

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000

Individual predictions based on nonlinear mixed modeling: application to prenatal twin growth

J. J. Stirnemann^{a,b*}, A. Samson^a, and J. C. Thalabard^a

The assessment of growth during fetal life and childhood commonly relies upon cross-sectional reference ranges or centiles. However, individual sequential predictions may help the time-wise assessment of a growth process. In twin pregnancies for example, which are at-risk of growth restriction, such predictions may improve the detection of abnormal trajectories. In this article we present a simple forecasting method, assuming that a given normal individual behaves in the same way as a reference population. We consider as a prediction in a given individual, the forecast of a future observation conditional to any previous observation and a set of population parameters obtained by nonlinear mixed modeling in a reference population. We suggest an estimator for this prediction without resorting to linear approximation and show that it enjoys interesting asymptotics when the amount of observations increases over time. Two independent real datasets of twin pregnancies with normal growth and outcome are used to illustrate the application of such predictions in prenatal growth. The first dataset is considered as a reference dataset and modeled using a 2-level nonlinear model. Illustration and validation of predictions is performed on the second dataset. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: Prediction; Nonlinear mixed model; MCMC; Growth; Fetus; Twin pregnancy

1. Introduction

Growth is a continuous dynamic process, which is usually repeatedly measured and assessed over time. Human growth modeling during fetal life and childhood commonly relies upon the regression of cross-sectional data, defining normal population ranges or centiles for practical clinical use [1, 2, 3, 4, 5]. However, modeling of cross-sectional data may be unfit for individual growth prediction and sequential assessment since they do not describe the dynamic aspect of growth. Regardless of the method used to estimate these centiles, several problems arise when using such references for growth follow-up. Cross-sectional centiles only offer a cross-sectional coverage, i.e. in a given population, 90% of individuals should fall within the 5th and 95th centile. Consequently, the coverage of a growth path of an individual monitored longitudinally in time is unknown. Furthermore, population cross-sectional centiles ignore any previous measurements of a given individual and make separate assessments for each new time-point. For example, as part of routine follow-up, fetal biometry is monitored sequentially by ultrasound throughout pregnancy and especially in twins which are at-risk of

^a Applied Mathematics, MAP5, UMR CNRS 8145, Université Paris Descartes

^b Department of Obstetrics and Maternal-Fetal Medicine, GHU Necker - Enfants Malades, Université Paris Descartes

* Correspondence to: MAP5, Centre universitaire des Saints-Pères, 45 rue des Saints-Pères, 75006 Paris. E-mail: j.stirnemann@gmail.com

growth restriction. Therefore, monitoring growth requires a tool able to assess individual growth trajectories and not only data-points.

Mixed effects models [6, 7, 8] are a powerful tool for the analysis of longitudinal growth data taking into account the variability between subjects. Linear mixed models have been extensively used for modeling growth, including polynomials or splines (see [9] for example). Compared to previously reported linear models, nonlinear mixed models (NLMM) present desirable properties. Indeed, Pinheiro and Bates [8] discuss several benefits of nonlinear models, i.e. parsimony of model parameters, better control of overfitting with a mechanistic interpretation of model parameters. These properties are especially relevant when the purpose of modeling is forecasting since they provide validity beyond the observed range of the data better than linear models [8]. For example, several studies have used NLMMs for modeling growth in a reference set of individuals [10, 11, 12].

The objective of the present study is to predict growth characteristics given previous observations for a twin pregnancy in the setting of nonlinear mixed effects models. In twin pregnancies, because growths of both fetuses are intrinsically correlated, a two-level mixed model is more appropriate. However, to our knowledge, mixed models have never been applied to fetal growth in twins. We consider the problem of forecasting a future observation y_t at time t of an individual (a pair of twins) with a given set of measurements $y_{1:n} = (y_1, \dots, y_n)$ at time t_n , with $t > t_n$. We assume that the growth process of a new pregnancy behaves in the same way as the reference population modelled by a NLMM with a set of parameter estimates θ , allowing a transfer of information across individuals, as proposed by Bortot *et al.* [13]. In the case of linear mixed models, the conditional expectation of a future observation is explicit [14, 15, 16, 9]. In the context of one-level linear mixed models, Pan and Goldstein suggested conditional norms to make predictions for an individual based upon the available data at a given time-point [16, 9]. However, this prediction is not explicit in the nonlinear case. Although a nonlinear extension of the Kalman filter may be used to approximate a predictive distribution [17], the convergence of this method is not established.

In this article, we suggest to predict a future measurement based on the estimation of $p(y_t|y_{1:n}, \theta)$ by a Markov Chain Monte Carlo (MCMC) algorithm. Compared to approaches using the Kalman filter or its extensions, this estimate can be used for any non-linear function and does not require a linear or linearized model. As in other filtering schemes, the estimated expected value of the predictive distribution may be seen as the prediction itself; the estimated variance describes the precision of the prediction.

This approach is applied to real fetal growth in twin pregnancies monitored by ultrasound. The population parameters are estimated in a reference dataset of longitudinal biometric measurements. Another independent set of twin pregnancies with normal growth and outcome is used for validation. Specific graphical tools are built to present the sequential estimations of growth parameters across time.

Following a brief description of the data in Section 2, Section 3 presents the general formulation of 2-level NLMMs. Section 4 defines predictions in the setting of NLMM and describes the method of estimation as well as asymptotic properties. Section 5 presents the application of this forecasting scheme to prenatal twin growth real data. The conclusions are discussed in Section 6.

2. Prenatal twin growth data

Two independent datasets of follow-up ultrasound in twin pregnancies are presented in Figure 1. Intrauterine growth is monitored sequentially by ultrasound using several biometric measurements, including the head perimeter, biparietal head diameter, abdominal circumference and femoral length. Throughout the article we will consider a univariate model based upon abdominal circumference, although extension to the multivariate case is straightforward as discussed in Section 4.3.

Dataset A comprises 277 twin pregnancies with normal growth and outcome. For both twins and each pregnancy, a normal birth weight was observed. Measurements were made available by a first ultrasound private practice in Paris,

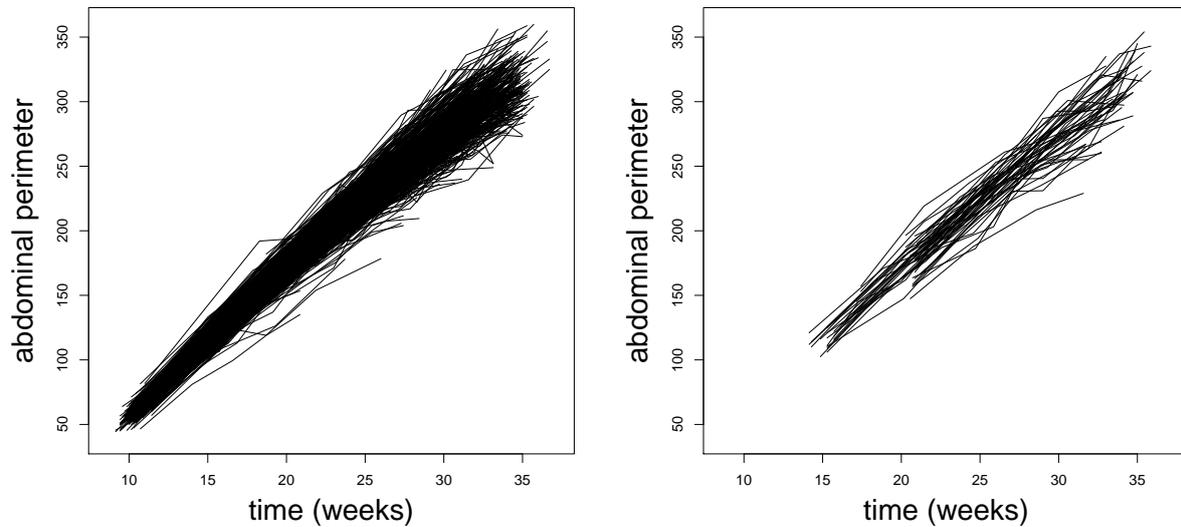


Figure 1. Individual growth of twins from two independent datasets: normal growth from dataset A (left), normal growth from dataset B (right)

France. Dataset B is a dataset of 27 twin pregnancies with normal growth and outcome, obtained in the same way as dataset A but from a different practice in Paris. In datasets A and B, observation times may be different from one pregnancy to another, but as part of routine follow-up, abdominal circumference is measured as early as 12 or 16 weeks (about 3 or 4 months) of gestation and at least monthly onwards until delivery. Therefore in a normal follow-up, twins are measured at 12, 16, 20, 24, 28, 32 and 36 weeks. By that time most twin pregnancies will deliver within the following two weeks. In both datasets A and B, drop-outs during follow-up are possible and the number of visits varied from 4 to 13 per pregnancy during gestation. The median (interquartile range) of time interval between visits was 4 (4-4.4) weeks and 4 (3.4-4.1) weeks in dataset A and B respectively.

3. Nonlinear mixed models for the reference population

The reference population of twin pregnancies (dataset A) is analyzed with a 2-level nonlinear mixed model (NLMM). For the sake of clarity, we restrain the presentation of the methodology to a scalar measurement (abdominal circumference in mm, denoted AC). The extension to a multivariate measurement (namely the head perimeter, biparietal head diameter, abdominal perimeter and femoral length) is presented in Section 4.3.

Let y_{ikj} be the observation for subject k ($k = 1, 2$), in unit i ($i = 1, \dots, N$) at time t_{ikj} ($j = 1, \dots, n_{ik}$). In the case of twin pregnancies, a unit represents a pregnancy and two subjects/twins. The 2-level NLMM can be written as:

$$y_{ikj} = f(\phi_{ik}, t_{ikj}) + g(\phi_{ik}, t_{ikj})\varepsilon_{ikj}, \quad \varepsilon_{ikj} \sim_{iid} \mathcal{N}(0, \sigma^2) \quad (1)$$

where ϕ_{ik} is the p -vector of growth parameters of subject k in pregnancy i and ε_{ikj} is the residual error. The function f is the structural nonlinear growth function and g is the variance function. Furthermore, we assume an additive model for the random effects:

$$\phi_{ik} = \mu + b_i + c_{ik}, \quad b_i \sim_{iid} \mathcal{N}(0, \Sigma), \quad c_{ik} \sim_{iid} \mathcal{N}(0, \Psi) \quad (2)$$

with μ the p -vector of means (fixed effects), b_i the random effect of size p of pregnancy i and c_{ik} the random effect of size p of twin k in pregnancy i . The variance matrices $\Sigma = (\Sigma_{\ell h})_{\ell=1, \dots, p; h=1, \dots, p}$ and $\Psi = (\Psi_{\ell h})_{\ell=1, \dots, p; h=1, \dots, p}$ are the inter- and

intra-pregnancy covariance matrices. Therefore, denoting \mathbf{v}' the transpose of vector \mathbf{v} , $\phi_i = (\phi'_{i1}, \phi'_{i2})' \sim \mathcal{N}((\mu', \mu')', \Omega)$ with Ω the $(2p \times 2p)$ matrix defined as:

$$\Omega = \begin{pmatrix} \Sigma + \Psi & \Sigma \\ \Sigma & \Sigma + \Psi \end{pmatrix}$$

The vector of unknown parameters is $\theta = (\mu, \Sigma, \Psi, \sigma^2)$.

In the following, we denote $p(\phi_{ik}; \theta)$ the Gaussian density of the individual parameters ϕ_{ik} and $p(y_{ikj} | \phi_{ik}; \theta)$ the Gaussian density $\mathcal{N}(f(\phi_{ik}, t_{ikj}), \sigma^2)$ of the observation y_{ikj} conditional on ϕ_{ik} . Note that $p(y_{ikj} | \phi_{ik}; \theta)$ depends on θ only through its component σ^2 .

Several algorithms [18] exist for the estimation of θ , including Gaussian quadrature [19], first-order conditional estimates [8] and stochastic approximation of the EM algorithm [20] (see [21] for a comparison of the algorithms). In the following, we consider that a NLMM is built on the reference population (dataset A). The set of reference individuals is denoted $I_A = \{1, \dots, N\}$.

4. Predictions for a new individual

Since the objective is to predict a future growth characteristics for a new pregnancy (i.e. that is not in I_A), we must explain what exactly we mean by prediction. Consistent with the general terminology of mixed effects models [8], wherein prediction refers to inference about random effect parameters as well as observations, under the rubric of prediction we include answers to two relevant questions for a given pregnancy: 1) what are the growth parameters for the two twins of that particular pregnancy; and 2) what will be the observed measurements at some future time. The answers will be based upon the model and model parameters θ estimated on the reference sample I_A and any data-points available for the specific individual we are considering. We consider that the estimation of θ is achieved from some reference dataset (dataset A in our case, see Section 3), and we consider θ to be known.

We assume that, for each new pregnancy, the corresponding twins have the same time-wise behaviour as the reference population (dataset A). We assume that n measurements are available for this pregnancy at time t_n . Let y_{kj} denote the measurement at time t_{kj} for $j = 1 \dots n$ of the k^{th} subject in this pregnancy. Denote $y_{k1:n} = (y_{k1}, \dots, y_{kn})$ the vector of observations for the k^{th} twin, $k = 1, 2$. Denote $y_{1:n} = (y'_{11:n}, y'_{21:n})'$ the available data vector for the whole pregnancy up to time t_n . Similarly, let ϕ_k denote the vector of parameters of the k^{th} new twin, $k = 1, 2$. We denote $\phi = (\phi'_1, \phi'_2)'$ the vector of the two twins.

Considering the general definition of prediction discussed above, our goals are to make inferences about 1) $\phi = (\phi'_1, \phi'_2)'$ and 2) a future vector of observations $y_t = (y_{1t}, y_{2t})$ for both twins at time $t > t_n$ conditionally on the available measurements $y_{1:n}$, under model parameters θ .

Our approach will be based on Bayesian predictive inference. For ϕ , inference will be based on the posterior distribution given $y_{1:n}$. Taking the model parameters θ as known, $p(\phi, \theta)$ is the relevant prior distribution for ϕ . Using Bayes' theorem the posterior distribution given $y_{1:n}$ is

$$p(\phi | y_{1:n}; \theta) = \frac{p(y_{1:n} | \phi; \theta) p(\phi; \theta)}{p(y_{1:n}; \theta)}, \quad (3)$$

where $p(y_{1:n}; \theta) = \int p(y_{1:n}, \phi; \theta) d\phi$ is the likelihood of the pregnancy data, and $p(y_{1:n} | \phi; \theta) = \prod_{j=1}^n p(y_j | \phi; \theta)$ by independence of the measurements conditional on ϕ . For a future observation y_t , inference will then be based on the predictive distribution

$$p(y_t | y_{1:n}; \theta) = \int p(y_t | \phi; \theta) p(\phi | y_{1:n}; \theta) d\phi \quad (4)$$

Two kinds of summaries of these distributions are useful. The first are the expected values

$$\mathbb{E}(\phi|y_{1:n}; \theta) \tag{5}$$

$$\mathbb{E}(y_t|y_{1:n}; \theta) \tag{6}$$

Expectation (5) may approximate the unknown pregnancy-specific parameter ϕ and (6) may approximate the measurement that will be obtained at the future time $t > t_n$. The appendix, to be discussed later, lends some support to these possibilities. The second summaries are called prediction intervals. A prediction interval is a range of values of ϕ or y_t with a specified probability coverage in (3) or (4), respectively. Such an interval provides a simple quantification of the precision of our current state of knowledge.

Of notice, since no hypothesis have been made regarding the prediction time t , a prediction may be done at any point on the time-line. However, regarding intrauterine growth, predictions obviously can not be extended past delivery. In a clinical perspective we may use predictions in several ways for growth monitoring: (i) One may be interested in one-step ahead predictions by comparing the observation made on a follow-up visit to a prediction or prediction interval estimated using the previous measurements. Considering growth monitoring as a screening tool for growth abnormalities, such prediction intervals should improve both the false negative rate by allowing an early detection of abnormal growth and false positive rates. Suspected impaired growth may warrant closer follow-up and second-line monitoring tools in order to avoid complications of intrauterine hypoxemia by anticipating delivery. On the contrary, such medical action may impose unnecessary ultrasound, maternal anxiety or even prematurity when a fetus is wrongfully classified as small. (ii) It might also be of interest to predict a biometric measurement such as weight at $t=9$ months of pregnancy, using a set of observations midway through pregnancy. Such a prediction is to be compared with the distribution of birthweight in normal newborns.

4.1. Calculation of predictions

In mixed-effects terminology, estimates of (5) and (6) are called predictions. In nonlinear models, such predictions are complex, as are the determinations of prediction intervals. When f is a linear function with regard to ϕ , the likelihood $p(y_{1:n}; \theta)$ has a closed-form solution as well as $p(y_t|y_{1:n}, \theta)$ (see [14, 15] for application to growth) and the best linear unbiased predictor (BLUP) is recovered (see [8]). Conversely, when f is nonlinear with respect to ϕ , $p(y_{1:n}; \theta)$ does not have a closed-form expression and therefore neither do $p(\phi|y_{1:n}; \theta)$ nor $p(y_t|y_{1:n}; \theta)$. However, previous studies have suggested estimators of such predictions in nonlinear mixed models. These approaches use first-order approximation of the model and comprise the so-called plug-in estimator $f(\hat{\phi}, t)$ where $\hat{\phi}$ is an estimator of $\mathbb{E}(\phi|y_{1:n}; \theta)$ (see [8]) and a nonlinear analog of the best linear unbiased predictor (BLUP) used in linear mixed models (see [22]). However, the simple plug-in estimator that relates $\hat{\phi}$ and $f(\hat{\phi}, t)$ will yield a biased prediction of $\mathbb{E}(y_t|y_{1:n}; \theta)$ since f is nonlinear in our case. In the following, we suggest two unbiased estimators $\hat{\phi}$ and \hat{y}_t of $\mathbb{E}(\phi|y_{1:n}, \theta)$ and $\mathbb{E}(y_t|y_{1:n}, \theta)$ respectively without resorting to approximation.

We suggest to resort to Monte Carlo integration to compute the distributions $p(\phi|y_{1:n}; \theta)$ and $p(y_t|y_{1:n}; \theta)$ described in (4) and (3). This can be achieved using a multidimensional Markov Chain Monte Carlo (MCMC) algorithm to simulate an M -sample $(\phi^m)_{m=1, \dots, M}$ of ϕ with a target distribution $p(\phi|y_{1:n}; \theta)$. Therefore, for any function F , we propose the following estimate of $\int F(\phi)p(\phi|y_{1:n}; \theta)d\phi$

$$\hat{F}_n^M(\phi) = \frac{1}{M} \sum_{m=1}^M F(\phi^m) \tag{7}$$

For a fixed number n of available measurements, the estimate $\hat{F}_n^M(\phi)$ is convergent when M increases, due to the ergodic theorem [23, 24]. Following [24], we can also deduce a central limit theorem and hence a confidence interval of this

estimator.

In the following, we use the estimator (7) for different functions F . When F is the identity function,

$$\hat{\phi}_n = \frac{1}{M} \sum_{m=1}^M \phi^m \quad (8)$$

is an estimate of the individual conditional expectation $\mathbb{E}(\phi|y_{1:n})$. When $F(u) = u^2 - \mathbb{E}^2(u)$, we obtain an estimate

$$\widehat{\text{var}}_n(\phi) = \frac{1}{M} \sum_{m=1}^M ((\phi^m)' \phi^m - \hat{\phi}'_n \hat{\phi}_n) \quad (9)$$

of $\text{var}(\phi|y_{1:n})$. Similarly we may choose $F(u) = f(u, t)$ which yields:

$$\hat{y}_{tn} = \frac{1}{M} \sum_{m=1}^M f(\phi^m, t) \quad (10)$$

an estimator of $\mathbb{E}(y_t|y_{1:n}, \theta)$. Considering $F(u) = f(u, t)^2 - \mathbb{E}^2(f(u, t))$, we get the following estimation of the variance of prediction $\text{var}(y_t|y_{1:n}, \theta)$:

$$\widehat{\text{var}}_n(y_t) = \frac{1}{M} \sum_{m=1}^M (f(\phi^m, t)^2 - \hat{y}_{tn}^2). \quad (11)$$

More generally, F may be any appropriate function that yields convergent estimators of the moments, quantiles, etc.

Regarding the 2-level structure of our model (1), it is worth recalling here that $\hat{\phi}_n$ and \hat{y}_{tn} are estimators for the whole pregnancy, meaning that predictions are jointly estimated for both twins. Predictions for the k^{th} twin are denoted $\hat{\phi}_{kn}$ and \hat{y}_{tkn} .

In the following, we chose to compute $(1 - \tau)$ -level prediction intervals (PI) of $(\phi|y_{1:n}; \theta)$ and $(y_t|y_{1:n}; \theta)$ by using symmetric quantiles of $F(\phi^m)$ ($F = Id$ and $F = f(u, t)$). Denoting $q(\cdot, u)$ the quantile function, we define PI as:

$$PI = [q(F(\phi^m), \tau/2); q(F(\phi^m), 1 - \tau/2)] \quad (12)$$

The choice of the critical value τ is debatable and its influence on detection of growth abnormalities is out of the scope of this paper. However, heuristically, we may consider 80% or 90% prediction intervals by analogy with existing cut-offs regarding normal growth.

The previous estimate has been implemented in language R, using a p -dimensional Metropolis-Hastings algorithm.

4.2. Asymptotic behavior and comparison with population reference ranges

Theoretical properties and proofs regarding the convergence of the distributions (4) and (3) when n grows to infinity are provided in Appendix A. It is shown that the expected value (5) and the variance of $p(\phi|y_{1:n})$ converge when the number of available data increases. It is not shown that the expected value converges to the true value $\mathbb{E}(\phi)$, but to an expectation of ϕ conditional on an infinite amount of observations. Simulations in Appendix B suggest however that this conditional expectation and the true unknown ϕ are the same. Therefore, the unbiased estimate $\hat{\phi}_n$ of (5) is an increasingly better predictor of the true ϕ as more data accumulates. We also prove that for any given time t , $\mathbb{E}(y_t|y_{1:n})$ converges with accruing data. Similar to the case with ϕ , \hat{y}_{tn} is an increasingly better predictor of the true unknown y_t as more data accumulates. In the homoscedastic case ($g(\phi, t) = 1$), the variance is bounded above σ^2 . This implies that the prediction variance is expectedly greater than the residual variance. Therefore in problems with a high residual noise, the prediction interval of a future measurement is large. These properties are illustrated by simulation in Appendix B.

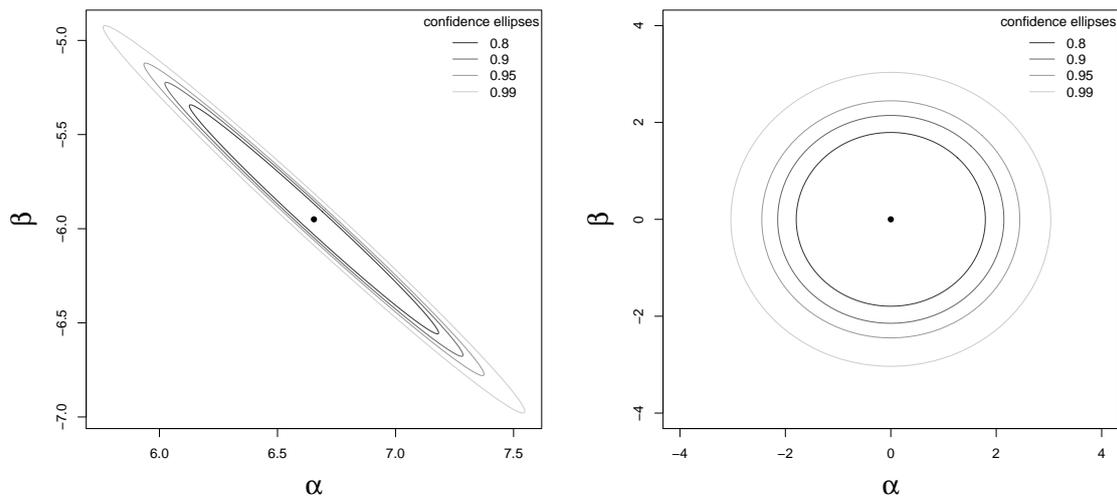


Figure 2. Reference ranges for ϕ defined by the confidence ellipses representing the population parameters θ in the plane defined by the $p = 2$ components of vector ϕ . The single point represents the population mean μ . Left-hand side: without centering and scaling; Right-hand side: after centering and scaling. The 80%, 90%, 95% and 99% confidence ellipses are given around the mean.

4.3. Extension to the multivariate case

Prenatal growth assessment by ultrasound usually relies upon four quantities (head perimeter, biparietal head diameter, abdominal perimeter and femoral length). These measurements may be seen as a multivariate vector of responses. Therefore, each observation y_{ikj} is an m -vector of measurements. The multivariate formulation of model (1) is thus based on μ an mp -vector of fixed effects; Ω an $2mp \times 2mp$ variance matrix; σ^2 the vector of residual variance. These parameters may be used directly to compute the estimate $\hat{F}_n^M(\phi)$. Therefore the multivariate case does not require specific theoretic modification although it may significantly increase the dimension of the problem and impact the MCMC estimation.

4.4. Graphical tools

We suggest to build a graphical tool to compare individual predictions to population reference values. The most intuitive graphical depiction is to compare individual predictions \hat{y}_{tn} to an overall population prediction interval, as routinely used in cross-sectional reference charts. For longitudinal data modeled by NLMM, this population interval may be computed numerically from the parameter estimates of the nonlinear mixed model defined in Section 3 and estimated in the reference population.

Another approach is to present the individual estimated parameters $\hat{\phi}_{kn}$. For example, considering a model requiring a vector ϕ of length $p = 2$ denoted $\phi = (\alpha, \beta)$, this may be done in two dimensions defined by each component of the vector ϕ . In this 2-dimensional plane, the reference ranges given by the population parameters (μ, Ω) are presented as Gaussian ellipses of given percentiles. Therefore, individual predictions $\hat{\phi}_{kn}$ are single points. Using this graphical display, an individual growth path is assessed by the departure of the prediction from the center. If the growth process is monitored iteratively, on clinical visits for example, the timewise evolution of each new prediction $\hat{\phi}_{kn}$ provides insight in the growth dynamics of the individual that is monitored. An example of growth monitoring using this graphical display is given in Section 5.3. To ensure clarity, the plane defined by the vector ϕ is centered and scaled according to the population distribution $\mathcal{N}(\mu, \Omega)$. Compared to traditional representations of growth data, this approach has the benefit of presenting filtered estimates of meaningful growth parameters while allowing a clear visual assessment of departure from reference value. Figure 2 presents the ellipses of population parameters before and after centering/scaling.

Table 1. Estimates of the 2-level NLMM parameter θ for the abdominal perimeter in real twin growth (dataset A)

Parameter	estimate (SE)	Parameter	estimate (SE)
μ_α	6.654 (0.011)	μ_β	-5.950 (0.013)
Σ_{11}	0.047 (0.017)	Ψ_{11}	0.039 (0.015)
Σ_{22}	0.064 (0.023)	Ψ_{22}	0.051 (0.019)
Σ_{12}	-0.054 (0.020)	Ψ_{12}	-0.044 (0.017)
σ^2	0.002 (3.10 ⁻⁵)		

5. Application to fetal growth in twins

We consider the two datasets of prenatal twin growth described in Section 2. Dataset A, which includes pregnancies with normal growth, is considered as the reference population. Predictions are estimated for pregnancies from dataset B.

5.1. Reference population analysis

A 2-level NLMM is used to estimate the population reference from dataset A. A Weibull growth function [10] is used as the structural growth function f :

$$f(\psi, t) = \alpha(1 - e^{-\beta t}) \quad (13)$$

with $\psi = (\alpha, \beta)$, $\alpha > 0$, $\beta > 0$ and $\phi = (\log \alpha, \log \beta)$ to ensure the positiveness of the parameters. The Weibull function (14) allows a mechanistic interpretation of growth considering α as the asymptote of growth and β as growth velocity. Estimation is performed with the SAEM algorithm for a 2-level NLMM [25]. A heteroscedastic variance model with $g(\phi, t) = f(\phi, t)$ is found as best fit (BIC = 25825 for the homoscedastic model, BIC = 25540 for the heteroscedastic model). Parameter estimates are presented in Table 1. The goodness of fit plots were satisfactory. External validation of both the model and corresponding population parameter estimates was performed on dataset B, using the normalized distribution prediction errors (*npde*) method described by Brendel *et al.* [26] with satisfactory results (data not shown). Noteworthy in Table 1, the magnitudes of intra-pregnancy and inter-pregnancy variances are within same ranges, and therefore should not be neglected.

5.2. Consistency of predictions in normally grown twins

Dataset B is a set of twin pregnancies with normal growth and outcome. For each pregnancy i and each twin $k = 1, 2$ in dataset B, we considered the fourth measurement of abdominal perimeter as unknown. Therefore, for each pregnancy, 3 observations were given and the fourth was held out. These four measurements are one month apart following routine ultrasound monitoring of twin pregnancies in France. In dataset B, the third and fourth measurements were made at 26.5 weeks (IQR=24.5-30) and 29.5 weeks (IQR=28.2-33) respectively. We wish to predict the fourth measurement, using a prediction interval and compare this interval to the true observation. Therefore, a 50% and a 90% prediction interval of $\mathbb{E}(y_{ik4}|y_{i11:3}, y_{i21:3})$ was computed for each of the $2N = 54$ fetuses in dataset B. We recall that predictions for both twins are estimated jointly due to the 2-level nature of model (1) accounting for the natural correlation between the twins. Therefore the prediction of y_{i14} for the first twin will necessarily be conditional to both $y_{i11:3}$ and $y_{i21:3}$. The prediction intervals showed satisfying coverage since the proportion of observations y_{ik4} belonging in their respective prediction interval were 94% and 53% for the 90% and the 50% prediction intervals respectively. These results emphasize that individual prediction intervals will yield an appropriate coverage for the individual growth path compared to population cross-sectional centiles which only yield point-wise coverage. In Figure 3, the computed 90% prediction intervals of $\mathbb{E}(y_{ik4}|y_{i11:3}, y_{i21:3})$ are plotted in nine twins from dataset B and compared to the value of y_{ik4} considered unknown in the estimation. For comparison, the routinely used 5th and 95th population percentiles computed from dataset A are also plotted. The increase in precision with 3 given observations using estimated individual prediction intervals is visible in all

cases compared to standard population percentiles. This illustrates with real data the theoretical considerations regarding asymptotical behavior of our predictions discussed in Section 4.2 and Appendix A.

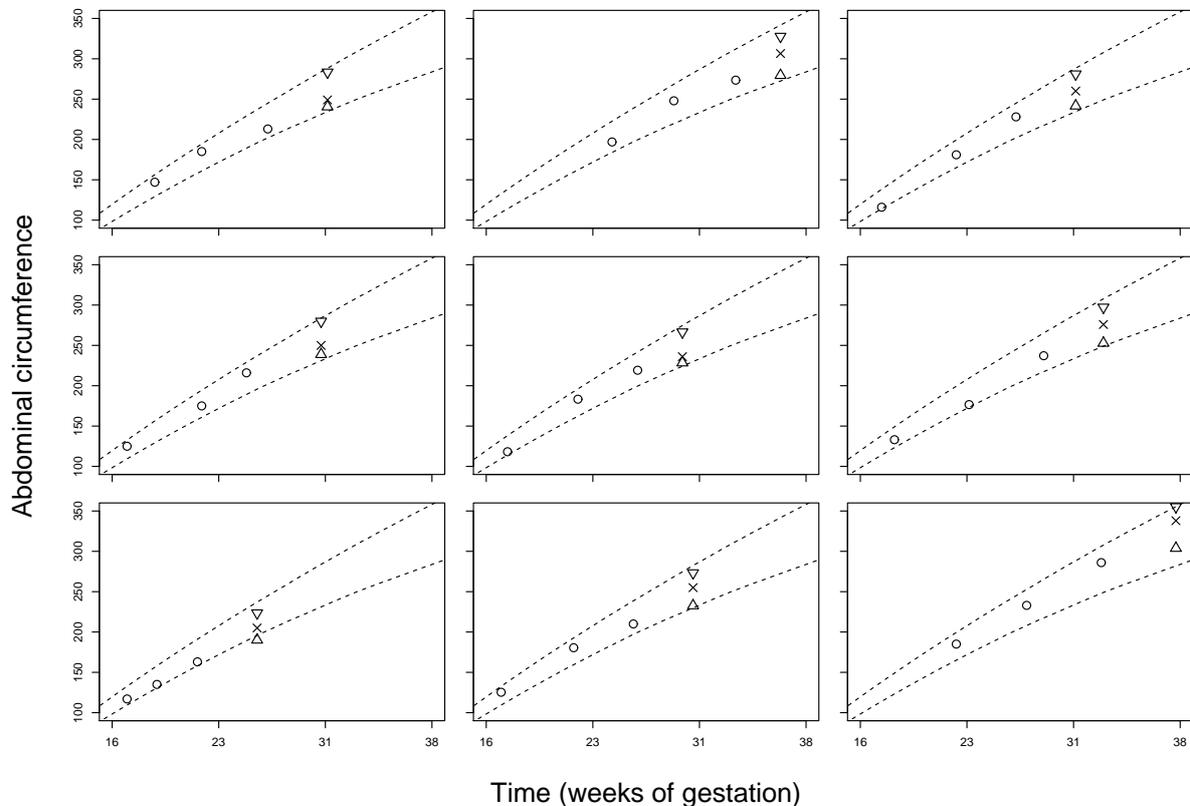


Figure 3. Observed data (circles) and hold-out data ("x") in 9 random fetuses from twin pregnancies from dataset B. The 90% prediction intervals of $\mathbb{E}(y_{ik4}|y_{i1:3}, y_{i21:3})$ are given for each (triangles). The 5th and 95th pointwise population centiles computed from dataset A are provided for comparison (dashed lines).

5.3. Illustration of predictions in abnormal growth

The illustration of individual predictions in the case of abnormal growth is presented in a study of a pregnancy with discordant growth that required preterm delivery, one twin showing normal growth and normal birth weight and the other showing severe intrauterine growth restriction. This pregnancy is compared to a normal pregnancy from dataset B. We consider one-step-ahead predictions in the setting of routine follow-up, using $\hat{\phi}_{kj}$ the estimates of $\mathbb{E}(\phi_k|y_{1:j}; \theta)$ for $j = 1, \dots, n$ and the graphical display described in Section 4.4. Therefore, in a clinical frame, for a single pregnancy followed throughout pregnancy the indice j is incremented for each visit yielding a new estimate $\hat{\phi}_{kj}$ for each new observation as it becomes available. As discussed in Section 4.2, predictions will become more precise as the available data increases. The normal case from dataset B had $n = 4$ follow-up visits at $t_{1:4} = (21, 26, 32, 34.7)$ weeks and the abnormal case had $n = 9$ visits at $t_{1:9} = (10.3, 15.6, 17, 19.4, 21.4, 23.3, 25.6, 28.6, 30.3)$ weeks. For both pregnancies, the 5th and 95th population pointwise percentile curves computed in dataset A are plotted with the observed twin measurements in Figure 4 (left-side panels), representing the usual tools used for growth follow-up. The concordant (top panels) and discordant growth (bottom panels) cases may be easily identified. The right-hand side of Figure 4 shows the estimates $\hat{\phi}_{kj}$ sequentially updated at each follow-up visit considering the remaining data $y_{j+1:n}$ unknown. These estimates are plotted in the standardized plane defined by the $p = 2$ components of ϕ , using the ellipses defined by the population parameters θ for reference. In the normal pregnancy, the values stay within normal ranges whereas early departure of one twin is found in the discordant case. Early departure from reference ranges is more visible using prediction estimates in the ϕ -plane (right-side figures) than using population reference ranges (left-side figures) since the noise of measurement error is

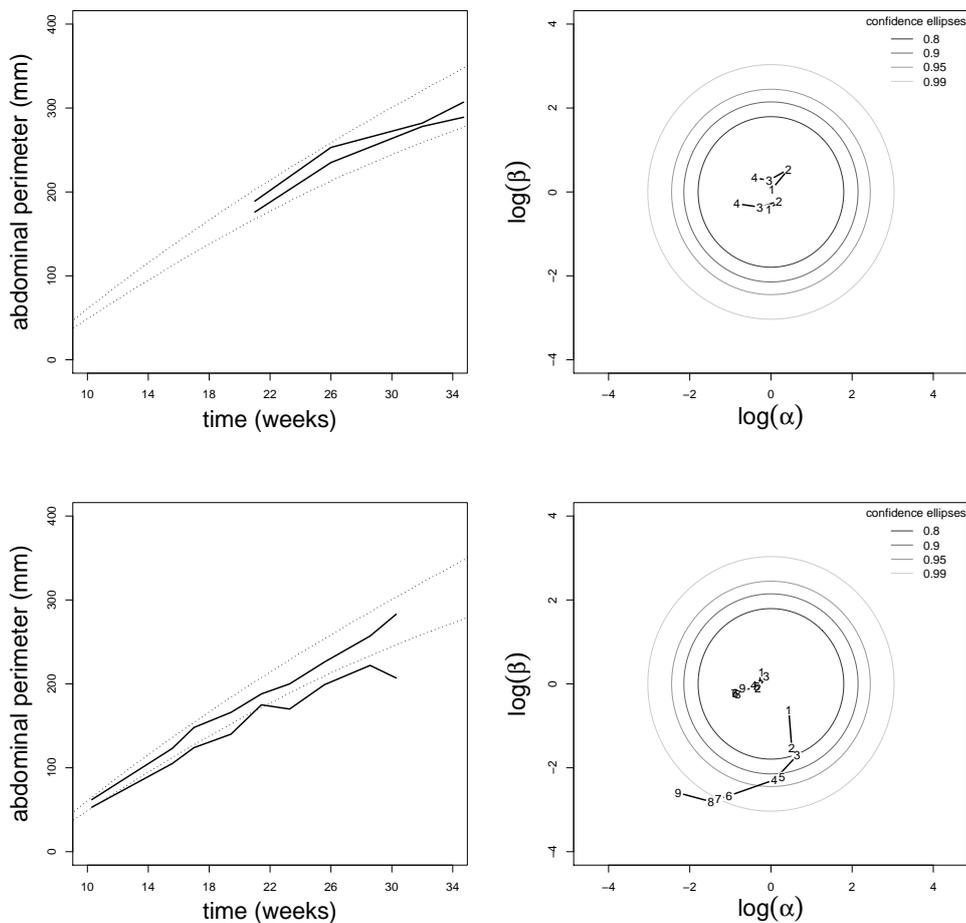


Figure 4. Growth trajectories in one normal pregnancy from dataset B (top panels) and an abnormal pregnancy (bottom panels). On the left-hand side, for each pregnancy, the raw observations of abdominal perimeter are plotted together with the 5th and 95th percentile of standard growth curves estimated on dataset A as in routine clinical practice. On the right-hand side, the individual estimate $\hat{\phi}_{k,j}$ is plotted across time ($j = 1 : n_i$) in the normalized plane defined by the $p = 2$ components of ϕ . In the normal pregnancy, $n = 4$ and in the abnormal case, $n = 9$. The indice of the visit is given on each point: in the normal pregnancy, $t_{1:4} = (21, 26, 32, 34.7)$ weeks and in the discordant pregnancy $t_{1:9} = (10.3, 15.6, 17, 19.4, 21.4, 23.3, 25.6, 28.6, 30.3)$. The 80%, 90%, 95% confidence intervals of the population parameters estimated on the reference population (dataset A) are given for reference.

filtered. Interestingly, in a mechanistic interpretation of growth dynamics, both components of ϕ decrease over time in the growth restricted fetus, indicating that both his asymptote and its velocity were abnormally low.

6. Conclusion

We have considered the problem of forecasting a future value of a growth process given previous measurements of an individual. We first assume that a NLMM has been built to analyze growth data from a reference population (dataset A). This reference population is used to choose the best regression function and to estimate the parameters (mean parameters, inter-patient variability, measurement noise). This step is performed using an estimation algorithm for NLMMs, such as the SAEM algorithm [20], the Gaussian quadrature [19, 27] or FOCE [28, 7]. In this paper, given the regression function and the estimated parameters, we propose an estimate of an individual's predictive distribution based on an MCMC algorithm.

The forecasting method advocated in this article overcomes any dilemma regarding the use of cross-sectional normal ranges for the assessment of an individual's longitudinal path by sequentially providing an individualized unbiased predictions of expected measures with consistent coverage. Moreover in the case of twin pregnancies, it provides

reference ranges that encompass the natural correlation between both fetuses. Since twins are at-risk of intrauterine growth restriction, specific models that allow reliable forecasting are needed.

The demonstration of convergence of predictions is a robust result since it implies that the prediction will improve with accruing data. In the setting of intrauterine growth monitoring, this means that individual prediction point-estimates and intervals will improve as measurements are made throughout pregnancy. Therefore, our method is clinically more relevant than cross-sectional growth curves. Indeed, convergence of the predictive distribution implies that estimated individual prediction intervals should systematically improve predictions compared to routinely used population reference ranges. This may have two strong clinical implications: i) improve the early recognition of growth anomalies and ii) prevent considering as *abnormal* growth patterns that would have been considered abnormal using population reference intervals. The most difficult step of our forecasting scheme relies in the definition of the *reference* population as it determines the set of parameters θ used for individual predictive distribution estimation. One must resort using surrogates to define this population. For example, in growth, one might choose a reference population according to several criteria: eventless pregnancy, delivery at term, normal delivery, normal baby development, ... Each of these criteria is debatable and one might want to add more stringent conditions. However, any modification in these criteria will change the estimated value of θ and therefore the interpretation of the forecasted distribution for new individuals.

We are aware of several obstacles that could limit a wider use of this forecasting methodology. The time required to obtain an accurate MCMC estimation of individual predictions may be a drawback. However, in our example, a satisfactory empirical estimation of $p(y_t|y_{1:n}; \theta)$ is achieved in less than one minute on a basic computer. This should be improved by implementations in compiled languages such as C for example. However, MCMC estimation may be unfit for rapidly accruing data. A sequential algorithm, such as filtering, could be more adapted to the problem. New particle filter algorithms may be seen as a promising alternative to MCMC. Finally, the main limit of our estimate is the preliminary estimation of parameters θ . The estimated parameter $\hat{\theta}$ that is plugged in the MCMC algorithm has an uncertainty that is not taken into account. This uncertainty is usually estimated through the standard error of $\hat{\theta}$. Our approach can be easily extended to take into account this uncertainty with a Bayesian approach. We could consider the parameter θ as random, with expectation $\hat{\theta}$ and standard deviation equal to the standard error of $\hat{\theta}$.

Nonetheless, individual forecasting is a promising tool. In fetal growth, it provides a truly individualized reference distribution that may be updated at each new ultrasound examination. Moreover, covariates may easily be added such as parental biometry or birth weights of previous babies, thus increasing the predictive value. The decisional aspect of sequential monitoring deserves to be investigated further. Although this is beyond the scope of this paper, when using tests in sequential monitoring, one must be aware of the inflation of the first-order error. In growth monitoring, for example, the purpose of sequential testing may be the early detection of growth restriction. Indeed, although previous studies have suggested specific models for the time-wise assessment of growth restriction [29], our method provides straightforward predictions of growth measurements at any given time-point. This approach may therefore help the early recognition and decision-process in intrauterine growth disorders by providing an intuitive tool for daily clinical use for any number of given measurements or time intervals between measurements.

Acknowledgements

We thank Dr. Bernard M.D., Dr. Levailant M.D. and Mrs. Heurteloup who provided the data for this study. The authors thank one of the referee for his (her) insightful comments which contributed to drastically improve the submitted manuscript.

References

1. Altman DG. Construction of age-related reference centiles using absolute residuals. *Statistics in Medicine* 1993; **12**(10):917–24.
2. Borghi E, de Onis M, Garza C, den Broeck JV, Frongillo EA, Grummer-Strawn L, Buuren SV, Pan H, Molinari L, Martorell R, *et al.*. Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. *Statistics in Medicine* 2006; **25**(2):247–65.
3. Cole TJ. Growth charts for both cross-sectional and longitudinal data. *Statistics in Medicine* 1994; **13**(23-24):2477–92.
4. de Onis M, Onyango AW, Borghi E, Garza C, Yang H. Comparison of the World Health Organization (WHO) child growth standards and the national center for health statistics/WHO international growth reference: implications for child health programmes. *Public Health Nutrition* 2006; **9**(7):942–7.
5. Wright EM, Royston P. A comparison of statistical methods for age-related reference intervals. *Journal of the Royal Statistical Society, Series A* 1997; **160**:47–69.
6. Davidian M, Giltinan D. *Nonlinear models to repeated measurement data*. Chapman and Hall: London, 1995.
7. Lindstrom MJ, Bates DM. Nonlinear mixed-effects models for repeated measures data. *Biometrics* 1990; **46**:673–687.
8. Pinheiro J, Bates D. *Mixed-effect models in S and Splus*. Springer-Verlag, 2000.
9. Pan H, Goldstein H. Multi-level repeated measures growth modelling using extended spline functions. *Statistics in Medicine* Dec 1998; **17**(23):2755–2770.
10. Jaffrézic F, Meza C, Lavielle M, Foulley JL. Genetic analysis of growth curves using the SAEM algorithm. *Genetics Selection Evolution* 2006; **38**(6):583–600.
11. Fang Z, Bailey RL. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *Forest Science* 2001; **47**:287–300(14).
12. Hall DB, Bailey RL. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *Forest Science* 2001; **47**:311–321(11).
13. Bortot P, Masarotto G, Scarpa B. Sequential predictions of menstrual cycle lengths. *Biostatistics* Apr 2010; **11**(4):741–755.
14. Fatti LP, Senaoana EM, Thompson ML. Bayesian updating in reference centile charts. *Journal of the Royal Statistical Society, Series A* 1995; **161**(1):103–15.
15. Rao CR. Prediction of future observations in growth curve models. *Statistical Science* 1987; **2**(4):434–447.
16. Pan H, Goldstein H. Multi-level models for longitudinal growth norms. *Statistics in Medicine* Dec 1997; **16**(23):2665–2678.
17. Overgaard R, Jonsson N, Tornoer C, Madsen H. Non-linear mixed-effects models with stochastic differential equations: Implementation of an estimation algorithm. *Journal of Pharmacokinetics and Pharmacodynamics* 2005; **32**:85–107.
18. Bauer R, Guzy S, C N. A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. *AAPS Journal* 2007; **9**(1):E60–E83.
19. Wolfinger R. Laplace's approximations for non-linear mixed-effect models. *Biometrika* 1993; **80**:791–795.
20. Kuhn E, Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis* 2005; **49**:1020–1038.
21. Girard P, Mentré F. A comparison of estimation methods in nonlinear mixed effects models using a blind analysis. *PAGE 14* 2005; **Abstr 834** [www.pagemeeeting.org/?abstract=834].
22. Hall DB, Clutter M. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. *Biometrics* Mar 2004; **60**(1):16–24.
23. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov chain Monte Carlo in practice*. Interdisciplinary Statistics, Chapman & Hall: London, 1996.
24. Robert C, Casella G. *Monte Carlo statistical methods*. Springer-Verlag: New York, 2002.
25. Panhard X, Samson A. Extension of the SAEM algorithm for nonlinear mixed models with two levels of random effects. *Biostatistics* 2009; **10**:121–135.
26. Brendel K, Comets E, Laffont C, Laveille C, Mentré F. Metrics for external model evaluation with an application to the population pharmacokinetics of gliclazide. *Pharmaceutical Research* Sep 2006; **23**(9):2036–2049.
27. Vonesh EF. A note on the use of Laplace's approximation for nonlinear mixed-effects models. *Biometrika* 1996; **83**:447–452.
28. Pinheiro J, Bates D. Approximations to the log-likelihood function in the non-linear mixed-effect models. *Journal of Computational and Graphical Statistics* 1995; **4**:12–35.
29. Hooper PM, Mayes DC, Demianczuk NN. A model for foetal growth and diagnosis of intrauterine growth restriction. *Statistics in Medicine* Jan 2002; **21**(1):95–112.

A. Asymptotical properties

Proposition 1 Denote $\mathcal{F}_n = \sigma(y_1, \dots, y_n)$ the natural filtration of the process for $n \geq 0$, and $\mathcal{F}_\infty = \bigcup_{n \geq 1} \mathcal{F}_n$. We have

(i) *Properties of the conditional distribution $p(\phi|y_{1:n})$. Let $\mathbb{E}_{\mathcal{F}_n}(\phi)$ and $\text{Var}_{\mathcal{F}_n}(\phi)$ denote the expectation and the variance of this distribution, respectively. We have*

- $\mathbb{E}_{\mathcal{F}_n}(\phi) \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}_{\mathcal{F}_\infty}(\phi)$
- $\text{Var}_{\mathcal{F}_n}(\phi) \xrightarrow[n \rightarrow \infty]{a.s.} \text{Var}_{\mathcal{F}_\infty}(\phi)$

(ii) *Properties of the predictive distribution $p(y_t|y_{1:n})$ for all $t > 0$:*

- $\mathbb{E}(y_t|y_{1:n}) \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}_{\mathcal{F}_\infty}(f(\phi, t))$
- $\forall n \in \mathbb{N}, \sigma^2 \mathbb{E}_{\mathcal{F}_\infty}(g^2(\phi_k, t)) \leq \text{Var}(y_{kt}|y_{1:n}) \leq \sigma^2 \mathbb{E}_{\mathcal{F}_\infty}(g^2(\phi_k, t)) + \mathbb{E}_{\mathcal{F}_\infty}(f^2(\phi_k, t)), \text{ for } k = 1, 2$

Proof:

(i) Let ξ be a function such that $\xi(\phi) \in \mathbb{L}^1$. Let's denote $X_n = \mathbb{E}(\xi(\phi)|\mathcal{F}_n)$. We have:

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}(\mathbb{E}(\xi(\phi)|\mathcal{F}_{n+1})|\mathcal{F}_n) = \mathbb{E}(\xi(\phi)|\mathcal{F}_n) = X_n$$

As X_n is a martingale, it converges and we have: $X_n \xrightarrow{a.s.} X_\infty$.

For $\xi(\phi) = \phi$, we have $\mathbb{E}(\phi|\mathcal{F}_n) \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}(\phi|\mathcal{F}_\infty)$. For $\xi(\phi) = \phi' \phi$ we have $(\phi' \phi|y_{1:n}) \xrightarrow[n \rightarrow \infty]{a.s.} (\phi' \phi|\mathcal{F}_\infty)$, hence the results regarding the variance.

(ii) Similarly, for $t > 0$ and $\xi(\phi) = f(\phi, t)$, $\mathbb{E}(y_t|y_{1:n})$ converges.

For all n we can write $\text{Var}(y_t|y_{1:n}) = \mathbb{E}(y_t^2|y_{1:n}) - \mathbb{E}^2(y_t|y_{1:n})$. For the first term:

$$\mathbb{E}(y_t^2|y_{1:n}) = \int y_t^2 p(y_t|y_{1:n}) dy_t = \iint y_t^2 p(y_t|\phi) dy_t p(\phi|y_{1:n}) d\phi.$$

This last expression simplifies since $\int y_t^2 p(y_t|\phi) dy_t = \mathbb{E}(y_t^2|\phi) = \sigma^2 g^2(\phi, t) + f^2(\phi, t)$. Therefore:

$$\begin{aligned} \mathbb{E}(y_t^2|y_{1:n}) &= \int \sigma^2 g^2(\phi, t) p(\phi|y_{1:n}) d\phi + \int f^2(\phi, t) p(\phi|y_{1:n}) d\phi \\ &= \sigma^2 \mathbb{E}(g^2(\phi, t)|\mathcal{F}_n) + \mathbb{E}(f^2(\phi, t)|\mathcal{F}_n) \end{aligned}$$

Similarly, in the second term:

$$\begin{aligned} \mathbb{E}(y_t|y_{1:n}) &= \int y_t p(y_t|y_{1:n}) dy_t = \int y_t \int p(y_t|\phi) p(\phi|y_{1:n}) d\phi dy_t \\ &= \iint y_t p(y_t|\phi) dy_t p(\phi|y_{1:n}) d\phi. \end{aligned}$$

This simplifies since $\int y_t p(y_t|\phi) dy_t = \mathbb{E}(y_t|\phi) = f(\phi, t)$.

$$\mathbb{E}(y_t|y_{1:n}) = \int f(\phi, t) p(\phi|y_{1:n}) d\phi = \mathbb{E}(f(\phi, t)|\mathcal{F}_n).$$

We finally have $\text{Var}(y_t|y_{1:n}) = \sigma^2 \mathbb{E}(g^2(\phi, t)) + \mathbb{E}(f^2(\phi, t)|\mathcal{F}_n) - \mathbb{E}^2(f(\phi, t)|\mathcal{F}_n)$.

Using Jensen's inequality with the convex function x^2 , the random variable $X = f(\phi, t)$ and the conditional expectation $\mathbb{E}(\cdot|\mathcal{F}_n)$, we have $\mathbb{E}^2(f(\phi, t)|\mathcal{F}_n) \leq \mathbb{E}(f^2(\phi, t)|\mathcal{F}_n)$ and hence the results. \square

B. Simulation study

Simulations are designed to illustrate the convergence properties described Proposition 1, using the estimators presented in (8) and (10) when the number of data increases. For the purpose of clarity, the simulations are performed using a 1-level NLMM, i.e. only one fetus per pregnancy. The results are similar with a 2-level NLMM.

B.1. Methods

To evaluate precisely the estimator, we assume in this simulation study that the population parameter θ is known from a reference population. Therefore the reference population does not need to be simulated, only the validation dataset is simulated. The moments of the predictive distribution estimated from the validation dataset are compared to their true values. Consequently, the estimated error includes only the estimation error of the predictive distribution and not the population parameter estimation error.

B.1.1. Data generation Data were generated according to the model described in Section 3 with only one level of variability, using a Weibull growth function [10]

$$f(\psi, t) = \alpha(1 - e^{-\beta t}) \quad (14)$$

with $\alpha > 0$, $\beta > 0$ and $\phi = (\log \alpha, \log \beta)$ to ensure the positiveness of the parameters. The error model is a homoscedastic model ($g(\psi, t) = 1$) and $\phi \sim \mathcal{N}(\mu, \Omega)$. Parameter values for $\theta = (\mu, \Omega, \sigma^2)$ are those obtained on a real dataset and described in Section 2. A dataset comprising $N = 100$ individuals with $n_i = n = 50$ equally spaced time-points was generated.

B.1.2. Definition of biases When n grows to infinity, the behavior of the estimation of each individual's predictive distribution $p(\phi_i | y_{i1:n})$ $i = 1, \dots, N$ was investigated by Monte Carlo estimation method. The bias of the estimates $\hat{\phi}_{ij}$ and $\hat{V}_{ij}(\phi)$ of $\mathbb{E}(\phi_i | y_{i1:n})$ and $V(\phi_i | y_{i1:n})$ defined by (7) were computed for each simulated individual $i = 1, \dots, N$ and for $j = 1, \dots, n$. The known value of simulated parameter for the i^{th} individual is denoted ϕ_i^{sim} . For a given number of available data $j = 1, \dots, n$, the empirical biases of $\hat{\phi}_{ij}$ and $\hat{V}_{ij}(\phi)$ are defined as:

$$\begin{aligned} (B_{\mathbb{E}})_j &= \frac{1}{N} \sum_{i=1}^N (\hat{\phi}_{ij} - \phi_i^{\text{sim}}) \\ (B_{\text{Var}})_j &= \frac{1}{N} \sum_{i=1}^N \hat{V}_{ij}(\phi) \end{aligned}$$

We also computed the relative Root Mean Squared Error (*RMSE*) between the true simulated observation y_{in}^{sim} at the n^{th} time-point and its prediction $\hat{f}_{ij}(\phi, t_n)$ obtained with $j = 1, \dots, n - 1$ data-points.

$$(RMSE)_j = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{y_{in}^{\text{sim}} - \hat{f}_{ij}(\phi, t_n)}{y_{in}^{\text{sim}}} \right)^2}$$

B.2. Results

The convergence of the estimated first two moments of the conditional distribution $p(\phi | y_{1:n})$ is depicted in Figure 5. The numerical convergence of $(B_{\mathbb{E}})_j$ towards zero describes the convergence of the expectation of the estimated conditional distribution when the number of given measurements increases (Proposition 1 (i)). Similarly, the variance of the estimated conditional distribution also rapidly converges towards zero thus demonstrating the increase in precision with accruing

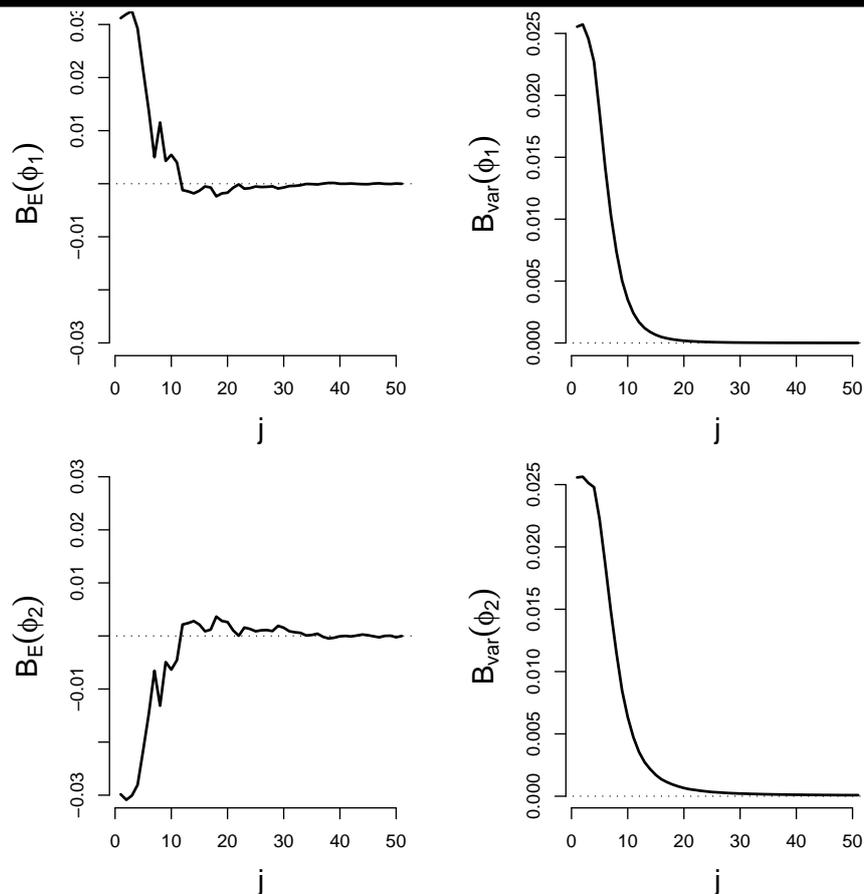


Figure 5. Simulated asymptotic behavior of the estimated predictions illustrating Proposition 1. Top panels: Asymptotic behavior of the bias of the estimated expectation $(B_E)_j$ of $\phi|y_{1:j}$ for $j = 1, \dots, n_i$, for each of the $p = 2$ components ($\log \alpha, \log \beta$) of ϕ respectively. Bottom panels: Asymptotic behavior of the bias of the variance $(B_{Var})_j$ of $\phi|y_{1:j}$ for $j = 1, \dots, n_i$. The X-axis describes the number of given observations j , the Y axis is the value of the bias.

data (Proposition 1 (i)). Both $(B_E)_j$ and $(B_{Var})_j$ are close to zero for $j \geq 15$ available measurements. Although 15 appear unreasonable in our context, the bias is already very limited, in the order of 0.03 in absolute value, with 3 or 4 measurements. Figure 6 plots $RMSE_j$ as a function of $j = 1, \dots, n - 1$. The convergence of $RMSE_j$ demonstrates the ability to predict a future observation with increasing precision when data add-up with time.

Using a 2GHz Intel Core 2 Duo T7250 CPU, the estimation of one predictive distribution with $M = 100000$ MCMC iterations requires 42 s.

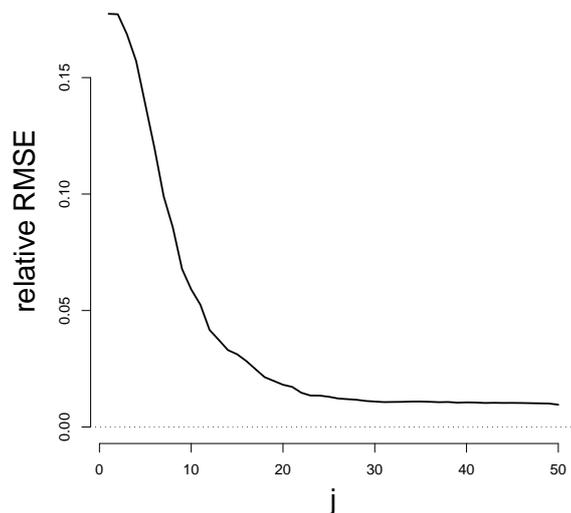


Figure 6. Asymptotical behavior of the standardized bias of the prediction $f(\hat{\phi}_i, t_n)$ of y_n when $j = 1, \dots, n - 1$, defined by the relative mean squared error (RMSE)