A functional mixed model for scalar on function regression with application to a functional MRI study

Wanying Ma, Luo Xiao*, Bowen Liu,

Department of Statistics, North Carolina State University, 2311 Stinson Drive, Raleigh, NC, 27606, USA
lxiao5@ncsu.edu

Martin A. Lindquist

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, MD, 21205, USA

SUMMARY
Motivated by a functional magnetic resonance imaging (fMRI) study, we propose a new functional mixed model for scalar on function regression. The model extends the standard scalar on function regression for repeated outcomes by incorporating subject-specific random functional effects. Using functional principal component analysis, the new model can be reformulated as a mixed effects model and thus easily fit. A test is also proposed to assess the existence of the subject-specific random functional effects. We evaluate the performance of the model and test via a simulation study, as well as on data from the motivating fMRI study of thermal pain. The data application indicates significant subject-specific effects of the human brain hemodynamics related

*To whom correspondence should be addressed.
to pain and provides insights on how the effects might differ across subjects.

\textit{Key words:} Functional data analysis; Functional principal component; Functional mixed model; Repeated measurements; fMRI; Variance component testing.

1. Introduction

Scalar on function regression models (Ramsay, 2006) are used to relate functional predictors to scalar outcomes and are becoming increasingly popular in statistical applications (e.g., Goldsmith \textit{and others}, 2011; Morris, 2015; Reiss \textit{and others}, 2017). These models have also been extended to data with repeated outcomes (e.g., Goldsmith \textit{and others}, 2012; Gertheiss \textit{and others}, 2013). However, existing models only model the effects of the functional predictor as fixed and do not allow for random functional effects that are either subject- or outcome-specific.

We are motivated by a functional magnetic resonance imaging (fMRI) study of thermal pain (Lindquist, 2012). We begin by briefly describing the study, which was performed on 20 participants. A number of stimuli, consisting of thermal stimulations delivered to the participants left forearm, were applied at two different levels (high and low) to each participant. The temperature of these painful (high) and non-painful (low) stimuli were determined using a pain calibration task performed prior to the experiment. After an 18s time period of thermal stimulation (either high or low), a fixation cross was presented for a 14s time period until the words “How painful?” appeared on the screen. After four seconds of silent contemplation, participants rated the overall pain intensity on a visual analog scale (VAS). The ratings took continuous values and were re-scaled within the range of 100 to 600. The experiment concluded with 10s of rest. During the course of the experimental trial, each subject’s brain activity was also measured using fMRI. Data was extracted from different known pain-responsive brain regions across the brain. Each time course consisted of 23 equidistant measurements made every 2s, providing a total of 46s of
brain activation, ranging from the time of onset of the application of the stimuli to the conclusion of the pain report. The same experiment was conducted multiple times on each participant, with the total number of the repetitions ranging from 39 to 48, thereby giving rise to an unbalanced design. To illustrate the structure of the data, Figure 1 shows the fMRI and the pain rating data for two subjects each with three repetitions, mimicking Figure 1 in Goldsmith and others (2012).

In previous work, Lindquist (2012) used this data set to study how brain activation affected the pain rating using a scalar on function regression model that treated the continuously observed fMRI data as a functional covariate and the subjective rating as a scalar response. However, they used a population model that did not allow for the subject-specific effect of the fMRI imaging on the pain rating to be appropriately modeled. In this work, we seek to determine whether the fMRI data affects the pain rating in a unified or subject-specific manner. For this purpose, we extend the scalar on function linear regression to a new functional mixed effects model for repeated outcomes, and develop a test to determine if the relation between the brain imaging data (more specifically, fRMI data at one brain region) and the pain rating is subject-specific or not. Suggested by the Associate Editor, we further extend the proposed model and test to simultaneously assess the association between pain rating and fMRI data at multiple brain regions.

Testing for the lack of an effect in a functional predictor, i.e., whether the coefficient function is exactly zero, has been well developed in the scalar on function regression literature. For example, Cardot and others (2003) developed a test using the covariance of the scalar response and the functional predictor. Swihart and others (2014) and McLean and others (2015) used the exact likelihood ratio tests of zero variance components (Crainiceanu and Ruppert, 2004). Kong and others (2016) proposed classical Wald, score and F-tests; see also Su and others (2017). However, these tests all focus on fixed functional effects and hence are not applicable to simultaneously testing a collection of random functional effects. Instead of directly testing if multiple random functional effects are all zero, we propose an equivalent test, which tests if the covariance func-
tion of the random functional effects is zero or not. The test can be further formulated as testing
whether multiple variance components are zero. Because existing tests for multiple variance com-
ponents are either computationally intensive or conservative (Qu and others, 2013; Drikvandi
and others, 2012; Baey and others, 2019), we propose an alternative test which is based on the
exact likelihood ratio test of one zero variance component (Crainiceanu and Ruppert, 2004) and
can be more powerful for finite sample data.

The remainder of this paper is organized as follows. In Section 2, we describe our proposed
model along with model estimation and also our test. In Section 3, we extend the proposed model
to deal with a multivariate functional predictor. In Section 4, we assess the numerical performance
of our model and test. In Section 5, we consider the motivating data application. We conclude
the paper with some discussion in Section 6.

2. Method

2.1 Functional mixed model for scalar on function regression with repeated outcomes

We begin by introducing notation. For subject $i$ ($i = 1, 2, \ldots, n = 20$), let $Y_{ij}$ denote the pain
rating at the $j$th repetition with $j = 1, 2, \ldots, n_i$ and $n_i$ denotes the number of repetitions for
subject $i$. Similarly, let $Z_{ij}$ denote the level of stimuli for the $j$th repetition for subject $i$, with
$Z_{ij} = 1$ representing high and $Z_{ij} = 0$ representing low. We shall first consider the fMRI time
series data at one brain region, which corresponds to a univariate functional predictor. In Section
3, we shall extend our model to fMRI data at multiple regions, which corresponds to a multivariate
functional predictor. Let $W_{ijk}$ denote the observed fMRI data at time $t_k = 2k$ seconds ($k = 1, 2, \ldots, K = 23$), which is assumed to be a noisy observation of the smooth functional data
$X_{ij}(t_k)$. Let $T = [0, 46]$ denote the time course of the experiment.

To model the subject-specific random effect of a functional predictor, we propose a new
functional mixed model extending the scalar on function linear regression for repeated outcomes.
The proposed model is

\[ Y_{ij} = \alpha + \alpha_i + Z_{ij}(\gamma + \gamma_i) + \int_{t \in T} \{ \beta(t) + Z_{ij}\delta(t) + \beta_i(t) \} \{ X_{ij}(t) - \mu(t) \} dt + \epsilon_{ij}, \quad (2.1) \]

where \( \alpha \) is the population intercept, \( \alpha_i \) is the subject-specific random intercept, \( \gamma \) is the population effect of the covariate \( Z_{ij} \), \( \gamma_i \) is the subject-specific random effect of \( Z_{ij} \), \( \mu(\cdot) \) is the mean function of the functional predictor \( X_{ij}(t) \), \( \beta(\cdot) \) is the population effect of the functional predictor, \( \delta(\cdot) \) is the interaction effect of the functional predictor and the scalar covariate, \( \beta_i(\cdot) \) is the subject-specific random effect of the functional predictor, and \( \epsilon_{ij} \) are independently and identically distributed (i.i.d.) random errors with distribution \( \mathcal{N}(0, \sigma^2_\epsilon) \). We assume that \( \alpha_i \) are i.i.d. with distribution \( \mathcal{N}(0, \sigma^2_\alpha) \), \( \gamma_i \) are i.i.d. with distribution \( \mathcal{N}(0, \sigma^2_\gamma) \), \( \beta_i(\cdot) \) are i.i.d. random functions following a Gaussian process over \( T \) with mean function \( \mathbb{E}\{\beta_i(t)\} = 0 \) and covariance function \( \text{cov}\{\beta_i(s), \beta_i(t)\} = C(s, t) \), and all random terms are mutually independent across subjects and from each other.

The proposed model is a functional analog to equivalent non-functional multi-subject models commonly used for fMRI data; see, e.g., Lindquist and others (2012). The term \( \delta(\cdot) \) in the model represents the stimuli-specific difference in response, which is typically the parameter of interest in many situations, and the term \( \beta_i(\cdot) \) corresponds to the subject-specific deviation from the population mean, the main interest of this work.

### 2.2 Model for the repeated functional predictor

The repeated functional predictor \( X_{ij}(t) \) might be correlated across repetitions, indexed by \( j \). Following Park and Staicu (2015) and Chen and others (2017), we consider a marginal functional principal component model where the functional predictor is projected onto a sequence of orthonormal marginal eigenfunctions and the associated scores are used to model the correlation
between the repeated functions. Specifically, the model takes the form

\[ W_{ijk} = X_{ij}(t_k) + e_{ijk}, \quad X_{ij}(t) = \mu(t) + \sum_{\ell \geq 1} \xi_{ij\ell} \phi_{\ell}(t), \quad (2.2) \]

where \( e_{ijk} \sim N(0, \sigma^2_e) \) are measurement errors that are independent across \( i, j \) and \( k \) and are independent from the true random functions \( X_{ij} \), \( \xi_{ij\ell} \) are random scores that are independent across \( i \) and \( \ell \) and \( \phi_{\ell}(\cdot) \) are orthonormal marginal eigenfunctions, i.e., \( \int_T \phi_{\ell_1}(t)\phi_{\ell_2}(t)dt = 1_{\ell_1 = \ell_2} \). Here \( 1_{\{\cdot\}} \) is 1 if the statement inside the bracket is true and 0 otherwise. The reason that the functions \( \phi_{\ell}(\cdot) \) are called marginal eigenfunctions and how they can be obtained will be explained soon. The dependence between repeated functions is then modeled via the scores. We use the exchangeable model \( \xi_{ij\ell} = \eta_{i\ell} + \zeta_{ij\ell} \), where \( \eta_{i\ell} \sim N(0, \sigma^2_{\eta\ell}) \) are independent across \( i \) and \( \ell \), and \( \zeta_{ij\ell} \sim N(0, \sigma^2_{\zeta\ell}) \) are independent across \( i, j \) and \( \ell \). The exchangeable model is reasonable for our fMRI data application; however, when the functional predictor is measured repeatedly along a longitudinal time or with a longitudinal covariate \( T_{ij} \), other model specifications such as unspecified or nonparametric covariances as functions of \( T_{ij} \) for the scores might be adopted and the proposed methods in the paper are still applicable. The proposed model is similar to the multi-level fPCA in Di and others (2009) and if \( \sigma^2_{\eta\ell} = 0 \), then the functional data are independent across repetitions. It follows that marginally \( X_{ij} \) are random functions from a Gaussian process with mean function \( E\{X_{ij}(t)\} = \mu(t) \) and covariance function

\[ \text{cov}\{X_{ij}(s), X_{ij}(t)\} = K(s, t) = \sum_{\ell \geq 1} \lambda_{\ell} \phi_{\ell}(s)\phi_{\ell}(t), \quad (2.3) \]

where \( \lambda_{\ell} = \sigma^2_{\phi\ell} + \sigma^2_{\zeta\ell} \). Equation (2.3) shows that \( \phi_{\ell}(\cdot) \) are indeed marginal eigenfunctions and can be obtained via the eigendecomposition of the marginal covariance function \( K(\cdot, \cdot) \).

2.3 Model estimation

The key is to reformulate model (2.1) into a linear mixed effects model using the marginal functional principal component analysis (fPCA) of the functional predictor \( X_{ij} \) described in Sec-
For model identifiability, we assume that the coefficient functions $\beta(\cdot)$ and $\delta(\cdot)$ can be represented as linear combinations of the eigenfunctions $\phi_\ell$ so that $\beta(t) = \sum_{\ell=1}^{\infty} \theta_\ell \phi_\ell(t)$ and $\delta(t) = \sum_{\ell=1}^{\infty} \delta_\ell \phi_\ell(t)$, where $\theta_\ell$ and $\delta_\ell$ are associated scalar coefficients to be determined. Similarly, let $\hat{\beta}_i(\cdot) = \sum_{\ell=1}^{\infty} \theta_{i\ell} \phi_\ell(t)$, where $\theta_{i\ell}$ are independent subject-specific random coefficients with distribution $\mathcal{N}(0, \tau_{i\ell}^2)$. Here the variance components $\tau_{i\ell}^2 \geq 0$ are to be determined as well. Then the induced covariance function $C(s,t)$ of the random functional effects equals $\sum_{\ell=1}^{\infty} \tau_{i\ell}^2 \phi_\ell(s)\phi_\ell(t)$. It follows that model (2.1) can be rewritten as

$$Y_{ij} = \alpha + \alpha_i + Z_{ij}(\gamma + \gamma_i) + \sum_{\ell=1}^{\infty} \xi_{ij\ell}(\theta_\ell + Z_{ij}\delta_\ell + \theta_{i\ell}) + \epsilon_{ij}. \quad (2.4)$$

Model (2.4) has infinitely many parameters and hence cannot be fit, a well known problem for scalar on function regression. Following the standard approach, we truncate the number of eigenfunctions for approximating the functional predictor, so that the associated scores and parameters for $\beta$ and $\beta_i$ are all finite dimensional. Specifically, let $L$ be the number of eigenfunctions to be selected. Then an approximate and identifiable model is given by

$$Y_{ij} = \alpha + \alpha_i + Z_{ij}(\gamma + \gamma_i) + \sum_{\ell=1}^{L} \xi_{ij\ell}(\theta_\ell + Z_{ij}\delta_\ell + \theta_{i\ell}) + \epsilon_{ij}. \quad (2.5)$$

Conditional on the scores $\xi_{ij\ell}$, model (2.5) is a linear mixed effects model and can be easily fit using standard mixed effects model software.

Equation (2.3) suggests that standard fPCA on $X_{ij}$ ignoring the dependence between repeatedly observed functions can be used to estimate the eigenfunctions $\phi_\ell$. Such an approach was proposed in Park and Staicu (2015) and Chen and others (2017). The fPCA on $X_{ij}$ can be conducted using a number of methods, e.g., local polynomial methods (Yao and others, 2005). We use the fast covariance estimation (FACE) method (Xiao and others, 2016), which is based on penalized splines (Eilers and Marx, 1996) and has been implemented in the R function “fpca.face” in the R package refund (Goldsmith and others, 2016). Then, we obtain the estimate of the mean function $\hat{\mu}$, estimates of the eigenfunctions, $\hat{\phi}_\ell$, estimates of the eigenvalues, $\hat{\lambda}_\ell$, and the
estimate of the error variance $\hat{\sigma}_e^2$. We predict the random scores $\xi_{ij\ell}$ using only the observations $\{W_{ij1}, \ldots, W_{ijK}\}$ and denote the prediction by $\hat{\xi}_{ij\ell}$. While the random scores can also be predicted using all observations from the $i$th subject, we have found in the simulations that such an approach may give unstable prediction and hence do not use it.

We select the number of eigenfunctions $L$ by percentage of variance explained (PVE); alternatively one may use AIC on the functional predictor (Li and others, 2013). We use a PVE value of 0.95. Denote the selected number by $\hat{L}$. Then a practical model for (2.5) is

$$Y_{ij} = \alpha + \alpha_i + Z_{ij}(\gamma + \gamma_i) + \sum_{\ell=1}^{\hat{L}} \hat{\xi}_{ij\ell}(\theta_\ell + Z_{ij}\delta_\ell + \theta_{i\ell}) + \epsilon_{ij}. \quad (2.6)$$

Denote the corresponding estimates of $\theta_\ell$ and $\delta_\ell$ by $\hat{\theta}_\ell$ and $\hat{\delta}_\ell$, respectively, and the prediction of $\theta_{i\ell}$ by $\hat{\theta}_{i\ell}$. Then, $\hat{\beta}(t) = \sum_{\ell=1}^{\hat{L}} \hat{\beta}_\ell \hat{\phi}_\ell(t), \hat{\delta}(t) = \sum_{\ell=1}^{\hat{L}} \hat{\delta}_\ell \hat{\phi}_\ell(t)$ and $\hat{\beta}_i(t) = \sum_{\ell=1}^{\hat{L}} \hat{\theta}_{i\ell} \hat{\phi}_\ell(t)$. Confidence bands for $\hat{\beta}(\cdot)$ and $\hat{\delta}(\cdot)$ can also be constructed and the details are omitted.

2.4 Test of random functional effect

Of interest is to assess if the functional effect is subject-specific or the same across subjects. In other words, if $\beta_i(t) = 0$ for all $i$ and $t \in \mathcal{T}$ in model (2.1) or $\beta_i(t) \neq 0$ for some $i$ at some $t \in \mathcal{T}$. Because $\beta_i$ are random coefficient functions, the test can be formulated in terms of its covariance function $C(s,t)$. The null hypothesis is $H_0 : C(s,t) = 0$ for all $(s,t) \in \mathcal{T}^2$ and the alternative hypothesis is $H_a : C(s,t) \neq 0$ for some $(s,t) \in \mathcal{T}^2$. Under $H_0$, $\beta_i(t) = 0$ for all $i$ and $t \in \mathcal{T}$ and model (2.1) reduces to a standard scalar on function linear regression model. Under the truncated model with $L$ functional principal components, $C(s,t) = \sum_{\ell=1}^{L} \tau_\ell^2 \phi_\ell(s)\phi_\ell(t)$, an equivalent test is $H_0' : \tau_\ell^2 = 0$ for all $\ell$ against $H_a' : \tau_\ell^2 > 0$ for at least one $\ell \leq L$. Thus, the test of random functional effect reduces to the test of zeroness of multiple variance components.

Several methods have been proposed for simultaneously testing multiple variance components, e.g., a permutation test (Drikvandi and others, 2012), a score test (Qu and others, 2013), and recently, an asymptotic likelihood ratio test (Baey and others, 2019). The permutation test in
Drikvandi and others (2012) is computationally intensive and the asymptotic LRT (Baey and others, 2019) tends to be conservative in our simulation study. A simple approach is to conduct test of zeroness of each variance component and then use a Bonferroni correction; this test will be referred to as the Bonferroni-corrected test hereafter. Alternatively, following McLean and others (2015) which tested the linearity of a bivariate smooth function, we use the working assumption

$$\tau^2_\ell = \tau^2$$ for all $\ell$, \hspace{1cm} (2.7)

and consider the corresponding test $H_0 : \tau^2 = 0$ against $H_a : \tau^2 \neq 0$. Under $H_0$, $H_0$ still holds. This test involves testing a single variance component and will be referred to as the equal-variance test. While $H_a$ is more general than $H_0$, it was noted in McLean and others (2015) that the equal-variance test could actually outperform the Bonferroni-corrected test even when the true variance components are not the same, i.e., (2.7) does not hold. We shall conduct extensive simulations to compare the performance of the asymptotic LRT, the Bonferroni-corrected test, and the proposed equal-variance test.

The latter two tests involve testing of zeroness of one variance component and we shall use the exact likelihood ratio test (LRT) in Crainiceanu and Ruppert (2004), which is implemented in the R package RLRsim (Scheipl and others, 2008). The advantage of the exact tests is that it is more powerful than asymptotic tests for finite sample data.

A practical issue with the equal-variance test is that standard testing procedures such as the LRT is not directly applicable to model (2.5) because the model has multiple additive random slopes. Therefore, we transform (2.5) into an equivalent mixed effect model under the assumption of (2.7), which has only one random slope term and can therefore easily be tested.

Under assumption (2.7), the random effects and random errors are independent from each other and satisfy the following distributional assumptions:

$$\alpha_i \sim N(0, \sigma^2_\alpha)$$, $$\gamma_{i\ell} \sim N(0, \sigma^2_\gamma)$$, $$\theta_{\ell\ell} \sim N(0, \tau^2)$$, $$\epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$$. \hspace{1cm} (2.8)
The goal of the equivalent model formulation is to convert a set of homoscedastic subject-specific random slopes in (2.5) into a simple random slope, so that the test on homoscedastic random slopes can be conducted using standard software.

Let $Y_i = (Y_{i1}, \ldots, Y_{iL})^T \in \mathbb{R}^{J_i}$, $Z_i = (Z_{i1}, \ldots, Z_{iL})^T \in \mathbb{R}^{J_i}$, $A_i = (\xi_{ij})_{j \in \mathbb{R}^{J_i} \times \mathbb{R}^{J_i} L}$, $B_i = (Z_{ij}^{\xi_{ij}})_{j \in \mathbb{R}^{J_i} \times \mathbb{R}^{J_i} L}$, and $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{iL})^T \in \mathbb{R}^{J_i}$. Also let $\theta = (\theta_1, \ldots, \theta_L)^T \in \mathbb{R}^L$, $\delta = (\delta_1, \ldots, \delta_L)^T \in \mathbb{R}^L$, and $\theta_i = (\theta_{i1}, \ldots, \theta_{iL})^T \in \mathbb{R}^L$. Then model (2.5) can be written in matrix form as follows:

$$Y_i = (\alpha + \alpha_i)1_{J_i} + Z_i(\gamma + \gamma_i) + A_i(\theta + \theta_i) + B_i\delta + \epsilon_i.$$ 

Let $\Delta_i = (1_{J_i}, Z_i, A_i, B_i) \in \mathbb{R}^{J_i \times (2+2L)}$ and $\eta = (\alpha, \gamma, \theta, \delta)^T \in \mathbb{R}^{2+2L}$. It follows that

$$Y_i = \Delta_i\eta + \alpha_i1_{J_i} + \gamma_iZ_i + A_i\theta_i + \epsilon_i. \tag{2.9}$$

Let $J_i = \max(J_i, L)$. Let $\bar{A}_i$ be $A_i$ if $J_i \leq L$ and otherwise $\bar{A}_i = [A_i, 0_{J_i \times (J_i - L)}]$. Then $\bar{A}_i \in \mathbb{R}^{J_i \times J_i}$. Similarly, let $\bar{\theta}_i = \theta_i$ if $J_i \leq L$ and otherwise $\bar{\theta}_i = (\theta_i^T, \nu_i^T)^T$, where $\nu_i \in \mathbb{R}^{J_i - L}$ is multivariate normal with zero mean and covariance $\tau^21_{J_i - L}$ and independent from all other random terms. The vector $\nu_i$ is used only to simplify the algebraic derivation. Then $A_i\theta_i = \bar{A}_i\bar{\theta}_i$ and $\bar{\theta}_i$ are independent and identically distributed multivariate normal with zero mean and covariance $\tau^21_{J_i}$ under the working assumption (2.7). Let $U_i D_i^2 V_i^T$ be the singular value decomposition of $\bar{A}_i$, where $U_i \in \mathbb{R}^{J_i \times J_i}$ and $V_i \in \mathbb{R}^{J_i \times J_i}$ are orthonormal matrices satisfying $U_i^T U_i = I_{J_i}$, $V_i^T V_i = I_{J_i}$, and $D_i = \text{diag}(d_{i1}, \ldots, d_{iL})$ is a diagonal matrix of the singular values of $\bar{A}_i$. Let $\tilde{Y}_i = (\tilde{Y}_{i1}, \ldots, \tilde{Y}_{iL})^T = U_i^T Y_i \in \mathbb{R}^{J_i}$, $\tilde{\theta}_i = (\tilde{\theta}_{i1}, \ldots, \tilde{\theta}_{iL})^T = V_i^T \theta_i$, and $\tilde{\epsilon}_i = (\tilde{\epsilon}_{i1}, \ldots, \tilde{\epsilon}_{iL})^T = U_i^T \epsilon_i$. Then a left multiplication of (2.9) by $U_i^T$ gives

$$\tilde{Y}_i = (U_i^T\Delta_i)\eta + (U_i^T1_{J_i})\alpha_i + (U_i^TZ_i)\gamma_i + D_i^2\tilde{\theta}_i + \tilde{\epsilon}_i,$$

or equivalently,

$$\tilde{Y}_{ij} = (U_{ij}^T\Delta_i)\eta + (U_{ij}^T1_{J_i})\alpha_i + (U_{ij}^TZ_i)\gamma_i + \sqrt{d_{ij}}\tilde{\theta}_{ij} + \tilde{\epsilon}_{ij}. \tag{2.10}$$
where $U_{ij}$ is the $j$th column of $U_i$. The specification (2.8) now becomes $\alpha_i \sim \mathcal{N}(0, \sigma_\alpha^2)$, $\gamma_i \sim \mathcal{N}(0, \sigma_\gamma^2)$, and the random terms are independent across $i$ and $j$, and are independent from each other. Model (2.10) can be fit using a standard mixed model, and then the test of $\tau^2 = 0$ can be conducted by the exact LRT (Crainiceanu and Ruppert, 2004).

3. Extension to multivariate functional predictor

Model (2.1) deals with only fMRI data at one brain region, and it is of interest to consider a model that incorporates fMRI data from multiple regions, i.e., to extend model (2.1) for multivariate functional data. Let $X_{ij}^{(m)}$ denote the $m$th functional predictor for region $m$ ($1 \leq m \leq M$), where $M$ is the number of regions to be modeled together. We extend model (2.1) so that

$$Y_{ij} = \alpha + \alpha_i + Z_{ij} (\gamma + \gamma_i) + \sum_{m=1}^{M} \int_{t \in T} \left\{ \beta_m(t) + Z_{ij} \delta_m(t) + \beta_{im}(t) \right\} \left\{ X_{ij}^{(m)}(t) - \mu_m(t) \right\} dt + \epsilon_{ij},$$

(3.11)

where the terms can be similarly interpreted as before. For the repeated multivariate functional predictor, we extend the decomposition model for repeated univariate functional data (Park and Staicu, 2015; Chen and others, 2017) so that

$$W_{ijk}^{(m)} = X_{ij}^{(m)}(t_k) + e_{ijk}^{(m)}, \quad X_{ij}^{(m)}(t) = \mu_m(t) + \sum_{\ell \geq 1} \xi_{ijk} \phi_{m\ell}(t),$$

where $\{\phi_{1\ell}(t), \ldots, \phi_{M\ell}(t)\}^T$ are multivariate eigenfunctions that satisfy $\sum_{m=1}^{M} \int_{T} \phi_{m\ell_1}(t) \phi_{m\ell_2}(t) dt = 1_{\ell_1 = \ell_2}$, $\xi_{ijk}$ are random scores that are modeled using an exchangeable model as in Section 2.3, and $e_{ijk}^{(m)} \sim \mathcal{N}(0, \sigma_{ek}^2)$ are measurement errors that are independent across $i, j, k$ and $m$. It follows that $\{X_{ij}^{(1)}, \ldots, X_{ij}^{(M)}\}^T$ is marginally following a multivariate Gaussian process with mean function $E\{X_{ij}^{(m)}(t)\} = \mu_m(t)$ and covariance function

$$\text{cov}\{X_{ij}^{(m1)}(s), X_{ij}^{(m2)}(t)\} = K_{m1,m2}(s, t) = \sum_{\ell \geq 1} \lambda_{\ell} \phi_{m1\ell}(s) \phi_{m2\ell}(t).$$

(3.12)

By letting $\beta_m(t) = \sum_{\ell=1}^{\infty} \theta_{\ell} \phi_{m\ell}(t)$, $\delta_m(t) = \sum_{\ell=1}^{\infty} \delta_{\ell} \phi_{m\ell}(t)$ and $\beta_{im}(t) = \sum_{\ell=1}^{\infty} \theta_{\ell} \phi_{m\ell}(t)$, model (3.11) reduces to (2.4). Equation (3.12) shows that $\{\phi_{1\ell}(t), \ldots, \phi_{M\ell}(t)\}^T$ are indeed marginal.
multivariate eigenfunctions.

Because of equation (3.12), to estimate the eigenfunctions $\phi_{m\ell}$, we may conduct multivariate fPCA on $\left\{ X_{ij}^{(1)}, \ldots, X_{ij}^{(M)} \right\}$, also ignoring the dependence between repeated multivariate functional data. We have extended the fast covariance estimation method (Xiao and others, 2016) to multivariate functional data and developed the corresponding R function, which gives estimate of the mean functions and the multivariate eigenfunctions. Alternatively, one may use the R package MFPCA which conducts multivariate fPCA for functions defined on different domains (Happ and Greven, 2018). Similar to before, the scores $\xi_{ij\ell}$ are predicted based on the observations at the $j$th visit for the $i$th subject.

4. A Simulation Study

In this section we conduct simulations to illustrate the performance of the proposed functional mixed model and compare the three tests described in Section 2.4 for testing the existence of random subject-specific functional effects. We shall focus on the models with a univariate functional predictor, but a simulation study with multivariate functional predictor is also conducted and the details are reported in Section S.2 of the Supplementary Materials.

4.1 Simulation settings

We let the domain of functional predictors be $T = [0, 1]$. Each simulated data set consists of $I$ subjects, with each subject having $J$ replicates. Specific values of $I$ and $J$ will be given later.

We generate the response $Y_{ij}$ using model (2.5). The model components are specified as $\alpha = 0.5$, $\gamma = 2$, $\alpha_i \sim N(0,1)$, $\gamma_i \sim N(0,1)$, $Z_{ij} \sim N(0,1)$, $\epsilon_{ij} \sim N(0,1)$. The values of $\tau_2^2$ will be specified later. We let $L = 3$, i.e., the functional predictor $X_{ij}$ has three functional principal components. The functional predictor $X_{ij}$ is generated by model (2.2) with $X_{ij}(t) = \sum_{\ell=1}^{L} \xi_{ij\ell} \phi_{\ell}(t)$ and $e_{ijk} \sim N(0, \sigma_e^2)$. Both independent and
correlated functional predictors are considered: (1) independent $X_{ij}(t)$: $\xi_{ijt} \sim N(0, \sigma^2_{\theta_{\ell}})$; (2) correlated $X_{ij}(t)$: $\xi_{ijt} = \eta_{i\ell} + \zeta_{ij\ell}$, where $\xi_{i\ell} \sim N(0, \sigma^2_{\theta_{\ell}})$ are independent across $i$ and $\ell$, and $\zeta_{ij\ell} \sim N(0, \sigma^2_{\zeta_{\ell}})$ are independent across $i, j$ and $\ell$. Here $\sigma^2_{\theta_{\ell}} = \sigma^2_{\zeta_{\ell}} = 0.5^{\ell}, \ell = 1, \ldots, L$, and the eigenfunctions are $\phi_1(t) = \sqrt{2}\sin(2\pi t)$, $\phi_2(t) = \sqrt{2}\cos(4\pi t)$, $\phi_3(t) = \sqrt{2}\sin(4\pi t)$. The noise variance $\sigma^2_e$ is chosen so that the signal to noise ratio in the functional data $r = \sigma^{-2} \int \mathcal{K}(t,t)dt$ equals either 0 or 3. Here $\mathcal{K}(s,t) = \text{cov}\{X_{ij}(s), X_{ij}(t)\}$ is the marginal covariance function. Note that $r = 0$ corresponds to smooth functional data without noises. Finally, the random scores $\theta_{i\ell}$ are generated by $\theta_{i\ell} \sim N(0, \tau^2_{\ell})$ with $\tau^2_{\ell} = 2^{1-\ell} \tau^2, \ell = 1, \ldots, L$. The quantity $\tau^2$ measures the level of variation of random subject-specific functional effect and will be specified later.

Given a fixed $\tau^2$, we simulate data using a factorial design with four factors: the number of subjects $I$, the number of replicates per subject $J$, the signal to noise ratio $r$ in the functional data, and the independent or correlated functional predictor $X_{ij}(t)$. A total of 24 different model conditions are used: $\{(I, J, r) : I \in \{20, 50, 200\}, J \in \{20, 50\}, r \in \{0, 3\}\}$ with functional predictor being either independent or correlated. Under each model condition, 20000 data sets are simulated for significance tests, and 1000 data sets are simulated for evaluating model estimation. For tests of subjects-specific random functional effects in the proposed model, simulated data with $\tau^2 = 0$ is used to evaluate the sizes of the tests, and simulated data with multiple values of $\tau^2$ are used to assess the power of tests. The power of the tests will also be assessed in additional settings for generating the random scores to accommodate some realistic situations, e.g., when the random scores corresponding to one of the eigenfunctions are exactly 0; see Section 4.2 for details. For model estimation, we set $\tau^2$ to be either 0.04 or 0.08.

### 4.2 Results on tests

Table S.1 in Section S.1 of the Supplementary Materials gives the sizes of the asymptotic LRT (denoted asLRT), the Bonferroni-corrected test and the equal-variance test at the 0.05 significance
level. Under various model conditions, the asymptotic LRT gives sizes much smaller than 0.05 and hence can be potentially conservative. The other two tests give sizes very close to the 0.05 level for independent functional predictor and then give slightly inflated sizes for correlated functional predictor. The results confirm the validity of the three tests for testing the proposed hypothesis.

Figure 2 shows the powers of the three tests as a function of $\tau^2$ for correlated functional predictor. All three tests have increased power when the number of subjects or the number of visits per subject increases. Moreover, they all have higher power when using smooth, i.e., noise-free, functional predictors compared with using noisy functional predictors, as is expected. Under all model conditions, the equal-variance test has higher power than the other two, especially when the number of subjects is small. This agrees with the finding in McLean and others (2015), although their settings are different from ours. The asymptotic LRT seems to have the lowest power among the three, showing that it is indeed conservative for finite sample data. It is also interesting to see that increasing the number of visits per subject seems to result in higher power of the tests than instead increasing the number of subjects. Indeed, with 20 subjects and 50 visits per subject, the power curve of the equal-variance test is close to 1 when $\tau^2$ is around 0.05, whereas with 50 subjects and 20 visits per subject, $\tau^2$ has to be 0.06 or larger to reach the same power. The findings remain the same for independent functional predictor and the corresponding power curves are given in Figure S.1 in Section S.1 of the Supplementary Materials.

In the above simulation, we have considered $\tau^2 = 2^{1-\ell} \tau^2$ for $\ell = 1, \ldots, L = 3$. Now we consider two additional scenarios: scenario 1 with $\tau_1^2 = \tau^2/4$, $\tau_2^2 = \tau^2/2$ and $\tau_3^2 = \tau^2$ and scenario 2 with $\tau_1^2 = \tau^2/2$, $\tau_2^2 = 0$ and $\tau_3^2 = \tau^2$. In scenario 1 the random scores for random functional effects have the smallest variation for the eigenfunction associated with the largest eigenvalue for the functional predictor, while in scenario 2 the random scores corresponding to the second eigenfunction are exactly 0. The power curves for univariate functional predictors are presented in Fig. S.7 - S.10 while for multivariate functional predictors they are in Fig. S.11-S.14 of the
Supplementary Materials. The figures show that the equal-variance test remains the best overall, among the three tests.

To summarize, the simulation study on the tests show that all three tests maintain proper size and have good power. The equal-variance test has the highest power and hence is preferred and will be used in the data application.

4.3 Results on estimation

We compare the proposed functional mixed effects model (2.1) (denoted FMM) with the standard scalar on function regression model (denoted FLM), i.e., $\beta_i(t) = 0$ in model (2.1), in terms of both estimation accuracy of the fixed population effects $\beta(t)$ and $\delta(t)$, and out-of-sample prediction accuracy of the response. For the former, we compute mean integrated squared error (MISE) defined as $\int (\beta(t) - \hat{\beta}(t))^2 dt$ for estimating $\beta(t)$, where $\hat{\beta}(t)$ is the estimate of $\beta(t)$ from either model. The MISE can be similarly computed for $\delta(t)$. For prediction, we use mean squared error (MSE). For each subject in the simulated data, we generate 10 new observations in order to evaluate subject-specific prediction accuracy.

Tables S.2 and S.3 in Section S.1 of the Supplementary Materials summarize the results when correlated functional predictor is used. Under each model condition, FMM outperforms FLM with a smaller MSE for predicting the response, and the two methods have comparable performance on estimating the fixed population functional effects with respect to their MISE. Both models have slightly better performance when the functional predictor is smooth without noises. As the sample size increases, both models achieve better performance for fixed effect estimation and response prediction. Increasing $\tau^2$ results in worse prediction result for the response in FLM while slightly deteriorating results for FMM, which indicates the better performance of FMM when there exists strong subject-specific random functional effect of the functional predictor. Results for the independent functional predictor are shown in Tables S.4 and S.5 in Section S.1.
5. Data Application

In this section, we analyze the data from the fMRI study of thermal pain \((n = 20)\) described in the Introduction. Recall, fMRI data were extracted from 21 different pain-responsive brain regions. The regions included the anterior insula (AINS), the dorsal anterior cingulate cortex (dACC), thalamus, parahippocampal gyrus (PHG), inferior frontal gyrus (IFG), occipital gyrus, corpus callosum, and the second somatosensory area (SII). These are all brain regions that are often categorized as belonging to the so-called “pain matrix”, which is a network of regions thought to generate pain from nociception (Petrovic and others, 2002). The time course extracted form each region consisted of 23 equidistant temporal measurements made every 2s, providing a total of 46s of brain activation, ranging from the time of the application of the heat stimuli to the pain report. We applied the proposed functional mixed model to 6 regions of interest (ROIs), which were found to give statistically significant population effects in Lindquist and others (2012); see Table 1 for a list of names of these ROIs.

We first conduct a joint analysis of all 6 ROIs using the functional mixed model. The residual plot in Figure S.4 in Section S.3 of the Supplementary Materials indicates that it is reasonable to assume normality of the random errors. The equal-variance test of zeroness of subject-specific functional effects gives a \(P\)-value of 0.010 (Table 1), hence favoring the proposed functional mixed model over the standard functional linear mixed model. In addition, the in-sample root mean squared estimation error of the responses for the functional mixed model is 63.55, much smaller than 76.35, the estimation error for the standard functional linear model.

Figure 3 plots the estimated subject-specific functional effects \(\beta(t) + \delta(t) + \beta_i(t)\) when the hot stimuli is applied. Overall the plots show highly diverse signals at the beginning of the trial, followed by strong positive signal in the middle of the trial, and slightly weaker signal towards the
end of trial. The delayed peak occurring in the time period immediately following the conclusion of the thermal stimuli (at time 18s) is consistent with the delayed nature of brain hemodynamics, which peaks roughly 6 seconds after peak neuronal activation, and is consistent with timings of other fMRI experiments (Lindquist and others, 2008). Notably, the secondary peak takes place around the time of the pain reporting (38–44s), perhaps signaling a contribution of activity during “pain recall”.

While the joint analysis and test indicate the existence of subject-specific random functional effects when multiple ROIs are considered together, they cannot assess the existence of subject-specific random effects for each individual ROI. Thus, we next carry out a separate analysis of the data using each ROI as a univariate functional predictor; the results are summarized in Table 1. The residual plots in Figure S.5 in Section S.3 of the Supplementary Materials also indicate it is reasonable to assume normality of random errors. Table 1 gives the root MSE (RMSE) of the estimation using FLM and FMM with each ROI. For ROIs right anterior insula (RAIns I) and right thalamus (RThal), FMM has smaller RMSE than FMM. Among the 6 ROIs, the models with ROIs RAIns I and RThal give significant subject-specific random functional effect at the 0.05 significance level. For these two ROIs, Figure 4 displays the estimated subject-specific functional effects $\beta(t) + \delta(t) + \beta_i(t)$ when the hot stimuli is applied.

Finally, we conduct a 2-cluster analysis of the random functional effects to understand how these effects differ. In each panel of Figure 4, the two clusters are denoted by either black solid curves or gray dashed curves. For both ROIs, it appears that subjects mostly differ in the timing of the delayed peak of brain hemodynamics, with one group having peaks around 22s and the other group having a much later peak, e.g., about 24s for ROI RAIns I. In addition, for ROI RThal, one group of subjects (gray curves) seems to have much pronounced delayed peak as well as strong signal during the application of the stimuli; these diverse subject-specific random curves indicate a better fit using the proposed model compared to the fixed population model.
6. Discussion

We proposed a functional mixed model to accommodate random functional effects of a univariate or multivariate functional predictor for scalar on function regression, along with a significance test of the random functional effects. Motivated by a fMRI study, we considered subject-specific random effects to assess if the association of the fMRI data with pain rating are subject-specific. We focused on functional data that are observed on a common grid, but the proposed model may be extended to handle sparse functional data. Indeed, the model estimation in Section 2.3 of the marginal decomposition model for repeated functional data can be adapted, e.g., using the FACE method for sparse functional data (Xiao and others, 2018) or for sparse multivariate functional data (Li and others, 2018). However, the random score prediction method adopted in the paper might not be optimal. Because of the sparsity of data, the predicted random scores will necessarily be shrunk to zero. Thus, it remains to be seen how the proposed model and test will perform for sparse functional data.

In the data application, we treated the fMRI data collected at multiple brain regions as multivariate functional data. One may also treat the data as two-way functional data or matrix-variate data as in Huang and others (2017). An interesting future research direction is to extend the proposed functional mixed model for repeated matrix-variate data.

7. Software

Software in the form of R code, together with a sample input data set and complete documentation is available at the Github website: https://github.com/lxiao5/fmm_sofr.

8. Supplementary Materials

Supplementary Materials containing additional simulation results and plots for the data application are available online at http://biostatistics.oxfordjournals.org.
We gratefully acknowledge the comments and suggestions of the Associate Editor and anonymous referee that led to a much improved paper. Conflict of Interest: None declared.

Funding

Luo Xiao was partially supported by Grant Number R01NS091307 from National Institute of Neurological Disorders and Stroke (NINDS) and Grant Number R56AG064803 from National Institute on Aging (NIA) and Martin A. Lindquist was partially supported by Grant Numbers R01EB016061 and R01EB026549 from National Institute of Health. This work represents the opinions of the researchers and not necessarily that of the granting organizations.

References


REFERENCES


REFERENCES


REFERENCES


XIAO, LUO, ZIPUNNIKOV, VADIM, RUPPERT, DAVID AND CRAINICEANU, CIPRIAN. (2016). Fast

Fig. 1. Data from the fMRI study: (a) and (c) give the fMRI time series at ROI LAIns for two subjects each with three repetitions; (b) presents the corresponding spaghetti plots of pain ratings.
Fig. 2. Power of three tests at the 5% level for the correlated functional predictor $X_{ij}(t)$ and as a function of $\tau^2$. Black lines are for smooth functional data, i.e., $r = 0$ while gray lines are for noisy functional data. Solid lines: equal-variance test; dashed lines: Bonferroni-corrected test; dot-dashed lines: asLRT.
Fig. 3. Estimated functional effect with a joint analysis of 6 ROIs using FMM. The black solid line is the population functional effect $\beta(t) + \delta(t)$ when the hot stimuli is applied; the gray dashed curves are the subject-specific random functional effect $\beta(t) + \delta(t) + \beta_i(t)$. 
Table 1. Results of FMM and FLM for separate and joint analysis of the fMRI data.

<table>
<thead>
<tr>
<th>ROI</th>
<th>Separate Analysis</th>
<th>P-value</th>
<th>RMSE</th>
<th>FLM</th>
<th>FMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAIns</td>
<td>0.447</td>
<td>81.31</td>
<td>80.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAIns_I</td>
<td>0.035</td>
<td>81.83</td>
<td>79.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAIns_II</td>
<td>0.456</td>
<td>81.19</td>
<td>81.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RThal</td>
<td>0.001</td>
<td>81.13</td>
<td>78.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSII_ParOperc</td>
<td>0.404</td>
<td>81.34</td>
<td>81.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dACC</td>
<td>0.082</td>
<td>81.29</td>
<td>79.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joint analysis</td>
<td>0.010</td>
<td>76.35</td>
<td>63.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Estimated functional effect with a separate analysis of ROI RAIns_I and ROI RThal using FMM. The thick solid line is the population functional effect $\beta(t) + \delta(t)$ when the hot stimuli is applied; the black solid curves are $\beta(t) + \delta(t) + \beta_i(t)$ for one cluster of the subjects and the gray dashed curves are another cluster.