Heteroscedastic Growth Curve Modeling with Shape-Restricted Splines

JIEYING JIAO, WENLING SONG, YISHU XUE, AND JUN YAN*

Abstract

Growth curve analysis (GCA) has a wide range of applications in various fields where growth trajectories need to be modeled. Heteroscedasticity is often present in the error term, which can not be handled with sufficient flexibility by standard linear fixed or mixed-effects models. One situation that has been addressed is where the error variance is characterized by a linear predictor with certain covariates. A frequently encountered scenario in GCA, however, is one in which the variance is a smooth function of the mean with known shape restrictions. A naive application of standard linear mixed-effects models would underestimate the variance of the fixed effects estimators and, consequently, the uncertainty of the estimated growth curve. We propose to model the variance of the response variable as a shape-restricted (increasing/decreasing; convex/concave) function of the marginal or conditional mean using shape-restricted splines. A simple iteratively reweighted fitting algorithm that takes advantage of existing software for linear mixed-effects models is developed. For inference, a parametric bootstrap procedure is recommended. Our simulation study shows that the proposed method gives satisfactory inference with moderate sample sizes. The utility of the method is demonstrated using two real-world applications.

KEYWORDS AND PHRASES: iteratively reweighted, linear mixed-effects model, parametric bootstrap.

1. INTRODUCTION

Growth curve analysis (GCA) plays a critical role in various fields such as agronomy [43], animal science [34], biology [38], clinical trials [52], and psychological studies [9, 24, 6], among others. A GCA provides information about not only the mean but also the variation of the growth trend of a certain population. For example, reference growth charts for children's height, weight, and other physical characterics are widely used in wellness checks. A growth chart typically depicts a collection of quantiles of the distribution of physical characteristics of the reference population as a function of age. Accurate characterizations of the growth trajectory in both the mean level and the variation level are needed to make valid inferences and draw meaningful conclusions. The mean level of a growth trajectory has been extensively studied with a variety of functional forms such as fractional polynomial [16, 37] and smoothing splines [8]. In contrast, the variation has been studied but far less extensively.

Heteroscedasticity is a commonly encountered challenge in GCA. The error variance can be modeled as a smooth function of time or the mean response. Kernel-based methods have been used, which led to uniformly consistent estimator of the variance function [5, 26]. Covariates could be incorporated into the variance by an additional regression for the dispersion [40, 27]. For predictive purposes, a parametric distribution at any time point is often desired. The lambda-mu-sigma (LMS) method handles heteroscedasticity along with non-normality. In particular, it assumes that, after being standardized by a time-specific median μ and Box–Cox transformed with a time-specific power λ, the response follows a normal distribution with mean zero and time-specific standard deviation σ [7, 8]. The functions μ, σ, and λ are assumed to evolve smoothly with time, which can be modeled by splines of time. Distributions other than the normal distribution can be used with time-specific parameters in the generalized additive modeling framework for location, scale, and shape (GAMLSS) [31, 33, 32]. Quantile regression is a distribution-free method that directly models the age-specific quantiles of the response, possibly conditioning on covariates [48].

Clustered data, which often arise in GCA, bring an additional challenge of handling the within-cluster dependence. The number of repeated measures over time on the same subject can be as many as 30–40 [42]. The linear mixed-effects model

---

*Corresponding author.

(LMM) introduces within-cluster dependence through cluster-level random effects. An additional model on the variance or scale leads to the mixed-effects location-scale model [35, 17, 18]. The variance model is formulated with covariates or through a scale-mixture with, for example, an inverse gamma scale. The `lme()` function in R package **nmle** [28] can fit LMMs with heteroscedasticity in a set of pre-programmed forms [56, p.71–100], and the best form can be selected using Akaike information criterion (AIC) [1] or Bayesian information criterion (BIC) [39]. The covariance structure can be modeled directly with nonparametric methods for flexible shapes [e.g., 13]. The normal distribution assumption of the response variable can be relaxed by using GAMLSS with random effects in the additive terms [41, p.247–252]. The generalized estimating equation (GEE) method [22] focuses on the marginal modeling. It has been generalized to handle heteroscedasticity and within-cluster correlations [51], but marginal models in general are not suitable for subject-specific predictions.

Despite the extensive GCA literature, there are two limitations in routine analyses. The first is that there is no convenient way to put shape restrictions, such as monotonicity and/or convexity/concavity, on the variance in addition to the mean of a growth curve. Existing methods such as GAMLSS [41] allow flexible shapes in the mean or variance structure, but there is no direct way yet to ensure shape restrictions. If effectively used, such restrictions could improve the efficiency in inferences. The second limitation is that the existing methods do not have the flexibility to allow the variance to be modeled as a function of the mean, which is common in generalized linear models. The R package **nmle** [28] only allows the mean for independent data and the marginal mean for clustered data in the variance structure, but not the conditional mean given random effects for clustered data. Variance as a function of the mean could, again, improve the efficiency in inferences, especially when the mean depends on multiple covariates.

To break the aforementioned two limitations, we propose to model the variance in GCA as a shape-restricted function of the growth level [21]. The shape restrictions includes monotonicity or/and convexity/concavity, accommodated with shape-restricted splines such as monotone splines [30] or convex splines [25] with evenly spaced knots and constrained parameters. For clustered data, the variance model can incorporate the growth level either through the marginal mean or the conditional mean given the cluster-level random effects. Either AIC or BIC can be used to select the degrees of freedom of the splines and to select between marginal and conditional mean models. The parameters are estimated in an iteratively reweighted fitting algorithm. The performance of the proposed methods is validated through an extensive simulation study and applications to two real examples.

The rest of this paper is organized as follows. Section 2 gives a review of the shape-restricted spline basis. Growth models with shape-restricted Heteroscedasticity for both independent and clustered data are presented in Section 3. A simulation study is reported in Section 4 to assess the performance of the methods. We illustrate the use of the proposed approach with the fetal pancreas length data and the chicken weight data in Section 5. A discussion concludes in Section 6. The computing code is publicly available at https://github.com/JieyingJiao/GCA_Code.

2. SHAPE-RESTRICTED SPLINES

Shape-restricted splines are piecewise polynomials, differentiable up to a certain degree, that satisfy restrictions on shapes such as monotonicity and convexity. They offer great flexibility in approximating unknown smooth curves with known shape restrictions. Specifically, a shape-restricted curve is approximated by a linear combination of a set of shape-restricted spline bases, where the coefficients are restricted to get the desired pattern. Before introducing our proposed method, which employs the I-spline bases and C-spline bases, we first briefly review how they are constructed.

To define shape-restricted spline bases, we start from M-splines. M-spline bases are standardized versions of B-spline bases so that they integrate to 1 [10]. An M-spline of degree k over an interval $[l, u]$ is defined recursively as

$$M_i^{(1)}(x) = \begin{cases} \dfrac{1}{t_{i+1} - t_i}, & t_i \leq x \leq t_{i+1}, \\ 0, & \text{otherwise}, \end{cases}$$

$$M_i^{(k)}(x) = \begin{cases} \dfrac{k[(x - t_i)M_i^{(k-1)}(x) + (t_{i+k} - x)M_{i+1}^{(k-1)}(x)]}{(k-1)(t_{i+k} - t_i)}, & t_i \leq x \leq t_{i+k} \\ 0, & \text{otherwise}, \end{cases}$$

$i = 1, \ldots, m + 2k$, where t_i's are the knots with

$$l = t_1 = \cdots = t_k < \cdots < t_{m+k+1} = \cdots = t_{m+2k} = u,$$

and m is the number of internal knots. The M-spline bases are positive over $[l, u]$. A linear combination of M-spline bases with nonnegative coefficients is non-negative. Same as B-splines, it is continuously differentiable up to $k - 1$ times for $k \geq 1$.

I-splines are integrals of M-splines [30]. The I-spline bases with degree k over the interval $[l, u]$ are

$$I_i^{(k)}(x) = \int_l^x M_i^{(k)}(s)\, \mathrm{d}s, \quad l \leq x \leq u, \qquad i = 1, \ldots, m + 2k.$$

Because their derivatives are M-splines, which are non-negative, I-spline bases can be used for modeling monotonic functions. A linear combination of I-spline bases with non-negative (or non-positive) coefficients are non-decreasing (or non-increasing). An intercept is always needed when using I-spline bases since their lowest order is linear.

C-splines are integrals of I-splines [25]. The C-spline bases with degree k over the interval $[l, u]$ are

$$C_i^{(k)}(x) = \int_l^x I_i^{(k)}(s)\, \mathrm{d}s, \quad l \leq x \leq u, \qquad i = 1, \ldots, m + 2k.$$

This set of bases does not have a linear or a constant term, both of which need to be added when fitting curves. With restrictions on the coefficients, C-splines can be used to approximate functions with specific combinations of monotonicity (increasing or decreasing) and shape (convexity or concavity) [46]. A commonly seen pattern in growth curve is non-decreasing concave, which can be implemented by restricting the first derivatives to be positive and second derivatives to be negative.

In implementation, we used the I-spline and C-spline bases from R package **splines2** [46]. As illustrated later, the degrees of freedom can be chosen by AIC or BIC. For a typical GCA, a moderately complicated pattern can be approximated by spline bases with a few (3–5) degrees of freedom.

3. GCA WITH SHAPED-RESTRICTED HETEROSCEDASTICITY

For clarity of presentation, we start from the independent data setting and then consider the clustered data setting. Inferences and model selection come next.

3.1 Model for Independent Data

Suppose that a random sample of n subjects are observed, with one measure each at a random time point. Specifically, let the observed data be $\{y_i, t_i : i = 1, \ldots, n\}$, where y_i is the measure at time t_i from the ith subject. Additional covariates are possible, such as gender or treatment group. In a unified linear model setting, the spline bases generated from the time and additional covariates can be represented by a p-dimensional $\mathbf{x}_i$ for the ith subject. To introduce heteroscedasticity, we use a smooth function $g(\nu_i)$ to characterize the standard deviation of the regression error of the ith subject, where v_i is some index variable. The index variable can be observed such as time, or unobserved such as the mean of the corresponding subject from the linear model. The smooth function $g(\cdot)$ is parametrized by shape-restricted splines.

Using I-splines as an example, a heteroscedastic linear model is

$$\begin{aligned}
y_i &= \mathbf{x}_i^\top \beta + \varepsilon_i, \quad i = 1, \ldots, n, \\
\varepsilon_i &\sim \mathrm{N}(0, g^2(v_i, \theta)), \\
g(v_i, \theta) &= \theta_0 + \sum_{k=1}^K \theta_k I_k(v_i),
\end{aligned} \tag{3.1}$$

where β is a p-dimensional regression coefficient vector for $\mathbf{x}_i$ (with shape restrictions if needed), ε_i is the normally distributed regression error with mean zero and standard deviation $g(\nu_i; \theta)$, $\{I_k(\cdot), k = 1, \ldots, K\}$ are I-spline bases with K degrees of freedom, and $\theta = (\theta_0, \ldots, \theta_k)$ is a $(K + 1)$-dimensional coefficient vector. The coefficients θ can be restricted to control the shape of the heteroscedasticity as a function of ν_i. For example, if the variance increases with the mean (or time), the coefficients θ can be restricted to be non-negative. Like in standard GCA, the covariates $\mathbf{x}_i$ in the mean model may include shape-restricted spline bases as well, in which case, their coefficients β need to be restricted similarly.

If concavity or convexity is desired, the I-splines can be replaced with C-splines and a linear term of time with appropriate restrictions on the coefficients, as introduced in Section 2.

3.2 Model for Clustered Data

When more than one measures were collected from each of the n subjects, the observed data will have a clustered structure. Let the observed data be $\{\mathbf{y}_i, \mathbf{t}_i, i = 1, \ldots, n\}$, where $\mathbf{y}_i$ is a vector of n_i dimensions showing the measures from the ith subject, and $\mathbf{t}_i$ is the corresponding time vector. As before, we still use spline bases to estimate the error variance,

---

Algorithm 1 Iteratively reweighted fitting algorithm for clustered data.

▷ **Input** $\{\mathbf{y}_i, \mathbf{X}_i, \mathbf{Z}_i, i = 1, \ldots, n\}$.
1: **procedure**
2:     Fit a linear mixed-effects model without weight.
3:     Get estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, estimate $\hat{\boldsymbol{\alpha}}$ of $\boldsymbol{\alpha}$, residuals $\boldsymbol{e}_i$, and fitted (marginal or conditional) mean $\hat{\boldsymbol{\mu}}_i$ $i = 1, 2, \ldots, n$.
4:     **repeat**
5:         Treat residuals $\boldsymbol{e}_i$'s as an observation from $\mathrm{N}(0, g^2(\hat{\boldsymbol{\mu}}_i, \boldsymbol{\theta}))$, $i = 1, 2, \ldots, n$.
6:         Get maximum likelihood estimate $\hat{\boldsymbol{\theta}}$ with monotone constraints that $\boldsymbol{\theta} > 0$.
7:         Fit a linear mixed-effects model with weight $\{g^{-1}(\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\theta}}), \ i = 1, 2, \ldots, n\}$
8:         Get updated $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\alpha}}$, $\boldsymbol{e}_i$, and $\hat{\boldsymbol{\mu}}_i$, $i = 1, 2, \ldots, n$.
9:     **until** $\hat{\boldsymbol{\beta}}$ converges.
10: **end procedure**
▷ **Output** $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\alpha}}$, and $\hat{\boldsymbol{\theta}}$.

---

but with a linear mixed-effects model to account for the dependence structure within the dataset. Fixed effects covariates can contain spline bases generated from time $\mathbf{t}_i$. Again, using I-Splines as an example, the model for the ith subject is

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \varepsilon_i, \quad i = 1, \ldots, n,$$
$$\mathbf{b}_i \sim \mathrm{N}(\mathbf{0}, \ \mathbf{B}(\boldsymbol{\alpha})),$$
$$(\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in_i})^\top = \varepsilon_i \sim \mathrm{MVN}\left(\mathbf{0}, \ \mathrm{diag}(g^2(v_{i1}, \theta), \ldots, g^2(v_{in_i}, \theta))\right),$$
$$g(v_{ij}, \theta) = \theta_0 + \sum_{k=1}^{K} \theta_k I_k(v_{ij}),$$

(3.2)

where $\mathbf{X}_i$ is an $n_i \times p$ covariate matrix for fixed effects, $\mathbf{Z}_i$ is an $n_i \times q$ design matrix for random effects, β is a p dimensional fixed effects vector, $\mathbf{b}_i$ is a q-dimensional random effects vector with covariance matrix $\mathbf{B}$ parameterized by vector $\boldsymbol{\alpha}$, and MVN is the multivariate normal distribution. Other notations are the clustered analogs to those in Equation (3.1).

A special choice for the index variable v_{ij} is the mean of the response variable. For a mixed-effects model, this mean can be conditional on the random effects or not. If random effects are conditioned on, the mean is

$$\mu_{i,\mathrm{c}} = E[\mathbf{y}_i \mid \mathbf{b}_i] = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i;$$

(3.3)

otherwise it is

$$\mu_{i,\mathrm{m}} = E[\mathbf{y}_i] = \mathbf{X}_i\beta.$$

(3.4)

For ease of referencing, we call them conditional mean and marginal mean, respectively. When the error variance changes with the marginal mean, the response variable still has a multivariate normal distribution. If the conditional mean is in the error variance structure, there is dependence between the random effects and the error term, and the response variable no longer has a multivariate normal distribution. To calculate the likelihood function for this situation, as needed in AIC and BIC calculations, numerical integration is needed. See details in Section 3.3.

3.3 Inference

The maximum likelihood method can be used to get parameter estimates in theory as long as appropriate restrictions on the coefficients are imposed to enforce the shape restrictions. To obtain the maximum likelihood estimator, we propose an iteratively reweighted fitting procedure that takes advantage of existing software packages for linear mixed-effects models allowing weights. The steps are summarized in Algorithm 1 for clustered data when the error variance is changing with the mean. We use $\boldsymbol{\mu}_i$ in the algorithm to represent either the conditional mean or the marginal mean, and $\hat{\boldsymbol{\mu}}_i$ for the estimated value of the mean. Algorithm for independent data is similar and simpler, and will not be repeated here. The shape restrictions on the heteroscedasticity (and the mean model) can be imposed with a constrained optimizer, such as the `constrOptim()` function in R.

To construct reference quantiles in a GCA, we suggest using parametric bootstrap. It avoids deriving the likelihood function and the Hessian matrix [36, p. 227], which is challenging when the error variance changes with the conditional mean. This is of particular importance when some of the estimated θ's are on the boundaries of the constrained parameter space [2]. Since we focus on the final fitted curve instead of the basis coefficients, bootstrap provides a natural solution for the quantitles of the fitted curve regardless of whether some of the estimated θ's are on the boundaries.

Algorithm 2 Steps to get one parametric bootstrap sample for clustered data.

▷ **Input** $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\alpha}}$, and $\hat{\boldsymbol{\theta}}$.

1: **procedure**

2:     Generate random effects $\boldsymbol{b}_i^*$'s from N$(0, \mathbf{B}(\hat{\boldsymbol{\alpha}}))$, $i = 1, 2, \ldots, n$.

3:     Let $\boldsymbol{\mu}_i^* = (\mu_{i1}^*, \ldots, \mu_{in_i}^*)$ be $\mathbf{X}_i\hat{\beta} + \mathbf{Z}_i\mathbf{b}_i^*$ when use conditional mean, or $\mathbf{X}_i\hat{\beta}$ when use marginal mean.

4:     Generate error terms $\boldsymbol{\varepsilon}_i^*$ from MVN $\left(\mathbf{0}, \ \text{diag}(g^2(\mu_{i1}^*, \hat{\boldsymbol{\theta}}), \ldots, g^2(\mu_{in_i}^*, \hat{\boldsymbol{\theta}}))\right)$, $i = 1, 2, \ldots, n$.

5:     Let $\boldsymbol{y}_i^* = \mathbf{X}_i\hat{\beta} + \mathbf{Z}_i\mathbf{b}_i^* + \varepsilon_i^*$, $i = 1, 2, \ldots, n$.

6:     Apply Algorithm 1 to $\{\boldsymbol{y}_i^*, \mathbf{X}_i, \mathbf{Z}_i, i = 1, 2, \ldots, n\}$ and record the output $\hat{\boldsymbol{\beta}}^*$, $\hat{\boldsymbol{\alpha}}^*$, and $\hat{\boldsymbol{\theta}}^*$.

7: **end procedure**

▷ **Output** One bootstrap copy $\{\hat{\boldsymbol{\beta}}^*, \hat{\boldsymbol{\alpha}}^*, \hat{\boldsymbol{\theta}}^*\}$.

Algorithm 2 summarizes the steps to get one parametric bootstrap sample for clustered data. Same as before, the algorithm for independent data will be similar and will not be displayed here. Repeating this process gives a large sample of bootstrap copies of the point estimates of the model parameters. Their empirical standard deviations are then used as the standard errors of the model parameter estimates. For mixed-effects models, parametric bootstrap method has better performance compared with bootstrap methods that only re-sample observations or residuals, as it produces more accurate standard deviation of estimated parameters, and closer-to-nominal coverage rates for confidence intervals [11, 50, 44].

After the model fitting process, model checking can be done using the residuals. The standardized residuals, i.e., the residuals divided by the estimated error standard deviation, should follow a standard normal distribution, and their normality can be checked visually using a normal Q-Q plot, or other normality testing methods.

3.4 Model Selection

With the same set of data, we can fit different models by using spline bases with different degrees of freedom and/or locations of knots. For clustered data, candidate models can either have the conditional mean or the marginal mean in the error variance function. In order to choose the best one from these potential models, AIC or BIC can be used as model selection criteria. They consider both the fitting accuracy and the model complexity:

$$\text{AIC} = -2\log L + 2P,$$
$$\text{BIC} = -2\log L + P\log n,$$

where L is the likelihood function of the fitted model, P is the number of parameters, and n is the sample size. Models with smaller AIC or BIC are preferred.

Special attention needs to be paid when the model has error variance changing with the conditional mean (3.3) since the error variance depends on the random effects. The likelihood is

$$L = \prod_{i=1}^{n} \int_{-\infty}^{+\infty} f(\mathbf{y}_i \mid \mathbf{b}_i) f(\mathbf{b}_i) \, d\mathbf{b}_i,$$

where $f(\mathbf{y}_i \mid \mathbf{b}_i)$ is the probability density function (pdf) of the response vector $\mathbf{y}_i$ conditioning on the random effects $\mathbf{b}_i$, and $f(\mathbf{b}_i)$ is the pdf of the random effects $\mathbf{b}_i$. From Model (3.2), it is easy to see that $f(\mathbf{b}_i)$ is N$(\mathbf{0}, \ \hat{\mathbf{B}})$, and $f(\mathbf{y}_i \mid \mathbf{b}_i)$ is N$(\mathbf{X}_i\hat{\beta} + \mathbf{Z}_i\mathbf{b}_i, g^2(\mathbf{v_i}, \hat{\theta}))$. There is no closed-form result for this integral, but it can be numerically calculated to obtain the likelihood.

4. SIMULATION STUDY

The proposed methods were validated with an extensive simulation study which covers the most commonly seen scenarios for both independent data and clustered data.

4.1 Independent data

This study mimics a scenario where each subject has one measure; see the fetal pancreas length application in Section 5.1. The data generating model was

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i, \quad i = 1, \ldots, n,$$

Table 1. Summary of simulation results for the independent data scenario: SE is the empirical standard error; $\hat{\text{SE}}$ is the average of bootstrap standard errors, and CP is the empirical coverage percentage of 95% confidence intervals.

n	g	coef	naive method ($\times 10^{-2}$)				proposed method ($\times 10^{-2}$)			
			bias	se	$\hat{\text{se}}$	CP	Bias	SE	$\hat{\text{SE}}$	CP
100	g_1	β_0	0.6	7.5	7.4	96.0	0.0	2.3	2.7	97.0
		β_1	−0.4	8.8	8.8	94.2	−0.1	7.1	7.0	94.5
		β_2	−0.4	8.0	7.9	93.7	−0.0	5.0	5.0	94.0
	g_2	β_0	−0.0	10.0	9.7	93.6	−0.1	2.3	2.0	90.6
		β_1	0.1	10.8	10.5	94.8	0.1	6.2	6.0	94.1
		β_2	0.1	10.6	10.2	92.5	0.2	4.3	4.1	93.5
	g_3	β_0	−0.1	8.0	7.5	94.6	−0.3	3.7	3.4	90.0
		β_1	−0.1	9.9	9.9	94.7	−0.1	9.3	9.0	93.7
		β_2	0.2	9.1	8.4	91.9	0.4	6.7	6.4	93.6
200	g_1	β_0	−0.0	5.5	5.2	93.5	−0.0	1.6	1.9	97.7
		β_1	0.0	6.2	6.2	95.0	−0.0	5.1	4.9	94.1
		β_2	−0.0	6.0	5.6	92.7	−0.0	3.5	3.5	94.6
	g_2	β_0	0.1	6.9	7.0	95.9	−0.0	1.5	1.4	92.3
		β_1	−0.0	7.2	7.5	95.0	0.1	4.1	4.2	95.3
		β_2	−0.1	7.4	7.3	94.5	0.0	2.9	2.9	94.7
	g_3	β_0	−0.1	5.5	5.3	94.8	−0.0	2.7	2.4	92.3
		β_1	0.2	6.9	7.0	95.0	−0.1	6.3	6.2	94.5
		β_2	0.1	6.1	5.9	94.3	0.0	4.7	4.5	92.7

where y_i is ith response, x_{1i} is a Bernoulli(0.5) variable, x_{2i} is a Uniform$(0,2)$ variable, $(\beta_0, \beta_1, \beta_2) = (1,1,1)$, and ε_i is a zero-mean normally distributed error term with heteroscedasticity. Specifically, the standard deviation of ε_i is $g(\mu_i)$, where $\mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}$. Three functional forms were considered for g: $g_1(\mu) = 0.25(\mu - 0.9)$, $g_2(\mu) = 0.02(\mu^3 + 1.2)$, and $g_3(\mu) = 0.1(5\Phi(\mu - 2)/0.3) + 1)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The specific parameter values in these functions were chosen such that the function values were positive, the ratio of the maximum value over the minimum value was around 30, and the resulting signal-to-noise ratio in the linear regression model was around 3. Two sample sizes were considered $n \in \{100, 200\}$.

For each configuration, 1000 datasets were generated. For comparison, both the naive linear regression model with heteroscedasticity ignored and the proposed GCA with shape-restricted heteroscedasticity were fitted to each dataset. For the naive model, the variances of the regression coefficients were obtained using the robust estimator to account for the heteroskedasticity [49, 23] as implemented in the R package `sandwich` [55, 53, 54]. In the proposed method, I-spline bases were used to enforce monotonicity. As more degrees of freedom were needed for fitting more complex patterns with satisfying accuracy, we picked quadratic I-spline basis with 2, 3, and 7 degrees of freedom when fitting the three $g(\cdot)$ patterns, respectively, with evenly spaced internal knots. Parametric bootstrap was used to calculate the standard deviation of the estimates and to construct 95% confidence interval of the regression coefficients and the error variance curve. The number of bootstrapping replicates was 1000.

Table 1 summarizes the empirical bias, empirical standard error (se), estimated standard error ($\hat{\text{se}}$), and the coverage percentage (CP) of the 95% confidence intervals of parameter estimates. The bias from the proposed method is close to zero under all the scenarios. The estimated standard deviations from parametric bootstrap $\hat{\text{se}}$ is close to the empirical value from the proposed method, and the CP is close to the nominal level 95%. The point estimate of β_2 has lower variation than that for β_1, which is expected because the continuous covariate x_2 provides more information than the binary x_1. As sample size increases, all standard errors decreases. In comparison with proposed method, the naive method leads to much higher standard deviations in the regression coefficient estimation. Although the 95% confidence intervals from the naive method seem to have appropriate coverage percentage, they are much longer than those from the proposed method as evident from the standard errors.

Figure 1 displays the estimated heteroscedasticity form and the averaged point-wise 95% confidence intervals from the proposed method using parametric bootstrap. The averaged estimates represented by the solid lines are close to the true curves in dashed lines, and the true curves lie within the 95% point-wise confidence intervals. The wider intervals near the right boundary, especially in the case of g_3, reflect that this is a challenging situation; the functional form is convex first and concave later, which needs more degrees of freedom to capture.
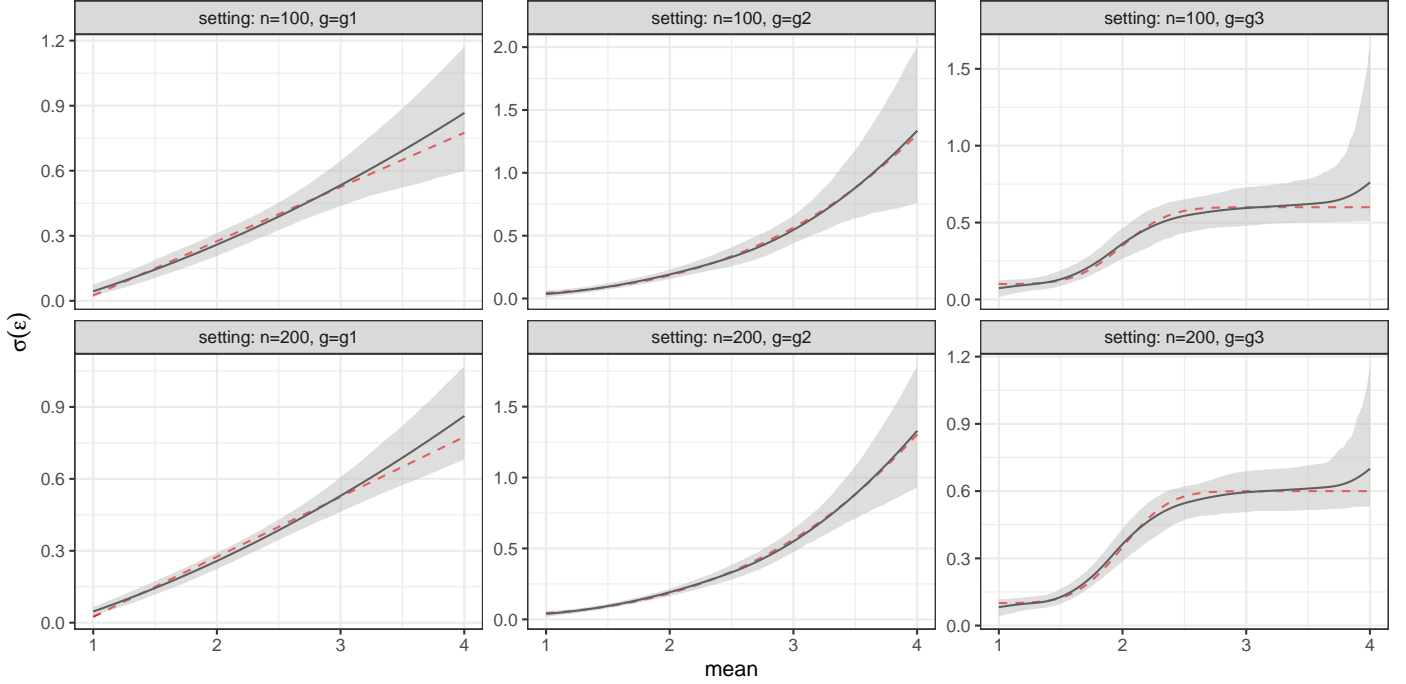
Figure 1: Error variance estimation for independent data simulation. The red dashed line is the truth, and the black solid line is the estimation. The grey region is the 95% point-wise confidence interval.

4.2 Clustered data

For clustered data, we considered a setting where each subject has repeated measures; see the chicken weight application in Section 5.2. The data generating model was a mixed-effects model with a random effect at the subject level,

$$y_{ij} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2ij} + b_i + \varepsilon_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, 5,$$

where y_{ij} is the response variable for the jth observation from the ith subject, x_{1i} is a subject-level covariate generated from Bernoulli(0.5), x_{2ij} is an observation-level continuous covariate generated from Uniform(0, 5), $(\beta_0, \beta_1, \beta_2) = (1, 1, 1)$, b_i is the subject-level random effect generated from $N(0, \sigma_b^2)$ with $\sigma_b = 0.1$, and ε_{ij} is a zero-mean normally distributed error term with heteroscedasticity. The standard deviation of ε_{ij} was set to be $g(\nu_{ij})$, where $g \in \{g_1, g_2, g_3\}$ is same as in the independent data setting, ν_{ij} is either the conditional mean or the marginal mean of the jth observation of the ith subject defined in Equation (3.3) and (3.4), respectively. The total number of subjects was set to be again $n \in \{100, 200\}$.

Three different estimation methods were compared. The first is the naive method that fits a linear mixed-effects model with constant error variance. To better capture the standard error, the robust sandwich estimator is used [45]. The second method uses the function `lme()` from the R package **nlme**, which provides some preset forms of heteroscedasticity. For settings where g_1 and g_2 are used and μ is the conditional mean, we used the true setup to specify the error variance when fitting the model. For other settings, the correct setup is not available in the `lme()` function, but we still used a power function of the conditional mean to specify the heteroscedasticity. The third method is the proposed method, where quadratic I-spline bases were used with degrees of freedom 5 for g_1 and g_2, and 7 for g_3. Internal knots of spline bases were chosen to be evenly spaced. The number of replications for parametric bootstrap was 1000. For each configuration, results were obtained for 1000 datasets.

Tables 2 and 3 summarize the simulation results for the variance as a function of marginal mean and conditional mean, respectively. All three methods seem to give unbiased point estimates for the regression coefficients, as they all have correct specification of the regression model. Their differences are in their uncertainty levels and coverage percentages of the 95% confidence intervals. The naive method has the worst performance since it did not consider the heteroscedasticity at all. The method with function `lme()` performs better than the naive method, but is still not satisfactory due to the misspecified heteroscedasticity form through the limited choices offered by `lme()`. Even in settings where the heteroscedasticity is correctly specified by the power function $g_2()$ with conditional mean as the index variable, see Table 3, its coverage

Table 2. *Summary of simulation results when error variance changes with the marginal mean: SE is the empirical standard error; $\hat{SE}$ is the average of bootstrap standard errors, and CP is the empirical coverage percentage of 95% confidence intervals.*

n	g	coef	naive method ($\times 10^{-2}$)				lme ($\times 10^{-2}$)				proposed method ($\times 10^{-2}$)			
			Bias	SE	$\hat{SE}$	CP	Bias	SE	$\hat{SE}$	CP	Bias	SE	$\hat{SE}$	CP
100	g_1	β_0	0.1	8.8	8.4	94.0	−1.6	6.9	6.6	93.4	−0.1	6.9	6.6	93.4
		β_1	−0.2	10.0	9.5	94.2	0.1	8.3	8.4	95.4	−0.2	8.3	8.2	94.3
		β_2	−0.1	3.5	3.3	93.3	0.2	3.0	2.8	92.3	−0.0	3.0	2.9	92.8
		σ_b	−0.7	9.9			0.4	8.7			−0.9	8.7		
	g_2	β_0	−0.4	9.5	9.7	95.2	−2.9	4.3	3.5	81.0	−0.1	4.0	3.7	93.0
		β_1	0.1	10.4	10.4	94.6	0.2	5.0	5.6	96.2	−0.1	4.9	4.9	94.7
		β_2	0.2	4.4	4.3	93.6	1.1	2.6	2.2	87.6	0.1	2.4	2.3	93.6
		σ_b	−0.4	10.8			4.6	2.8			−0.1	5.4		
	g_3	β_0	−0.1	7.1	7.1	95.0	−3.8	6.4	5.1	82.5	−0.1	5.4	5.2	93.8
		β_1	0.3	9.0	9.0	95.4	1.3	6.3	7.2	97.6	0.1	5.8	6.1	95.2
		β_2	0.0	3.1	3.1	94.5	1.1	2.8	2.4	88.8	0.0	2.5	2.6	95.1
		σ_b	−1.2	9.2			6.0	6.6			−0.8	6.7		
200	g_1	β_0	0.2	6.1	6.0	94.9	−1.3	4.9	4.6	92.9	0.2	4.8	4.7	94.0
		β_1	−0.4	6.8	6.8	95.3	−0.0	5.8	5.9	95.5	−0.3	5.8	5.8	95.2
		β_2	0.0	2.3	2.4	95.7	0.3	2.0	2.0	94.3	−0.0	2.0	2.0	94.8
		σ_b	−0.8	8.6			0.5	7.3			−0.5	7.5		
	g_2	β_0	0.0	6.9	6.9	94.7	−2.7	3.1	2.5	75.7	0.1	2.8	2.7	93.7
		β_1	−0.0	7.3	7.4	94.8	0.2	3.7	4.0	96.3	−0.1	3.6	3.5	94.2
		β_2	0.0	3.1	3.1	94.3	1.0	1.8	1.6	87.8	−0.0	1.6	1.6	94.1
		σ_b	−0.9	9.1			4.9	2.0			0.8	3.9		
	g_3	β_0	0.1	4.9	5.0	95.2	−3.9	4.6	3.6	74.6	0.0	3.8	3.7	94.5
		β_1	−0.0	6.3	6.4	95.2	1.3	4.8	5.1	96.0	0.1	4.5	4.3	94.0
		β_2	−0.0	2.2	2.2	95.4	1.1	2.0	1.7	86.4	−0.0	1.8	1.8	95.5
		σ_b	−1.3	8.1			7.1	4.8			−0.5	5.7		

percentages for β_2 are much lower than 95%. In contrast, in all three settings, the proposed method gives point estimates with lower standard errors as well as confidence intervals with coverage percentage close to 95%. As the sample size n increases, the performance becomes better as expected.

Also reported in Tables 2 and 3 are estimates of the standard deviation σ_b of the random effects. It is known that the standard errors of the random-effects variance parameters are hard to get and confidence intervals constructed from profile likelihood or parametric bootstrap are preferred [4]. So here we focus on the point estimate of σ_b. When the heteroscedasticity takes more complicated forms such as g_2 and g_3, the proposed method has much smaller bias than the lme() method. In some settings such as $g = g_2$; the lme() estimates of σ_b have bias but lower variation compared to those from the proposed method. This echos that caution is needed when using standard errors of the random effect variance. The empirical standard errors of the point estimates from the proposed method decrease as the sample size increases, but apparently not at the rate of $1/\sqrt{n}$, suggesting that a larger sample size is needed for the asymptotic properties of the random-effect variance estimator to hold.

Figure 2 displays the fitted heteroscedasticity from the proposed method. The two panels show the results with the index variables being the marginal mean and the conditional mean, respectively. Similar to the independent data scenarios, the estimated curve is close to the true curve, and is within the averaged 95% point-wise confidence intervals. The intervals are narrower than those in the independent data scenarios as the repeated measures provide more information.

5. APPLICATIONS

5.1 Fetal Pancreas Length

Fetus pancreatic dysplasia and hypertrophy, i.e., abnormality of fetal pancreas, are associated with congenital malformations [29, 19]. The fetal pancreas growth curve is a critical tool for prenatal screening for disorders. The growth pattern of fetal pancreas' lengths during the prenatal period has been investigated [12, 20, 15], but not the changing variation of the measurements. The proposed method allows capturing the fetal pancreas' growing patterns in both the mean level and the variation level.

Table 3. *Summary of simulation results when the error variance changes with the conditional mean: SE is empirical standard error; $\hat{\text{SE}}$ is the average of bootstrap standard errors; and CP is the empirical coverage percentage of 95% confidence interval.*

n	g	coef	naive method ($\times 10^{-2}$)				lme ($\times 10^{-2}$)				proposed method ($\times 10^{-2}$)			
			Bias	SE	$\hat{\text{SE}}$	CP	Bias	SE	$\hat{\text{SE}}$	CP	Bias	SE	$\hat{\text{SE}}$	CP
100	g_1	β_0	−0.3	8.5	8.5	95.1	−1.8	6.7	6.6	93.9	−1.3	6.8	6.8	94.4
		β_1	0.5	9.4	9.5	94.6	0.7	8.0	8.3	95.3	0.5	8.0	8.3	95.7
		β_2	0.1	3.4	3.3	95.1	0.4	2.9	2.8	94.9	0.3	2.9	2.9	94.5
		σ_b	−1.5	9.4			−0.3	8.4			−1.4	8.7		
	g_2	β_0	0.4	9.8	9.7	94.2	−2.5	4.3	3.5	83.1	−0.5	4.1	3.8	92.7
		β_1	−0.2	10.3	10.4	94.4	0.1	5.1	5.6	97.2	−0.4	5.0	5.0	94.8
		β_2	−0.1	4.4	4.3	93.9	0.9	2.6	2.2	87.5	0.0	2.4	2.3	92.7
		σ_b	−0.7	10.7			4.5	3.0			−0.1	5.4		
	g_3	β_0	0.0	7.4	7.1	93.7	−3.7	6.6	5.1	81.2	−0.3	5.6	5.3	94.0
		β_1	−0.3	9.4	9.0	93.9	0.9	6.4	7.2	97.6	−0.5	6.2	6.1	94.6
		β_2	0.2	3.2	3.1	94.0	1.2	3.0	2.4	87.3	0.0	2.7	2.6	95.4
		σ_b	−1.9	9.0			6.0	6.7			−1.0	6.7		
200	g_1	β_0	0.0	6.0	6.0	94.9	−1.4	4.9	4.6	92.4	−1.0	4.8	4.8	94.7
		β_1	0.0	6.7	6.7	94.7	0.4	5.8	5.9	95.4	0.3	5.8	5.8	94.2
		β_2	0.0	2.4	2.4	95.1	0.3	2.0	2.0	94.3	0.2	2.0	2.1	95.6
		σ_b	−1.9	8.1			0.1	7.2			−0.9	7.6		
	g_2	β_0	0.1	6.9	6.9	93.8	−2.9	3.0	2.5	73.5	−0.7	2.8	2.7	92.7
		β_1	−0.1	7.4	7.4	95.7	0.4	3.7	4.0	96.8	−0.2	3.6	3.5	95.2
		β_2	−0.0	3.1	3.1	94.5	1.1	1.8	1.6	85.0	0.0	1.7	1.6	94.9
		σ_b	−0.9	9.4			4.6	2.0			0.3	4.1		
	g_3	β_0	0.1	5.3	5.0	93.4	−4.1	4.9	3.6	72.8	−0.3	3.9	3.8	94.0
		β_1	−0.2	6.8	6.4	93.6	1.3	4.8	5.1	95.4	−0.1	4.4	4.3	94.7
		β_2	0.0	2.4	2.2	93.4	1.2	2.1	1.7	83.8	−0.1	1.9	1.8	94.3
		σ_b	−1.5	8.1			7.2	5.1			−0.4	5.6		

Data were collected from 44 pregnant women at different stages of pregnancy who visited the Second Hospital of Jilin University in China during April to July of 2012. No patient showed any external pathology or anomaly. The dataset contains a single measure of the fetal pancreas' length from each patient. Figure 3(I) shows the pancreas' lengths in millimeters versus the pregnant duration in days. It is reasonable to assume that the growth speed slows down as the fetus matures, in which case the growth curve would be increasing and concave. It is also reasonable to assume that the variation increases with time. Therefore, we use C-splines to model the mean growth level and use I-splines for the heteroscedasticity. Specifically, the model is

$$y_i = \beta_0 + \beta_1 t_i + \sum_{k=1}^{K_1} \beta_{k+1} C_k(t_i) + \varepsilon_i,$$

$$\varepsilon_i \sim \text{N}\left(0, \quad \left(\theta_0 + \sum_{k=1}^{K_2} \theta_k I_k(t_i)\right)^2\right),$$

where y_i is the i_{th} length measurement, t_i is the corresponding time with linear coefficient β_1, $C_k(t_i)$ is the kth C-splines basis evaluated at t_i with coefficient β_{k+1}, $k = 1, \ldots, K_1$, ε_i is the independent error term with heteroscedasticity, and $I_k(t_i)$ is the kth I-spline basis evaluated at t_i with coefficient θ_k, $k = 1, \ldots, K_2$. The internal knots of both spline bases were evenly spaced. The degrees of freedom K_1 and K_2 for the C-splines and I-splines, respectively, were both chosen to be 4 by BIC.

Figure 3 shows fitted results for the growth curve of the pancreas length. By shape restrictions, the fitted curve is increasing and concave while the variance is increasing over time. The estimated curve and the point-wise confidence intervals in panel (I) accurately capture the mean and variation pattern. The residual plot with fitted 95% confidence intervals in panel (II) shows good performance on estimating the heterogeneous pattern of error standard deviation. The Q-Q plot of the standardized residuals in panel (III) shows no alarming deviation from the normality.
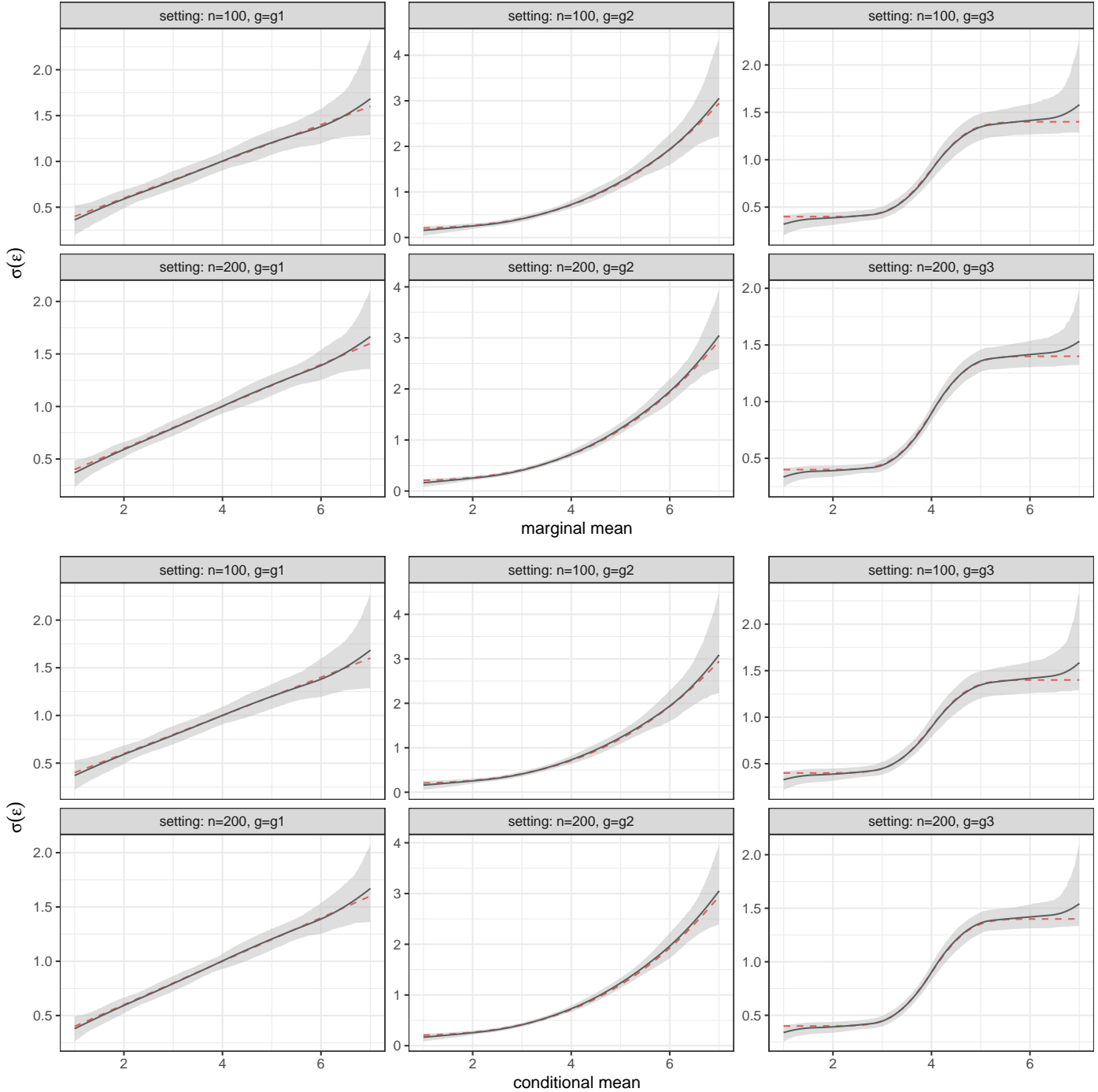
Figure 2: Heteroscedasticity estimation for clustered data when it changes with marginal mean (upper) and conditional mean (lower). The red dashed line is the truth, and the black solid line is the estimation. The grey region is the averaged 95% point-wise confidence interval.

5.2 Chicken Weight

The chicken weight data, which is available in R package **datasets**, is a classic example of clustered data for GCA [14, p. 4]. It contains the body weights in grams of 50 chicks measured at birth date and every other day thereafter until day 20, plus an additional measurement on day 21. These 50 chicks were divided into four groups to have different protein
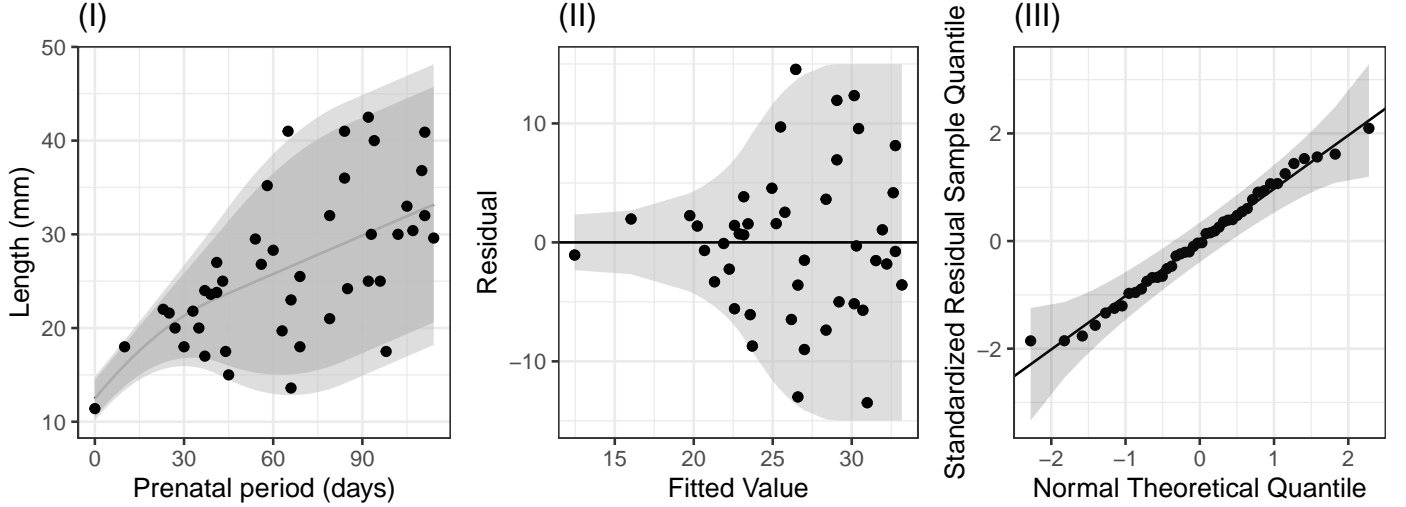
Figure 3: Pancreas length data analysis: (I) original data with estimated mean and 90% and 95% point-wise confidence intervals; (II) the residual versus fitted mean plot with 95% point-wise confidence interval; (III) Q-Q plot of standardized residual.
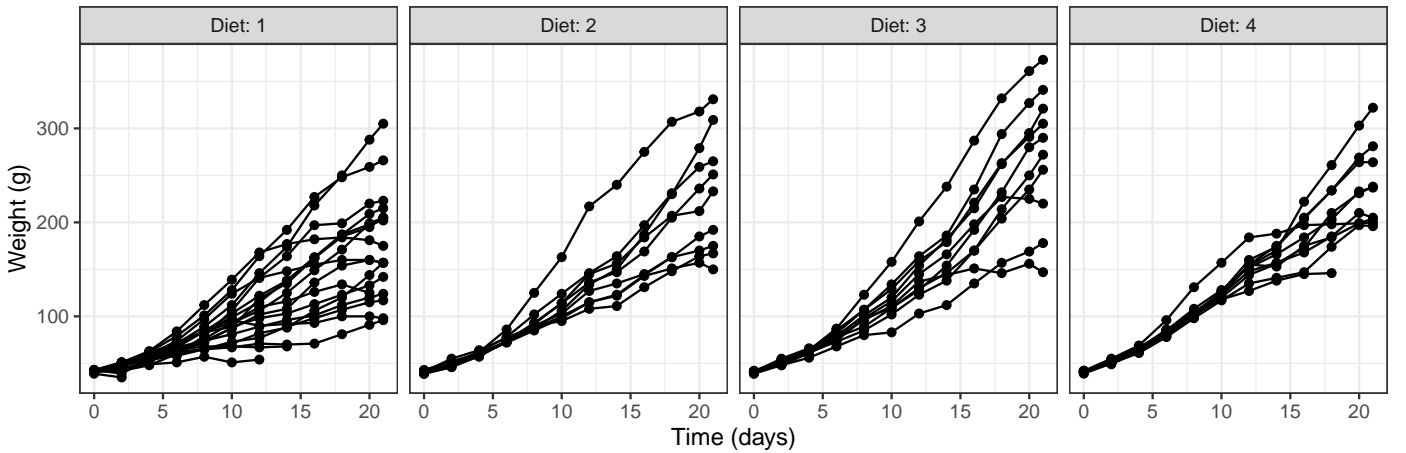
Figure 4: Weights over time in days for chicks in the four diet groups in the chicken weight dataset.

diets, and the scatter plots of weight versus time for each group are shown in Figure 4. We removed Chick No. 24 as an outlier from the original dataset since its weight stopped increasing after day 6. All four groups show increasing trends on both the mean the variation level. Previous studies focused on building regression models on weight gain [14], i.e., the weight change, but no work has been done to directly capture the heteroscedasticity in GCA.

We fitted Model (3.2) to this dataset. The model includes fixed effects consisting of linear I-spline bases of time and their interactions with the diet, as well as a chick-level random effects on the slope of time. Heteroscedasticity is characterized by an I-spline of an index variable, which is either the marginal mean or the conditional mean of the linear effects model. Specifically, the model is

$$y_{ij} = \beta_0 + \sum_{k=1}^{K_1} \left(\beta_k I_{1k}(t_{ij}) + \beta_k^{(2)} I_{1k}(t_{ij}) D_i^{(2)} + \beta_k^{(3)} I_{1k}(t_{ij}) D_i^{(3)} + \beta_k^{(4)} I_{1k}(t_{ij}) D_i^{(4)} \right) + b_i t_{ij} + \varepsilon_{ij},$$
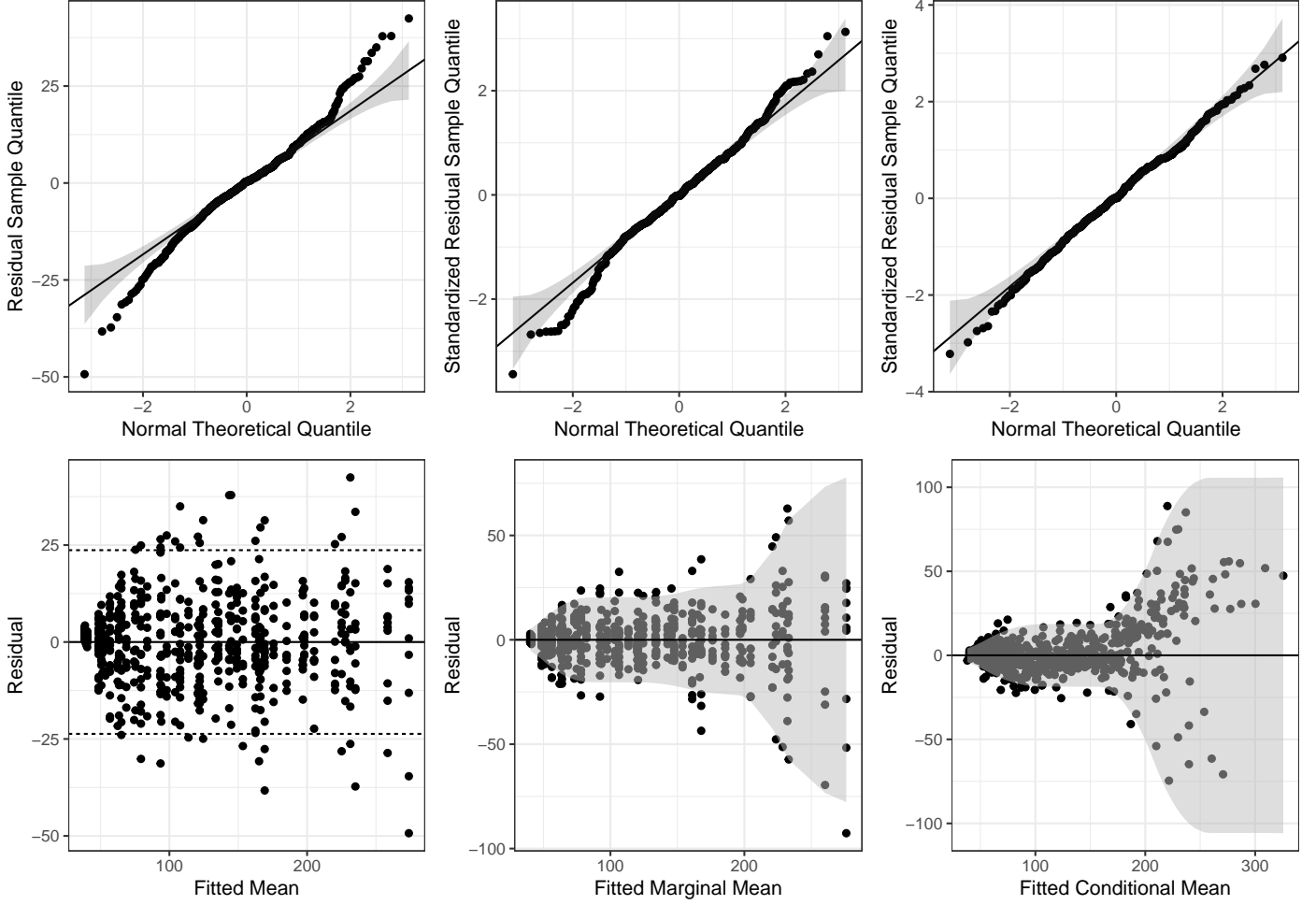$$b_i \sim \mathrm{N}(0, \sigma_b^2),$$

Figure 5: Diagnostics of chicken weight data analysis. Left: naive model; Middle: heteroscedastic model with variance changing with marginal mean; Right: heteroscedastic model with variance changing with conditional mean. Upper: Q-Q plots of standard residual; Lower: residual versus fitted means, with the 95% point-wise confidence intervals.

$$\varepsilon_{ij} \sim \mathrm{N}\left(0, \quad \left(\theta_0 + \sum_{k=1}^{K_2} \theta_k I_{2k}(\nu_{ij})\right)^2\right),$$

where y_{ij} is the weight of chick i at time t_{ij}, $\{I_{1k}(t_{ij}) : k = 1, \ldots, K_1\}$ is a set of I-splines bases with K_1 degrees of freedom used in the mean model, $(D_{ij}^2, D_{ij}^{(3)}, D_{ij}^{(4)})$ are the dummy variables of diet using diet 1 as the reference level, b_i is a normally distributed chick-level random effect with variance σ_b^2, the error term ε_{ij} is normal with mean zero and variance changing with index variable ν_{ij} (either marginal or conditional mean), $\{I_{2k}(\mu_{ij}) : k = 1, \ldots, K_2\}$ is a set of I-spline bases with K_2 degrees of freedom used in the heteroscedasticity model, and the regression coefficients to be estimated are $\{\beta_k, \beta_k^{(2)}, \beta_k^{(3)}, \beta_k^{(4)} : k = 1, \ldots, K_1\}$ and $\{\theta_k : k = 1, \ldots, K_2\}$. The internal knots of spline bases were always chosen to be evenly spaced.

We first need to decide whether to use the marginal mean or the conditional mean as the index variable in the heteroscedasticity model. For both situations, the BIC chose the same number of degrees of freedom of the I-splines. The I-spline bases for the mean pancreas length had degree 0 with 3 degrees of freedom. The I-spline bases for the variance model had degree 1 and 9 degrees of freedom. Since both the marginal mean model and the conditional mean model had the same number of parameters, we can choose the best model by only comparing the log-likelihood. The conditional mean model had a log-likelihood of -2164.899, which is significantly higher than that of the marginal mean model, -2199.183. Both of them were much higher than -2315.220, the log-likelihood of the naive model that did not consider
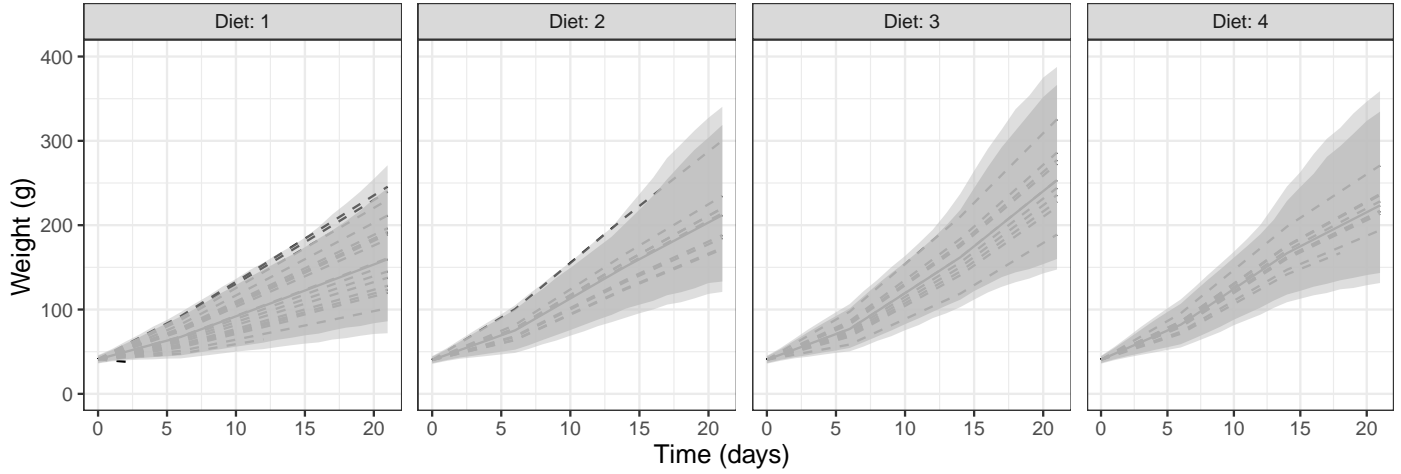
Figure 6: Fitted growing curves for all four Diet groups overlaid with the point-wise 90% and 95% confidence intervals . The solid lines are the fixed effects, and the dashed lines represent the chicks with random effects on Time.

heteroscedasticity.

Figure 5 shows the diagnostic plots for the three models. For the naive model, the residual plot suggests increasing variance as fitted value increases and the Q-Q plot suggest heavier tail than the normal distribution. After considering heteroscedasticity with the proposed method, the model with variance changing with the marginal mean model still has heavy tail problem as seen from the Q-Q plot. The model with variance changing with the conditional mean, however, shows no obvious deviation from the normal distribution in the Q-Q plot of the standardized residuals. These diagnostics are consistent with the model comparison results in terms of log-likelihood.

Point-wise quantiles are of important practical value in GCA since they can be used as reference to check if an individual is in the normal range. For this linear mixed-effects model, we approximated the quantiles by generating 10,000 individuals using the random effects and for each of them, simulating their grow curve using the fitted model. The upper and lower 5% quantiles will give us 90% confidence interval, and the upper and lower 2.5% quantiles will form the 95% confidence interval. Finally, the fitted growth curves for the four diet groups along with their point-wise 90% and 95% confidence intervals from the model with residual variance changing with the conditional mean are shown in Figure 6. Also overlaid are the fitted individual curves for all the chicks in each diet group. It indicates that diet 3 works best for the chicks with fastest growing speed and highest weight at the end of the experiment, while diet 1 is the least favorable.

6. DISCUSSION

Shape-restricted splines provide great flexibility in incorporating prior knowledge about the shapes of the curves to be fitted. With the recently available R package **splines2** [46], such fitting is facilitated in routine data analysis. GCA is an important area where shape restrictions often need to be enforced. In addition to the mean growth level, the heteroscedasticity can also have shape restrictions. Such shape restrictions are enforced through constrained optimizations in an iteratively reweighted fitting procedure, which takes advantage of existing software routines that allow weights. For clustered data, the variance of the error term can be changing with either the marginal mean or the conditional mean. In the latter case, the likelihood is hard to calculate as the marginal distribution of the response vector is no longer within the multivariate normal family. This is not too much of an inconvenience because the iteratively reweighted fitting procedure does need to evaluate this likelihood. It is only needed in calculating model comparison criteria, which only needs to be done once for each fitted model.

Although proposed in the context of GCA, our method is applicable to the general setting of linear mixed-effects models or multi-level models with shape restricted heteroscedasticity. In fact, our simulation studies were done in a general setting. More accurate point and interval estimators are expected when the heteroscedasticity is appropriately accounted for. The parametric bootstrap for inferences also works well in providing valid uncertainty measures for the estimated parameters. Alternative approaches to shape restrictions are possible. For example, isotonic regression [3, 47] can be used to enforce monotonicity, but its implementation and extension to curvature restrictions may not be as simple. Beyond linear models, application of shape restrictions in generalized additive models for location, scale, and shape [31]

or quantile regressions [48] merits further investigation.

REFERENCES

[1] AKAIKE, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In *Selected Papers of Hirotugu Akaike* 199–213 Springer.

[2] ANDREWS, D. W. K. (1999). Estimation When a Parameter is on a Boundary. *Econometrica* **67**(6) 1341–1383. Accessed 2023-05-02.

[3] BARLOW, R. E. and BRUNK, H. D. (1972). The Isotonic Regression Problem and Its Dual. *Journal of the American Statistical Association* **67**(337) 140–147.

[4] BOLKER, B. (2016). *Wald Errors of Variances*. Online; accessed Jan 3, 2022.

[5] CARROLL, R. J. (1982). Adapting for Heteroscedasticity in Linear Models. *The Annals of Statistics* **10**(4) 1224–1233.

[6] CARTER, R. L., RESNICK, M. B., ARIET, M., SHIEH, G. and VONESH, E. F. (1992). A Random Coefficient Growth Curve Analysis of Mental Development in Low-Birth-Weight Infants. *Statistics in Medicine* **11**(2) 243–256.

[7] COLE, T. J. (1988). Fitting Smoothed Centile Curves to Reference Data. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* **151**(3) 385–418.

[8] COLE, T. J. and GREEN, P. J. (1992). Smoothing Reference Centile Curves: The LMS Method and Penalized Likelihood. *Statistics in Medicine* **11**(10) 1305–1319.

[9] CURRAN, P. J., OBEIDAT, K. and LOSARDO, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognition and Development* **11**(2) 121–136.

[10] CURRY, H. B. and SCHOENBERG, I. J. (1966). On Pólya Frequency Functions IV: the Fundamental Spline Functions and Their Limits. *Journal d'Analyse Mathématique* **17**(1) 71–107.

[11] DAS, S. and KRISHEN, A. (1999). Some Bootstrap Methods in Nonlinear Mixed-Effect Models. *Journal of Statistical Planning and Inference* **75**(2) 237–245.

[12] DESDICIOGLU, K., MALAS, M. A. and EVCIL, E. (2010). Foetal Development of the Pancreas. *Folia Morphologica* **69** 216–224.

[13] DIGGLE, P. J. and VERBYLA, A. P. (1998). Nonparametric Estimation of Covariance Structure in Longitudinal Data. *Biometrics* **54**(2) 401–415.

[14] HAND, D. J. and CROWDER, M. J. (1996) *Practical Longitudinal Data Analysis*. Routledge, Boca Raton.

[15] HATA, K., HATA, T. and KITAO, M. (1988). Ultrasonographic Identification and Measurement of the Human Fetal Pancreas in Utero. *International Journal of Gynecology & Obstetrics* **26**(1) 61–64.

[16] HEALY, M. J. R., RASBASH, J. and YANG, M. (1988). Distribution-Free Estimation of Age-Related Centiles. *Annals of Human Biology* **15**(1) 17–22.

[17] HEDEKER, D., MERMELSTEIN, R. and DEMIRTAS, H. (2008). An Application of a Mixed-Effects Location Scale Model for Analysis of Ecological Momentary Assessment (EMA) Data. *Biometrics* **64**(2) 627–34.

[18] HEDEKER, D., MERMELSTEIN, R. J. and DEMIRTAS, H. (2012). Modeling Between-subject and Within-subject Variances in Ecological Momentary Assessment Data Using Mixed-effects Location Scale Models. *Statistics in Medicine* **31**(27) 3328–3336.

[19] HILL, L. M., PETERSON, C., RIVELLO, D., HIXSON, J. and BELFAR, H. L. (1989). Sonographic Detection of the Fetal Pancreas. *Journal of Clinical Ultrasound* **17**(7) 475–479.

[20] KRAKOWIAK-SARNOWSKA, E., FLISIŃSKI, P., SZPINDA, M., SARNOWSKI, J., LISEWSKI, P. and FLISIŃSKI, M. (2005). Morphometry of the Pancreas in Human Foetuses. *Folia Morphologica* **64** 29–32.

[21] LAMBERT, P. C., ABRAMS, K. R., JONES, D. R., HALLIGAN, A. W. F. and SHENNAN, A. (2001). Analysis of Ambulatory Blood Pressure Monitor Data Using a Hierarchical Model Incorporating Restricted Cubic Splines and Heterogeneous Within-Subject Variances. *Statistics in Medicine* **20**(24) 3789–3805.

[22] LIANG, K. -Y. and ZEEGER, S. L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* **73**(1) 13–22.

[23] MACKINNON, J. G. and WHITE, H. (1985). Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics* **29**(3) 305–325. https://doi.org/10.1016/0304-4076(85)90158-7.

[24] MCARDLE, J. J. and NESSELROADE, J. R. (2003). Growth Curve Analysis in Contemporary Psychological Research. *Handbook of Psychology: Research Methods in Psychology* 447–480.

[25] MEYER, M. C. (2008). Inference Using Shape-Restricted Regression Splines. *The Annals of Applied Statistics* **2**(3) 1013–1033.

[26] MULLER, H. -G. and STADTMULLER, U. (1987). Estimation of Heteroscedasticity in Regression Analysis. *The Annals of Statistics* **15**(2) 610–625.

[27] NELDER, J. A. and LEE, Y. (1998). Joint Modeling of Mean and Dispersion. *Technometrics* **40**(2) 168–171.

[28] PINHEIRO, J., BATES, D., DEBROY, S., SARKAR, D. and R CORE TEAM (2022). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-157. https://CRAN.R-project.org/package=nlme.

[29] QUINN, A., BLANCO, C., PEREGO, C., FINZI, G., LA ROSA, S., CAPELLA, C., GUARDADO-MENDOZA, R., CASIRAGHI, F., GASTALDELLI, A., JOHNSON, M., DICK, E. and FOLLI, F. (2012). The Ontogeny of the Endocrine Pancreas in the Fetal/Newborn Baboon. *The Journal of Endocrinology* **214** 289–299. https://doi.org/10.1530/JOE-12-0070.

[30] RAMSAY, J. O. (1988). Monotone Regression Splines in Action. *Statistical Science* **3** 425–461.

[31] RIGBY, R. A. and STASINOPOULOS, D. M. (2005). Generalized Additive Models for Location, Scale and Shape (with Discussion). *Applied Statistics* **54** 507–554.

[32] RIGBY, R. A. and STASINOPOULOS, D. M. (2014). Automatic Smoothing Parameter Selection in GAMLSS With an Application to Centile Estimation. *Statistical Methods in Medical Research* **23**(4) 318–332.

[33] RIGBY, R. A., STASINOPOULOS, D. M. and VOUDOURIS, V. (2013). Discussion: A Comparison of GAMLSS with Quantile Regression. *Statistical Modelling* **13**(4) 335–348.

[34] ROBERT-GRANIÉ, C., HEUDE, B. and FOULLEY, J. -L. (2002). Modelling the Growth Curve of Maine-Anjou Beef Cattle Using Heteroskedastic Random Coefficients Models. *Genetics Selection Evolution* **34**(4) 1–23.

[35] ROBERT-GRANIÉ, C., HEUDE, B. and FOULLEY, J. -L. (2002). Modelling the Growth Curve of Maine-Anjou Beef Cattle Using Heteroskedastic Random Coefficients Models. *Genetics Selection Evolution* **34**(4) 423.

[36] Rossi, R. J. (2018) *Mathematical Statistics: An Introduction to Likelihood Based Inference.* John Wiley & Sons, Hoboken, NJ.

[37] Royston, P. and Wright, E. M. (1998). A Method for Estimating Age-Specific Reference Intervals ('Normal Ranges') Based on Fractional Polynomials and Exponential Transformation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **161**(1) 79–101.

[38] Sandland, R. L. and McGilchrist, C. A. (1979). Stochastic Growth Curve Analysis. *Biometrics* **35**(1) 255–271.

[39] Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics* 461–464.

[40] Smyth, G. K. (1989). Generalized Linear Models with Varying Dispersion. *Journal of the Royal Statistical Society: Series B (Methodological)* **51**(1) 47–60.

[41] Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V. and De Bastiani, F. (2017) *Flexible Regression and Smoothing: Using GAMLSS in R.* CRC Press, Boca Raton, FL.

[42] Stone, A., Shiffman, S., Atienza, A. and Nebeling, L. (2007) *The Science of Real-Time Data Capture: Self-Reports in Health Research.* Oxford University Press, Oxford, England.

[43] Strydhorst, S., Hall, L. and Perrott, L. (2018). Plant Growth Regulators: What Agronomists Need to Know. *Crops & Soils* **51**(6) 22–26.

[44] Thai, H. -T., Mentré, F., Holford, N. H., Veyrat-Follet, C. and Comets, E. (2013). A Comparison of Bootstrap Approaches for Estimating Uncertainty of Parameters in Linear Mixed-Effects Models. *Pharmaceutical Statistics* **12**(3) 129–140.

[45] Wang, T. and Merkle, E. C. (2018). merDeriv: Derivative Computations for Linear Mixed Effects Models with Application to Robust Standard Errors. *Journal of Statistical Software, Code Snippets* **87**(1) 1–16. https://doi.org/10.18637/jss.v087.c01.

[46] Wang, W. and Yan, J. (2021). Shape-Restricted Regression Splines with R Package splines2. *Journal of Data Science* **19**(3) 498–517.

[47] Wang, X. and Li, F. (2008). Isotonic Smoothing Spline Regression. *Journal of Computational and Graphical Statistics* **17**(1) 21–37.

[48] Wei, Y., Pere, A., Koenker, R. and He, X. (2006). Quantile Regression Methods for Reference Growth Charts. *Statistics in Medicine* **25**(8) 1369–1382.

[49] White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* **48**(4) 817–838.

[50] Wu, H. and Zhang, J. -T. (2002). The Study of Long-Term HIV Dynamics Using Semi-Parametric Non-Linear Mixed-Effects Models. *Statistics in Medicine* **21** 3655–75.

[51] Yan, J. and Fine, J. (2004). Estimating Equations for Association Structures (Pkg: P859-880). *Statistics in Medicine* **23**(6) 859–874.

[52] Zee, B. C. (1998). Growth Curve Model Analysis for Quality of Life Data. *Statistics in Medicine* **17**(5–7) 757–766.

[53] Zeileis, A. (2004). Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software* **11**(10) 1–17. https://doi.org/10.18637/jss.v011.i10.

[54] Zeileis, A. (2006). Object-Oriented Computation of Sandwich Estimators. *Journal of Statistical Software* **16**(9) 1–16. https://doi.org/10.18637/jss.v016.i09.

[55] Zeileis, A., Köll, S. and Graham, N. (2020). Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R. *Journal of Statistical Software* **95**(1) 1–36. https://doi.org/10.18637/jss.v095.i01.

[56] Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. and Smith, G. M. (2009) *Mixed Effects Models and Extensions in Ecology With R.* Springer, New York.

Jieying Jiao. Department of Statistics, University of Connecticut, USA.
E-mail address: jieying.jiao@uconn.edu

Wenling Song. Department of Obstetrics, The First Hospital of Jilin University, China.
E-mail address: songwenlingcarol@163.com

Yishu Xue. Department of Statistics, University of Connecticut, USA.
E-mail address: yishu.xue@uconn.edu

Jun Yan. Department of Statistics, University of Connecticut, USA.
E-mail address: jun.yan@uconn.edu