1.05	1.42	3.07	2.56	3.22
1	$\sqrt{\hat{\Sigma}_{0jj}}$	5.86	6.15	3.68	NA	4.68	NA	$<10^{-2}$
2	$\hat{\beta}_j$	-5.46	7.62	45.78	-7.45	3.44	NA	-8.89
2	s.e.	1.41	1.78	0.99	1.39	2.25	NA	2.54
2	$\sqrt{\hat{\Sigma}_{0jj}}$	5.69	5.40	3.58	NA	3.61	NA	$<10^{-2}$
3	$\hat{\beta}_j$	-4.71	5.94	43.01	NA	-0.04	NA	-9.75
3	s.e.	1.49	1.97	0.96	NA	2.69	NA	2.67
3	$\sqrt{\hat{\Sigma}_{0jj}}$	5.89	7.63	4.88	NA	6.35	NA	4.76

that the mean *stinc* and *stdev* values are  $(s_1, s_2)$  and  $(s_1 + ds_1, s_2 + ds_2)$  in the knee and the foot group respectively; the variances of these two angles are the same for both groups. Models 2 and 3 are derived from model 1 by setting  $ds_2 = 0$  and  $ds_1 = ds_2 = 0$ , respectively. Under model 3, the mean values of the five angles of the ankle model are the same in both groups. The LME and the PLME algorithms gave slightly different numerical results; the PLME estimates are presented as a larger maximum for the likelihood  $L_{1p}$  was obtained with this algorithm. The estimated residual standard error  $1/\sqrt{2\hat{\kappa}}$  was 0.024 radians, or 1.36 degrees for the three models.

In model 1, the  $z$ -statistic for testing the null hypothesis of no *stdev* effect is  $Z_{\text{obs}} = -2.89/2.56 = -1.13$  for a  $p$ -value of 0.26. This high  $p$ -value suggests setting  $ds_2 = 0$ ; this leads to model 2 where the test for a  $s_1$  between group difference has a  $p$ -value smaller than  $10^{-4}$ ; this is highly significant, even when accounting for the underestimation of the standard errors highlighted in Table 1. So model 2 provides the best fit, in agreement with the findings of Pierrynowski et al. (2003) who also noted the significant difference in *stinc* between the two groups. The nonlinear mixed effects model allows to test for an  $s_2 = \textit{stdev}$  difference; this was not possible with the unpenalized estimates since they were numerically unstable.

**8. A reduced model.** In his seminal work, Inman dealt with four angles, *ttinc*, *ttdev*, *stinc* and *stdev*. The fifth angle  $\gamma_0$  of (3.1) does not play any role in his investigations. This section suggests a reduced model featuring 4 anatomical angles instead of 5. It investigates whether this model leads to better estimates of the anatomical angles.

In the reference position the leg and the foot reference frames have the same orientation, thus, the predicted value  $\Psi = \mathbf{I}$  is possible. Therefore, for some angles  $\alpha$  and  $\phi$ , one has

$$\mathbf{I} = \mathbf{A}(t_1, t_2)\mathbf{R}(\alpha, x)\mathbf{R}(\gamma_0, z)\mathbf{R}(\phi, y)\mathbf{B}(s_1, s_2)^\top.$$

This equation implies that  $\gamma_0$  is the  $Z$ -Cardan angle of  $\mathbf{A}(t_1, t_2)^\top\mathbf{B}(s_1, s_2)$  in the  $X - Z - Y$  convention. Thus,  $\gamma_0$  can be expressed in terms of  $\mathbf{A}(t_1, t_2)$  and  $\mathbf{B}(s_1, s_2)$  as

$$(8.1) \quad \gamma_0(\mathbf{A}_1, \mathbf{B}_2) = -\arcsin(\mathbf{A}_1^\top\mathbf{B}_2).$$

Fitting a reduced model where  $\gamma_0$  depends of  $(s_1, s_2, t_1, t_2)$  through (8.1) is easily carried out in the framework of Section 3. All the previous derivations hold with the reduced model provided that the matrix  $\mathbf{X}$  is redefined in such a way that

its  $i$ th row is given by the  $4 \times 1$  vector:

$$\mathbf{X}_i = -\cos\{(\theta_i^z - \gamma_0)/2\} \\ \times \begin{pmatrix} -\left(\frac{\sin(t_1)\tan(t_2)}{D_i^2}\mathbf{A}_2 + \frac{1}{D_i}\mathbf{A}_3\right)^\top \left(\frac{\mathbf{R}_i}{\sqrt{1-(\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2}} - \frac{\mathbf{I}}{\cos \gamma_0}\right) \mathbf{B}_2 \\ \frac{\cos(t_1)\{1+\tan^2(t_2)\}}{D_i^2} \mathbf{A}_2^\top \left(\frac{\mathbf{R}_i}{\sqrt{1-(\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2}} - \frac{\mathbf{I}}{\cos \gamma_0}\right) \mathbf{B}_2 \\ \mathbf{A}_1^\top \left(\frac{\mathbf{R}_i}{\sqrt{1-(\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2}} - \frac{\mathbf{I}}{\cos \gamma_0}\right) \left(\frac{\sin(s_1)\tan(s_2)}{D_s^2} \mathbf{B}_1 + \frac{1}{D_s} \mathbf{B}_3\right) \\ -\frac{\cos(s_1)\{1+\tan^2(s_2)\}}{D_s^2} \mathbf{A}_1^\top \left(\frac{\mathbf{R}_i}{\sqrt{1-(\mathbf{A}_1^\top \mathbf{R}_i \mathbf{B}_2)^2}} - \frac{\mathbf{I}}{\cos \gamma_0}\right) \mathbf{B}_1 \end{pmatrix}.$$

The reduced model has been applied to the data on the 65 volunteers presented in Section 7.1. It yielded poor results; the within subject variability was larger than that with the five-parameter model of Section 4. The average values failed to reproduce Inman's results. To investigate this failure, we calculated the differences  $\hat{\gamma}_0 + \arcsin(\hat{\mathbf{A}}_1^\top \hat{\mathbf{B}}_2)$  obtained with the five-parameter model for the 65 subjects of Section 7.1. The average difference is  $-2.30$  degree (s.e. = 0.34); only 7 of the 65 differences have positive values.

The assumption underlying the reduced model, that the leg and the foot are aligned in the reference position, is not met. The reference position is measured when the experimental subject is standing, so that its ankle is loaded. However, the data is collected on an unloaded ankle moving freely when the subject is sitting. The nonnull value of  $\gamma_0 + \arcsin(\mathbf{A}_1^\top \mathbf{B}_2)$  might be explained by a slight change of the relative orientation of the two reference frames when the ankle goes from a loaded to an unloaded position. Thus, the reduced model is not suitable for the data analyzed in this work. It should, however, be considered when investigating the rotation axes of a loaded ankle. The value of  $\gamma_0 + \arcsin(\mathbf{A}_1^\top \mathbf{B}_2)$  may quantify rearfoot flexibility which is of interest to foot care professionals [Mansour et al. (2007)].

**9. Discussion.** This work presented a solution to the estimation of the directions of the two rotation axes of the ankle. The key element is the estimation criterion given by a penalized likelihood. This penalized likelihood is associated to a Bayesian model for the ankle joint and to a nonlinear mixed effects directional model that allows estimation of the between ankle variability of the rotation axes within a population. Simulations have shown that the population means and the population variances can be estimated in a reliable way. When used on a data set collected on a sample of volunteers, the nonlinear mixed effects directional model produced mean and variance estimates that were similar to those presented by Inman (1976). The good match with Inman's clinically accepted findings (see Table 3) provides empirical evidence that a two-axis (revolute) mechanistic model of the ankle [see equation (3.1)] is indeed appropriate for the ankle. Section 7 shows

that the proposed nonlinear mixed effects directional model can be extended to compare the ankle's axes in several populations. The Bayesian model of Section 4 might very well solve the problem of estimating in vivo the location of the ankle's rotation axes.

Future work of this model includes a detailed investigation of the within subject stability of the Bayesian estimates of Section 4, using both right and left foot data. The estimation of the translation parameters of the van den Bogert, Smith and Nigg (1994) ankle's model will also be investigated and the application of the reduced model to data collected on a loaded ankle will also be considered.

**Acknowledgments.** We are grateful to a referee for his perceptive comments who motivated Section 8 and to Sophie Baillargeon for carrying out the reduced model analysis.

## REFERENCES

- BINGHAM, C., CHANG, T. and RICHARDS, D. (1992). Approximating the matrix Fisher and Bingham distributions: Applications to spherical regression and procrustes analysis. *J. Multivariate Anal.* **41** 314–337. [MR1172902](#)
- CHIRIKJIAN, G. S. and KYATKIN, A. B. (2001). *Engineering Applications of Noncommutative Harmonic Analysis*. CRC Press, Boca Raton, FL. [MR1885369](#)
- GROOD, E. S. and SUNTAY, W. J. (1983). A joint coordinate system for the clinical description of three dimensional motion: Application to the knee. *J. Biomech. Eng.* **105** 136–144.
- INMAN, V. T. (1976). *The Joints of the Ankle*. Williams and Wilkins, Baltimore, MD.
- KIM, P. T. (1991). Decision theoretic analysis of spherical regression. *J. Multivariate Anal.* **38** 233–240. [MR1131717](#)
- LEÓN, C., MASSÉ, J.-C. and RIVEST, L.-P. (2006). A statistical model for random rotations. *J. Multivariate Anal.* **97** 412–430. [MR2234030](#)
- LEWIS, G. S., COHEN, T. L., SEISLER, A. R., KIRBY, K. A., SHEEHAN, F. T. and PIAZZA, S. J. (2009). In vivo test of an improved method for functional location of the subtalar joint axis. *J. Biomechanics* **42** 146–151.
- LEWIS, G. S., SOMMER, H. J. and PIAZZA, S. J. (2006). In vitro assessment of a motion-based optimization method for locating the talocrural and subtalar joint axes. *J. Biomech. Eng.* **128** 596–603.
- LINDSTROM, M. J. and BATES, D. M. (1990). Nonlinear mixed-effects models for repeated measures data. *Biometrics* **46** 673–687. [MR1085815](#)
- MANSOUR, E., BEGON, M., FARAHPOUR, N. and ALLARD, P. (2007). Forefoot–rearfoot coupling patterns and tibial internal rotation during stance phase of barefoot versus shod running. *Clin. Biomech.* **22** 74–80.
- MARDIA, K. V. and JUPP, P. E. (2000). *Directional Statistics*. Wiley, New York. [MR1828667](#)
- MCCARTHY, J. M. (1990). *Introduction to Theoretical Kinematics*. MIT Press, Cambridge, MA. [MR1084375](#)
- PIERRYNOWSKI, M. R., FINSTAD, E., KEMECSEY, M. and SIMPSON, J. (2003). Relationship between the subtalar joint inclination angle and the location of lower-extremity injuries. *J. Amer. Pediatr. Med. Assoc.* **93** 481–484.
- PINHEIRO, J. C. and BATES, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer, New York.

- RANCOURT, D., RIVEST, L. P. and ASSELIN, J. (2000). Using orientation statistics to investigate variations in human kinematics. *J. Roy. Statist. Soc. Ser. C* **49** 81–94. [MR1817876](#)
- RIVEST, L.-P., BAILLARGEON, S. and PIERRYNOWSKI, M. (2008). A directional model for the estimation of the rotation axes of the ankle joint. *J. Amer. Statist. Assoc.* **103** 1060–1069. [MR2462888](#)
- RIVEST, L.-P. and CHANG, T. (2006). Regression and correlation for  $3 \times 3$  rotation matrices. *Canad. J. Statist.* **34** 184–202. [MR2323992](#)
- VAN DEN BOGERT, A. J., SMITH, G. D. and NIGG, B. M. (1994). In vivo determination of the anatomical axes of the ankle joint complex: An optimization approach. *J. Biomechanics* **27** 1477–1488.

M. HADDOU  
L.-P. RIVEST  
DÉPARTEMENT DE MATHÉMATIQUES  
ET DE STATISTIQUE  
UNIVERSITÉ LAVAL  
QUÉBEC (QUÉBEC), G1V 0A6  
CANADA  
E-MAIL: [mohammed.haddou.1@ulaval.ca](mailto:mohammed.haddou.1@ulaval.ca)  
[Louis-Paul.Rivest@mat.ulaval.ca](mailto:Louis-Paul.Rivest@mat.ulaval.ca)

M. PIERRYNOWSKI  
SCHOOL OF REHABILITATION SCIENCE  
INSTITUTE OF APPLIED HEALTH SCIENCES  
MCMASTER UNIVERSITY  
1400 MAIN STREET WEST  
HAMILTON, ONTARIO, L8S 1C7  
CANADA  
E-MAIL: [pierryn@mcmaster.ca](mailto:pierryn@mcmaster.ca)