Received: 2 July 2023 | Revised: 15 November 2023 | Accepted: 15 December 2023

DOI: 10.1002/sim.10007

RESEARCH ARTICLE

Statistics in Medicine WILEY

Classified functional mixed effects model prediction

Xiaoyan Liu | **Jiming Jiang**

Statistics Department, University of California, Davis, California, USA

Correspondence
Xiaoyan Liu, Statistics Department, University of California, Davis, CA, USA.
Email: liux0116@gmail.com

Funding information
National Science Foundation, Grant/Award Numbers: DMS-1510219, DMS-1713120

In nowadays biomedical research, there has been a growing demand for making accurate prediction at subject levels. In many of these situations, data are collected as longitudinal curves and display distinct individual characteristics. Thus, prediction mechanisms accommodated with functional mixed effects models (FMEM) are useful. In this paper, we developed a classified functional mixed model prediction (CFMMP) method, which adapts classified mixed model prediction (CMMP) to the framework of FMEM. Performance of CFMMP against functional regression prediction based on simulation studies and the consistency property of CFMMP estimators are explored. Real-world applications of CFMMP are illustrated using real world examples including data from the hormone research menstrual cycles and the diffusion tensor imaging.

KEYWORDS
classification, CMMP, functional mixed effects model, mean squared prediction error

1 | INTRODUCTION

1.1 | Motivation

Developing technology has nowadays allowed collections of various types of functional data in increasing numbers of scientific studies, where observations bear two typical features: (1) data are densely measured over numerous grid points, and (2) the observational units of interest are curves with some between-curve variations at specific grid points.[1,2] In biomedical research, it has been quite common to observe biomarker profiles of individuals as curves measured over time, and it is natural to model each curve as a function of time, especially when there are complex variations in the longitudinal pattern. A typical example is the study of the hormone profiles of healthy women during their menstrual cycles.[3] Figure S1 in Appendix in the Supplemental Information shows the hormone release profiles of 91 cycles collected from 51 subjects, grouped by conceptive status, with each curve corresponding to a unique cycle. The profiles show generally nonlinear trend over time in each group with distinct individual variations among cycles as well as between the groups. Some biomarker profiles feature highly densely observed data with irregular patterns such as peaks or spikes over finely spaced grid points. A common example of such type of profiles is a proteomic spectrometry that consists of spectra data generated from surface-enhanced laser desorption and ionization time-of-flight (SEIDI-TOF) mass spectrometry that measures intensities of a certain biomarker protein over mass-to-charge (M/Z) ratios. It has a broad application in various biomedical research, especially cancer discrimination studies.

With the aforementioned examples, we may consider two motivating challenges associated with functional data analyses in longitudinal studies. One typical question of interest, as we already mentioned, is to study the mean profile and individual deviations from the mean. Traditional parametric models, such as the linear mixed effects model, are not flexible enough to adequately capture the complex nonlinear trend of the curves and also face the problem of curse of dimensions. Hence in such cases functional mixed effects models (henceforth referred to as functional mixed models, or

FMEM) with non-parametric fitting approaches are more appropriate to use. Another challenge is motivated by curve predictions and classifications. With an explosive increase in the quantity and availability of data brought by innovations of information technology in modern scientific research, there has been a growing demand for making accurate predictions at subject levels. As one of the purposes in precision medicine studies, new patients are classified into subgroups of the pre-identified patient population based on random individual characteristics so that effective treatments tailoring down to subgroups of individuals can be developed.[4] With all these being mentioned, prediction mechanisms implemented with FMEMs are important to explore.

In the field of functional data analysis, classification has long been one important topic of interest. Most of the existing relevant work has been focused on classifying a curve of unknown class to a predefined discrete class. Muller,[5] for example, proposed a longitudinal data classification based on generalized linear models which links class identities to functional predictors. Under the framework of FMEM, Zhu[6] discussed classification using Gaussian wavelet-based functional mixed models and robust wavelet mixed models for complex, high-dimensional functional data. Both work talked about classification at a group mean level. Yet not much has been explored in terms of curve matching problems driven by individual characteristics of the curve. Jiang et al.[7] first proposed a classified mixed effects model (CMMP) prediction method that matches a group of new observations based on the mixed effects predicted by a linear mixed-effects model to one of the predefined groups in a given training set. A few extensions based on the original work have been published since then.[8–10]

In this paper, our goal is to develop the classified functional mixed model prediction (CFMMP) method that extends the established classified mixed model prediction (CMMP) method under the framework of linear mixed effects model (LMEM) to the framework of FMEM. We will begin with revisiting the essential concepts of FMEM, followed by literature reviews in selected implementation methods of FMEM. Before we get into the details of the methodology of our proposed CFMMP, we will briefly introduce CMMP and some of its extensions. The paper will cover the theoretical properties of the proposed CFMMP, and simulation studies to evaluate the performance of CFMMP against functional regression prediction (FRP). Two real-world data sets will be used to illustrate the application of CFMMP.

1.2 | Functional mixed effects models

1.2.1 | Cubic polynomial spline smoothing

There has been extensive research done in the field of functional mixed effects models. Some of the work modeled the functional fixed effects or the functional random effects as parametric functions and some as nonparametric functions for example, Wang.[11] The limitation with this type of implementations is that the subject profiles may not be sufficiently modeled in the case where the fixed effects and the random effects are of more arbitrary forms. Improvements have been made in the flexibility of the fitting through approximating the fixed effects with nonparametric functions and the subject-level curves with select smoothing methods[12] or both fixed and random effects with smoothing splines.[3] However, these methods modeled each fixed effect and random effect as a fixed function, which may not sufficiently model the random effects and tends to underestimate of the between-curve variations. Guo[13] adapted Brumback[3] method to a more general framework that allowed more flexible design matrices for the functional fixed and random effects, and was able to yield relatively straight-forward inferences on the estimated functions. According to Guo,[13] given a set of observations $\boldsymbol{Y} = \{Y_i(\boldsymbol{t}_i)\}_{i=1}^{n}$ where $Y_i(\boldsymbol{t}_i) = (Y_i(t_1), Y_i(t_2), \ldots, Y_i(t_{ni}))'$ for $i = 1, \ldots, n$, with $Y_i(t_{ij})$ $(j = 1, \ldots, n_i)$ denoting the observation of the ith curve observed at grid point t_{ij}, the general form of functional mixed effects model is given by

$$Y_i(\boldsymbol{t}_i) = X_i\boldsymbol{\beta}(\boldsymbol{t}_i) + Z_i\boldsymbol{\alpha}_i(\boldsymbol{t}_i) + \boldsymbol{e}_i, \quad i = 1, \ldots, n, \tag{1}$$

where $X_i = [x_{i1}, x_{i2}, \ldots, x_{ip}] \otimes I_{n_i}$ is the design matrix of the functional fixed effects and is invariant of $\boldsymbol{t}$ for a given i. $Z_i = [z_{i1}, z_{i2}, \ldots, z_{iq}] \otimes I_{n_i}$ is the design matrix of the functional random effects. $\boldsymbol{e}_i = \{e_{ij}\}_{j=1}^{n_i}$ are the random errors following $N(\boldsymbol{0}, \sigma_e^2 I_{n_i})$. $\boldsymbol{\beta}(\boldsymbol{t}_i)$ $(n_i p \times 1)$ is the functional fixed effect reflecting the mean profiles and $\boldsymbol{\alpha}_i(\boldsymbol{t}_i) = (\alpha'_{i1}(\boldsymbol{t}_i), \ldots, \alpha'_{iq}(\boldsymbol{t}_i))'$ is the $n_i q \times 1$ functional random effects capturing the individual deviations from the mean profiles. The functional random effect is a collection of a series Gaussian process $(\alpha_{i1}(\boldsymbol{t}_i), \ldots, \alpha_{iq}(\boldsymbol{t}_i))'$ and

$$\boldsymbol{\alpha}_i(\boldsymbol{t}_i) \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_i) \tag{2}$$

with

$$\mathbf{\Sigma}_i = G \otimes R_i, \tag{3}$$

where G is the variance-covariance component matrix with the (i,j)th element being σ_i^2 when $i = j$, and σ_{ij} when $i \neq j$, for $i, j = 1, \ldots, q$. R_i is the correlation matrix. When there are no missing observations in the training data, all curves are observed over the same grid points $T = \{t_1, \ldots, t_m\}$.

Two estimation procedures were proposed by Guo[13] with respect to the general functional form (1). The first method relates the functional form of the fixed and the random effects to cubic smoothing splines. Specifically, each component of the functional fixed and the random effects evaluated at a single grid point t can be modeled as a Gaussian stochastic process using Wahba's Bayesian approach.[14] The implementation method based on this approach is highly restricted by dimension of the data as the algorithm involves inversion of high-dimensional covariance-variance matrix. Therefore, Guo proposed a more efficient estimation method based on Kalman-filtering by expressing model (1) in the form of a single series of multivariate state vectors.[13] The estimation procedures are then achieved by Kalman filtering that includes a forward filtering step followed by a backward smoothing. Marginal likelihood is used for parameter estimation. Compared with the cubic smoothing-spline based approach, the multivariate state-space-model (SSM) approach reduced the operation time from $O((mn)^3)$ to $O(mn^3)$[10,11] to model n curves over m grid points. Hypothesis testing and inferences of the functional effects were also discussed by Guo.[13]

1.2.2 | Wavelet-based functional mixed models

Though flexible in modeling the functional curves, smoothing-spline based FMEM assumes global smoothing properties for all functional effects and hence does not work well with functional data that are highly densely observed and featuring irregular local patterns such as spikes or peaks. Morris[1] developed a wavelet-regression-based functional mixed model that transforms an FMEM to a linear mixed effects model via discrete wavelet transformation (DWT). The DWT procedure decomposes the functional fixed and random effects into the sum of the products between a set of coefficients and a set of orthonormal wavelet bases. The estimation procedure involves a Bayesian wavelet-based approach that applies non-linear shrinkage to fixed effects and uses separate variance components for different functional random effects, thus allowing more generalized estimation of the functional fixed effects as well as the correlation structures of the functional random effects. Markov chain Monte Carlo (MCMC) methods are used to obtain the posterior samples which is then projected back to the original data by inverse discrete wavelet transformation to get the approximated form of the posterior mean of the functional fixed and random effects. The parameters are then estimated through marginalized likelihood functions. Some of the extended work of DWT in FMEM includes the more general estimation and inference method as well as the pertinent hypothesis testing procedure proposed by Antoniadisa and Sapatinas[15] for the parameters in the profile and restricted profile log-likelihood estimation based on the wavelet-based functional mixed models proposed by Morris.[1]

1.2.3 | Multilevel functional principle component analysis

There have been also efforts made to analyze more complexed functional mixed effects models. Di et al[16] proposed a functional principle component analysis (FPCA) framework to estimate a two-way ANOVA functional model (aka the "hierarchical functional model") that involves nested functional random effects. In their proposed framework, the functional fixed effects and random effects are first decomposed into a functional form of mean and covariance through Karhunen–Loève (KL) expansion. The mean and the covariance function are then estimated through method of moments, with parameter estimates of the covariance function being obtained through eigenanalysis. The principle component scores are estimated through MCMC or the best linear unbiased prediction.

1.2.4 | Multivariate functional mixed model

Though the FPCA method can be used to capture the heterogeneity of patterns in the outcomes as well as correlations among them, it does not explicitly model the correlations as is done in parametric models. Li et al[17] proposed a

multivariate functional mixed model framework that models outcomes through subject-specific deviations and subject- and outcome-specific deviations from the mean. They proposed a joint survival and multivariate latent process model by incorporating the MFMM with disease progression time-to-event data. This joint model models multiple longitudinal outcomes as multivariate sparse functional data. It featured reduced covariance functions compared to the multivariate FPCA method and separates the shared latent process from the outcome-specific latent processes, thus increasing the feasibility and interpretability of the model. According to Li et al, the mean and covariance functions were estimated through smoothing splines with the parameters estimated via the Monte Carlo EM algorithm.

2 | CLASSIFIED FUNCTIONAL MIXED EFFECTS MODEL PREDICTION

Our proposed method adapted the CMMP to the framework of FMEM through applying the classification to curves based on the functional mixed effects prediction, thus allowing a more flexible classification paradigm with potentially improved accuracy. In this section, we will first derive the algorithm of CFMMP, assuming that there is an exact match between the training set and the new observations, that is, the new subject indeed belongs to one group of the training set (matched scenario). In practice, there may not exist an exact match between the new group and the training data, whereas an approximate match always exists. As shown by Jiang et al,[7] even approximate match can help in improving prediction accuracy. This will be demonstrated in our simulation studies in Section 4.

2.1 | Prediction of functional mixed effects

Assuming there is a set of training curves $\{y_i(\boldsymbol{t}_i)\}_{i=1}^{n} = \{y_i(t_1), \ldots, y_i(t_{n_i})\}_{i=1}^{n}$ satisfying model (1), with each training curve representing a unique group. Our goal is to match a group of test data $\{y_{\text{new},j}(\boldsymbol{t}_j)\}_{j=1}^{n_{\text{new}}} = \{y_{\text{new},j}(t_1), \ldots, y_{\text{new},j}(t_{m_j})\}_{j=1}^{n_{\text{new}}}$ to one of the training groups. The training curves and the test curves are all observed over a fixed interval $\boldsymbol{t}$. Under the matched scenario, the test group also satisfies model (1) and is assumed to belong to training group I with $I \in \{1, \ldots, n\}$, and the value of I is unknown. It follows that

$$y_{\text{new},j}(\boldsymbol{t}_j) = X_{\text{new}}\boldsymbol{\beta}(\boldsymbol{t}_j) + Z_{\text{new}}\boldsymbol{\alpha}_I(\boldsymbol{t}_j) + \boldsymbol{e}_{\text{new},j}, \quad j = 1, \ldots, n_{\text{new}}, \tag{4}$$

where $X_{\text{new}}, Z_{\text{new}}$ are known design matrices and are shared by the same group of new curves, and $\boldsymbol{e}_{\text{new},j} = \{\{e_{\text{new},ij}\}_{i=1}^{m_j}\}_{j=1}^{n_{\text{new}}}$ are independently distribution random errors with $E(e_{\text{new},ij}) = 0$ and $\text{Var}(e_{\text{new},ij}) = \sigma^2_{\text{new}}$. We assume all curves in the test set are assumed to share the random effects. By taking the average across the common grid points $\boldsymbol{s} \subset T$ of the new curves, we get

$$\bar{y}_{\text{new}}(\boldsymbol{s}) = \frac{1}{n_{\text{new}}}\sum_{j=1}^{n_{\text{new}}} y_{\text{new},j}(\boldsymbol{s}), \tag{5}$$

and use this as a naive estimator of the functional mixed effect of the test (new) group as discussed by Jiang et al.[7] It follows that the functional mixed effect of the new curve to be predicted is:

$$\theta(\boldsymbol{s}) = E(\bar{\boldsymbol{y}}_{\text{new}}(\boldsymbol{s})|\boldsymbol{\beta}(\boldsymbol{t}_i), \boldsymbol{\alpha}_I(\boldsymbol{t}_i)) = X_{\text{new}}\boldsymbol{\beta}(\boldsymbol{s}) + Z_{\text{new}}\boldsymbol{\alpha}_I(\boldsymbol{s}). \tag{6}$$

The functional mixed effect of the new set predicted by matching to training group i then is $E(\theta(\boldsymbol{s})|\boldsymbol{y}_i)$. Based on the assumptions of the model (1), it can be shown that

$$\theta(\boldsymbol{s}) \sim N(X_{\text{new}}\beta(\boldsymbol{s}), Z_{\text{new}}\boldsymbol{\Sigma}_s Z'_{\text{new}}),$$

According to the joint distribution of random Gaussian processes,[18] it follows that

$$\begin{pmatrix} \boldsymbol{y}_i \\ \theta(\boldsymbol{s}) \end{pmatrix} \sim \text{N}\left(\begin{pmatrix} X_i\boldsymbol{\beta}(\boldsymbol{t}_i) \\ X_{\text{new}}\boldsymbol{\beta}(\boldsymbol{s}) \end{pmatrix}, \begin{pmatrix} Z_i\boldsymbol{\Sigma}_{\boldsymbol{t}_i,\boldsymbol{t}_i}Z_i' + \sigma_e^2 I_{m_i} & Z_i\boldsymbol{\Sigma}_{\boldsymbol{t}_i,\boldsymbol{s}}Z'_{\text{new}} \\ Z_{\text{new}}\boldsymbol{\Sigma}_{\boldsymbol{s},\boldsymbol{t}_i}Z_i' & Z_{\text{new}}\boldsymbol{\Sigma}_{\boldsymbol{s},\boldsymbol{s}}Z'_{\text{new}} \end{pmatrix}\right),$$

where

$$\boldsymbol{\Sigma}_{\boldsymbol{u},\boldsymbol{v}} = G \otimes R(\boldsymbol{u},\boldsymbol{v}), \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{t_i}, \boldsymbol{s}$$

with $R(\cdot)$ being the reproducing kernel function specified in Section 2.1. For the convenience of computation and theoretical proof, let $\boldsymbol{s}_i$ be the common grid points between the target test curve and training group i, and let X_i, Z_i and X_{new}, Z_{new} be the corresponding design matrices for the training set and the test curve, respectively. Applying this representation to corresponding parts in (6), we can show that, by Gaussian theory of Gaussian process regression:

$$E(\theta(\boldsymbol{s_i})|y_i(\boldsymbol{t_i})) = X_{\text{new}}\boldsymbol{\beta}(\boldsymbol{s}_i) + Z_{\text{new}}\boldsymbol{\Sigma}_{\boldsymbol{s}_i,\boldsymbol{s}_i} Z_i'(Z_i\boldsymbol{\Sigma}_{\boldsymbol{s}_i,\boldsymbol{s}_i}Z_i' + \sigma_e^2 I_{m_i})^{-1}(\boldsymbol{y}_i(\boldsymbol{s}_i) - X_i\boldsymbol{\beta}(\boldsymbol{s}_i)). \tag{7}$$

It is easy to show that, again by Gaussian theory, the second part of the summation form on the right side of (7) is $Z_{\text{new}}E(\boldsymbol{\alpha}_i(\boldsymbol{s}_i)|\boldsymbol{y_i}(\boldsymbol{s}_i))$, with $E(\boldsymbol{\alpha}_i(\boldsymbol{s}_i)|\boldsymbol{y_i}(\boldsymbol{s}_i))$ being the best predictor (BP) of $\boldsymbol{\alpha}_i(\boldsymbol{s}_i)$ under FMEM (1). The right side of (7) is the BP of $\theta(\boldsymbol{s}_i)$. By replacing the unknown parameters with their corresponding consistent estimators by using an appropriate model fitting method in Section 2.1, we obtain the empirical BP (EBP) of the functional fixed effect $\tilde{\boldsymbol{\beta}}(\boldsymbol{s}_i)$ and the functional random effect $\tilde{\alpha}_i(\boldsymbol{s}_i)$. Then the EBP of the functional mixed effect, $\tilde{\theta}_{(i)}(\boldsymbol{s}_i)$, of the test curve based on training group i can be expressed as:

$$\tilde{\theta_{(i)}}(\boldsymbol{s}_i) = X_{\text{new}}\tilde{\boldsymbol{\beta}}(\boldsymbol{s}_i) + Z_{\text{new}}\tilde{\boldsymbol{\alpha}}_i(\boldsymbol{s}_i), \tag{8}$$

Next we will show how to estimate I, the group index to which the test data are classified to, using the similar approach proposed for CMMP by Jiang et al.[7]

In order to find $\hat{I}$, the estimator of I, we consider the mean squared prediction error (MSPE) function of the EBP of $\theta(\boldsymbol{s}_i)$, defined as follows:

$$\text{M}(\theta(\cdot), i) = \|\theta(\boldsymbol{s}_i) - \tilde{\theta}_{(i)}(\boldsymbol{s}_i)\|^2, \tag{9}$$

where $\|\cdot\|^2$ is the L^2-norm. By $\bar{y}_{\text{new}}(\boldsymbol{s}_i)$ defined in (5) as a naive estimator of $\theta(\boldsymbol{s}_i)$, $\hat{I}$ is then obtained by minimizing the MSPE of the predicted functional mixed effects over all training groups. That is,

$$\hat{I} = \underset{1\le i\le n}{\text{argmin}} \|\bar{y}_{\text{new}}(\boldsymbol{s}_i) - \tilde{\theta}_{(i)}(\boldsymbol{s}_i)\|^2. \tag{10}$$

By replacing i in (8) with $\hat{I}$ and denoting the common grid points between the new curve and the $\hat{I}$th training group, the resulted classified functional mixed-effects predictor (CFMEP) is $\tilde{\theta}_{\hat{I}}(\boldsymbol{s})$.

2.2 | Theoretical properties of CFMMP under matched scenario

Theorem 1. *Continue with expression* (7)*, assuming Assumptions (i)–(v) in the supplemental material hold, it can be shown that the CFMEP is a consistent estimator, that is,*

$$\|\tilde{\theta}_{(\hat{I})}(\boldsymbol{s}) - \theta(\boldsymbol{s})\|^2 \xrightarrow{\text{P}} 0. \tag{11}$$

Proof of Theorem 1 is given in Appendix in the Supplemental Information. According to the subset argument,[19] if the estimator of a certain parameter of interest derived based on a subset of a sample are consistent, the estimator of the same parameter based on the entire sample is also consistent. Therefore, by showing that the CFMEP of the new group based on a subset of the training data is consistent, we conclude that the CFMEP based on the full training set is also a consistent estimator of the true functional mixed effects of the new group.

2.3 | Prediction of future curves

In addition to curve classification, prediction of future curves is also a motivating question of interest. Assuming there is a group of curves $\{y_{\text{f},f}(\boldsymbol{t}_f)\}_{f=1}^{n_{\text{f}}}$ with $\boldsymbol{t}_f = t_1, \ldots t_{m_f}$ to be predicted that belong to an unknown group I which matches one

of the group (curve) in a training set: $\{y_i(\boldsymbol{t}_i)\}_{i=1}^{n}$. Suppose that $y_{\mathrm{f}}(\boldsymbol{t}_f)$ satisfies (4), that is,

$$\boldsymbol{y}_{\mathrm{f}}(\boldsymbol{t}_f) = \boldsymbol{X}_{\mathrm{f}}\boldsymbol{\beta}(\boldsymbol{t}_f) + \boldsymbol{Z}_{\mathrm{f}}\boldsymbol{\alpha}_I(\boldsymbol{t}_f) + \boldsymbol{e}_{\mathrm{f}}, \tag{12}$$

where $X_{\mathrm{f}}, Z_{\mathrm{f}}$ are the design matrices specified by the future observations. Our interest is to predict $\theta_{\mathrm{f}}(\boldsymbol{t}_f)$, the functional mixed effects of the future curves. Under the proposed model, it follows that $\theta_{\mathrm{f}}(\boldsymbol{t}_f) = X_{\mathrm{f}}\boldsymbol{\beta}(\boldsymbol{t}_f) + Z_{\mathrm{f}}\boldsymbol{\alpha}_I(\boldsymbol{t}_f)$. Based on the model assumptions, the BP of future curve $y_{\mathrm{f}}(\boldsymbol{t}_f)$ is:

$$E(y_{\mathrm{f}}(\boldsymbol{t}_f)|y_1(\boldsymbol{t}_1), \ldots, y_n(\boldsymbol{t}_n)) = E(\theta_{\mathrm{f}}(\boldsymbol{t}_f)|y_1(\boldsymbol{t}_1), \ldots, y_n(\boldsymbol{t}_n)). \tag{13}$$

Similar to Jiang et al,[7] it can be shown that the EBP of the prediction of future curve $y_{\mathrm{f}}(\boldsymbol{t}_f)$ is the same as the CFMEP. By replacing the parameters in (13) with their consistent estimators following the same estimation method of CFMEP, we get the CFMMP of future curve $y_{\mathrm{f}}(\boldsymbol{t}_f)$, $\tilde{\theta}_{(\hat{I})}(\boldsymbol{t}_f)$, where $\hat{I}$ is obtained in the same way as (10).

2.4 | CFMMP under unknown-matched scenario

So far we have been assuming that there is an exact match between the training set and the new curves. Now we discuss the situation where the true matching status is unknown. In this case, we do not know whether there is an exact match between the training groups and the new group, and the truth might be that there is a match (matched scenario) or there is no such a match (no-matched scenario). Therefore, when the matching status is unknown, we implement the CFMMP by considering both matched and no-matched assumptions and compare the CFMEP between the two scenarios. The implementation of CFMMP under the matched situation is covered in Section 2.1. For the no-matched case, that is, $I \notin \{1, \ldots, n\}$, using the same notations in Section 4.1, the covariance between $\alpha_I(\boldsymbol{s}_I)$ and $\alpha_i(\boldsymbol{t}_i)(i = 1, \ldots, n)$ is zero. By Gaussian theory, it follows that $E(\alpha_I(\boldsymbol{s})|\boldsymbol{y}_i(\boldsymbol{s})) = 0$. Hence, the BP of $\theta(\boldsymbol{t})$ under the no-matched scenario is

$$E(\theta(\boldsymbol{s}_I)|\boldsymbol{y}_i(\boldsymbol{s}_i)) = X_{\mathrm{new}}\boldsymbol{\beta}(\boldsymbol{s}_i), \tag{14}$$

and by replacing $\boldsymbol{\beta}(\boldsymbol{s}_i)$ with its consistent estimator $\tilde{\boldsymbol{\beta}}(\boldsymbol{s}_i)$ we obtain the EBP, $\tilde{\theta}_0(\boldsymbol{s}_i) = X_f\tilde{\boldsymbol{\beta}}(\boldsymbol{s}_i)$, the of the CFMEP under the no-matched scenario.

Consider the MSPE function $\mathrm{M}(\theta(\cdot), \tilde{\theta}(\boldsymbol{s}_i)) = \|\theta(\boldsymbol{s}_i) - \hat{\theta}(\boldsymbol{s}_i)\|^2$, where $\hat{\theta}(\boldsymbol{s}_i)$ is the EBP of $\theta(\boldsymbol{s}_i)$ based on training group i. Let $\bar{y}_{\mathrm{new}}(\boldsymbol{s}_i)$, the average observations over grid points $\boldsymbol{s}_i$, be a naive estimator of $\theta(\boldsymbol{s})$, and $\tilde{\theta}_I(\boldsymbol{s}_I)$ be the CFMEP under matched-case assumption, omitting the grid point index i, the CFMEP under the matching-status-unknown scenario, $\tilde{\theta}(\boldsymbol{s})$, is obtained by the minimizer of the empirical MSPE between the matched and no-matched scenarios, that is:

$$\underset{\hat{\theta}(\boldsymbol{s})\in\{\tilde{\theta}_0(\boldsymbol{s}),\tilde{\theta}_{(\hat{I})}(\boldsymbol{s})\}}{\operatorname{argmin}} \quad \mathrm{M}(\bar{y}_{\mathrm{new}}(\boldsymbol{s}), \hat{\theta}(\boldsymbol{s})). \tag{15}$$

2.5 | Theoretical properties of CFMEP under unknown-matched scenario

Let $\tilde{\theta}_0(\boldsymbol{t})$ be the CFMEP of $\theta(\boldsymbol{t})$, the true functional mixed effects of the new curve under unknown-matched scenario. Assuming the following conditions are satisfied:

(i) $\boldsymbol{\alpha}_{\mathrm{new}}(\boldsymbol{t}_j), i = 1, \ldots, n, \boldsymbol{\alpha}(\boldsymbol{t}_i), j = 1, \ldots, n_{\mathrm{new}}$ are independent with the $\boldsymbol{e}_{\mathrm{new},j}, \boldsymbol{e}_i$ and $\boldsymbol{\alpha}_{\mathrm{new}}(\boldsymbol{t}_j)$ is independent with the training data under the no-matched scenario, and $\bar{\boldsymbol{e}}_{\mathrm{new}} = \frac{1}{n_{\mathrm{new}}}\Sigma_{j=1}^{n_{\mathrm{new}}}\boldsymbol{e}_{\mathrm{new},j}$ is independent with $\alpha_i(\boldsymbol{t}_i)$ and $\boldsymbol{e}_i$ for all i's.
(ii) Each component of $\boldsymbol{X}_i$ and $\boldsymbol{X}_{\mathrm{new}}$ is bounded.
(iii) $\hat{\boldsymbol{\beta}}(\boldsymbol{t}_i)$ is a consistent estimator of $\boldsymbol{\beta}(\boldsymbol{t}_i)$.
(iv) $m \to \infty, (\log m)^{2\nu}/n_{\min} \to 0$, where $n_{\min} = \min_{1\le i\le m} n_i$, and $n_{\mathrm{new}} \to \infty$.

It can be shown that the CFMEP under unknown-matched scenario is point-wise consistent, that is,

$$E\{(\tilde{\theta}_0(t_{ij}) - \theta(t_{ij}))^2\} \to 0, \quad \text{for } i = 1, \ldots, n_j, \quad j = 1, \ldots, n_{\mathrm{new}}.$$

As the proof of point-wise consistency of $\tilde{\theta}_{\text{new}}(\boldsymbol{t})$ can be well adapted from the method of the proof of the consistency of CMMP under unknown-matched scenario, we skipped the details of the proof in this paper. For details of the proof please refer to the proof of the consistency under unknown-matched case for CMMP (Theorem 2 in the Supplementary Material by Jiang et al[7]).

3 | PREDICTION BANDS

3.1 | Prediction band of CFMEP of the functional mixed effects

Let $\tilde{\theta}(\boldsymbol{t})$ be a consistent estimator of $\theta(\boldsymbol{t})$. Considering the following expression of MSPE:

$$\begin{aligned}\text{MSPE} &= E(\tilde{\theta}(\boldsymbol{t}) - \theta(\boldsymbol{t}))^2 \\ &= E([E(\tilde{\theta}(\boldsymbol{t}) - \theta(\boldsymbol{t}))]^2|\boldsymbol{y}(\boldsymbol{t})) + E(\text{Var}(\tilde{\theta}(\boldsymbol{t})|\boldsymbol{y}(\boldsymbol{t}))).\end{aligned}$$

Since $\tilde{\theta}(\boldsymbol{t})$ is a consistent estimator, that is, $E(\tilde{\theta}(\boldsymbol{t})) = \theta(\boldsymbol{t})$, it follows that

$$\text{MSPE} = E(\text{Var}(\tilde{\theta}(\boldsymbol{t})|\boldsymbol{y}(\boldsymbol{t}))).$$

Assuming there is a new curve that satisfies (4) and a set of training curves that satisfy (1). Using the same notations as model (1) and model (4), let $\boldsymbol{t}$ be the common grid points shared by training group i and the test curve. By the joint distribution of Gaussian processes $y_i(\boldsymbol{t})$ and $\theta_{\text{new}}(\boldsymbol{t})$, we get the conditional variance

$$\text{Var}(\theta_{\text{new}}(\boldsymbol{t})|y_i(\boldsymbol{t})) = \text{Var}(\alpha_{\text{new}}(\boldsymbol{t})|y_i(\boldsymbol{t})) \tag{16}$$

with

$$\text{Var}(\alpha_{\text{new}}(\boldsymbol{t})|y_i(\boldsymbol{t})) = Z_{\text{new}}\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}}}Z'_{\text{new}} - Z_{\text{new}}\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}},\boldsymbol{t}_i}Z'_i(Z_i\boldsymbol{\Sigma}_{\boldsymbol{t}_i}Z'_i + \sigma_e^2 I_{m_i})^{-1}Z_i\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}},\boldsymbol{t}_i}Z'_{\text{new}}.$$

where $\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{t}_i}$ is the variance matrix of $\alpha_{\text{new}}(\boldsymbol{t})$ and $\alpha_i(\boldsymbol{t})$ respectively. $i\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}},\boldsymbol{t}_i}$ is the variance-covariance matrix between the functional random effect of the new curve and that of the training curve i.

Under the matched scenario, assuming the true matched training group is I, it follows that $\alpha_{\text{new}}(\boldsymbol{t})$ shares the same underlying distribution as $\alpha_I(\boldsymbol{t})$, that is,

$$\text{Var}(\alpha_{\text{new}}(\boldsymbol{t})|y_I(\boldsymbol{t})) = \text{Var}(\alpha_I(\boldsymbol{t})|y_I(\boldsymbol{t})), \tag{17}$$

where I can be estimated by CFMMP.

Under the no-matched scenario, $\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}},\boldsymbol{t}_i} = \boldsymbol{0}$, which gives

$$\text{Var}(\tilde{\alpha}_{\text{new}}(\boldsymbol{t})|y_I(\boldsymbol{t})) = Z_{\text{new}}\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}}}Z'_{\text{new}}. \tag{18}$$

In this case there exists training group $\hat{I}$ identified by CFMMP that yields the closest match (ie, minimizes the estimated MSPE). We then estimate $\boldsymbol{\Sigma}_{\boldsymbol{t}_{\text{new}}}$ by estimating the posterior variance of the functional random effect $\text{Var}(\alpha_{\hat{I}}(\boldsymbol{t})|y_{\hat{I}}(\boldsymbol{t}))$.

Let $\tilde{\theta}_{\hat{I}}(\boldsymbol{t})$ be the CFMEP of $\theta(\boldsymbol{t})$. By replacing the unknown parameters in the posterior variance matrices and σ_e^2 in (16) with their consistent estimators estimated using one of the FMEM implementation methods in Section 1.2 (eg, Kalman filtering through SSMs), we get the estimated posterior variance $\hat{\text{Var}}(\alpha_{\hat{I}}(\boldsymbol{t})|y_{\hat{I}}(\boldsymbol{t}))$. Then the $100 \times (1-\alpha)\%$ confidence band for $\tilde{\theta}(\boldsymbol{t})$ is

$$\tilde{\theta}(\boldsymbol{t}) \pm \Phi^{-1}(1-\alpha/2)\left\{\hat{\text{Var}}(\alpha_{\hat{I}}(\boldsymbol{t})|y_{\hat{I}}(\boldsymbol{t})\right\}^{-\frac{1}{2}}. \tag{19}$$

3.2 | Prediction band of future observations

Consider a group of future curves $\boldsymbol{y}_{\text{f}}(\boldsymbol{t}_f) = (y_{\text{f},1}(\boldsymbol{t}_1), \ldots, y_{\text{f},n_{\text{f}}}(\boldsymbol{t}_{n_{\text{f}}}))'$ that share the same individual profiles and satisfy (12). The future prediction is for the mixed effect $\theta_{\text{f}}(\boldsymbol{t})$ of the new curve $\bar{\boldsymbol{y}}_{\text{f}}(\boldsymbol{t})$ defined similarly as (5). Let $\tilde{\theta}_{\text{f}}(\boldsymbol{t})$ be the CFMEP of

$\bar{\boldsymbol{y}}_{\mathrm{f}}(\boldsymbol{t})$ based on training group $y(\boldsymbol{t})$ satisfying (1). It can be shown that the MSPE of $\tilde{\theta}_{\mathrm{f}}(\boldsymbol{t})$ is

$$\begin{aligned} E(\bar{y}_{\mathrm{f}}(\boldsymbol{t}) - \tilde{\theta}_{\mathrm{f}}(\boldsymbol{t}))^2 &= E(\theta_{\mathrm{f}}(\boldsymbol{t}) + \boldsymbol{e}_{\mathrm{f}}(\boldsymbol{t}) - \tilde{\theta}_{\mathrm{f}}(\boldsymbol{t}))^2 \\ &= \mathrm{Var}(\tilde{\theta}_{\mathrm{f}}(\boldsymbol{t})|\boldsymbol{y}(\boldsymbol{t})) + n_{\mathrm{f}}^{-1}\sigma_e^2\boldsymbol{I}, \end{aligned}$$

With the parameters replaced by their consistent estimators, the $100 \times (1-\alpha)\%$ prediction band for $\bar{y}_{\mathrm{f}}(\boldsymbol{t})$ is

$$\bar{y}_{\mathrm{f}}(\boldsymbol{t}) \pm \Phi^{-1}(1-\alpha/2)\left\{\hat{\mathrm{V}}\mathrm{ar}(\tilde{\theta}_{\mathrm{f}}(\boldsymbol{t})|\boldsymbol{y}(\boldsymbol{t})) + n_{\mathrm{new}}^{-1}\hat{\sigma}_e^2\mathbf{1}\right\}^{-\frac{1}{2}}. \tag{20}$$

4 | SIMULATION STUDY

In this section, we will compare the performance of CFMMP with functional regression prediction (FRP) through finite-sample simulation studies. We consider an FMEM similar to the one proposed by Liu and Guo[20] in Section 4, and illustrate CFMMP for both the matched and no-matched scenarios.

Our training data are generated from the following FMEM:

$$y_i(\boldsymbol{t}_i) = 5x_{i1}sin(2\pi\boldsymbol{t}_i) + 5x_{i2}sin(\pi\boldsymbol{t}_i) + r_i(\boldsymbol{t}_i) + \boldsymbol{e}_i, \tag{21}$$

where $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})'$ for $i = 1, \ldots, n, n_i \in (1, \ldots, m)$, with each i representing one unique group. For a fixed i, x_{i1} is a random draw from a uniform distribution on $[-0.5, 0.5]$ and x_{i2} is assigned as 0 or 1 equally likely. The functional fixed effects are represented as $\beta_1(\boldsymbol{t}_i) = 5sin(2\pi\boldsymbol{t}_i)$ and $\beta_2(\boldsymbol{t}_i) = 5sin(\pi\boldsymbol{t}_i)$. $\boldsymbol{e}_i \sim N(\mathbf{0}, 1)$. $r_i(\boldsymbol{t}_i)$ is the functional random intercept and follows $N(\mathbf{0}, \sigma_\alpha^2\Sigma_{\boldsymbol{t}_i})$ with σ_α^2 being the variance component and $\Sigma_{\boldsymbol{t}_i}$ being the correlation matrix of group i. The (i,j)th component of $\Sigma_{\boldsymbol{t}_i}$ is computed by reproducing kernel functions through direct sum.[20,21] The test curve (group) $y_{\mathrm{new}}(\boldsymbol{t})$ has the same underlying fixed effects generated under the same FMEM as (21). For the matched scenario, the functional random effect of the test group is the same as one of the groups from the training set; under the unknown-matched scenario, the random effect is generated randomly from the same distribution as $r_i(\boldsymbol{t}_i)$ in (21). For all the simulation scenarios covered in this section, we carry out 1000 runs and will assume an unknown-matched scenario for the CFMMP algoirthm, regardless of the actual scenarios adopted for simulating the new group and the training groups. For the purpose of efficient computation, the SSM approach with Kalman filtering discussed in sect. 2.1.2. was used for model representation and parameter estimation. The SSM algorithm has been implemented in the Statistical Analysis System (SAS) software through PROC SSM.[22] The functional fixed and random effects effects are approximated by cubic smoothing splines and auto-regressive model of order 1 (AR(1)). The empirical MSPE is computed to compare the performance of CFMMP and FRP. The simulated data of an example training sample of 30 groups under scenario $\sigma_\alpha^2 = 500$ and $n_i = 20$ (Figure S2) based on the above setups as well as all the simulation results discussed below are included in the Appendix in the Supplemental Information.

4.1 | Performance of CFMMP under matched scenario

First we would like to evaluate the impact of sample size n_{new} of the test group on the performance of CFMMP. We set $n_i = m = 20$, and simulate the test data under the matched scenario. Appendix in the Supplemental Information shows the empirical MSPE of FRP and CFMMP as well as the percentage increase in the MSPE of FRP over CFMMP with $n = 30$ training groups for $n_{\mathrm{new}} = 1, 5, 20$ and $\sigma_\alpha^2 = 90, 500, 900$, respectively. It shows that CFMMP consistently yields a lower MSPE compared with FRP. When the rest of the simulation conditions remain equal, this performance difference increases as the variance components of functional random effect becomes larger, and same trend is observed with increasing sample size of the test group. This is expected as FRP does not consider individual-curve deviations and mainly replies on the mean profile of the training group, which all test curves in the same group share. Therefore, the performance of FRP does not change much with the variance of the functional random effects of the test curves, while the MSPE of CFMMP reduces with increasing sample size of the test group. This trend is enhanced with larger test sample size.

When we increase the number of grid points of $\boldsymbol{t}$ for both the training set and the test curve to $m = 40$, and carry out the simulations under $n_{\mathrm{new}} = 1$ and $n = 30$, the resulted MSPE of FRP has greatly increased with denser

observations (Appendix Table S2). In contrast, the performance of CFMMP is not affected as much, with only slightly increases in the MSPE. The chance that CFMMP identifies a matching under the matched scenario remains consistently at relatively high percentages.

4.2 | Comparing matched and no-matched scenarios

In this part of the simulation study, we compare the performance under the matched and the no-matched scenarios and evaluate how the performance comparison is affected by functional random effects and sample size of the training set. Under the no-matched scenario, our new data were still simulated from model (21) but are independent from the training data. We carry out simulations under the following scenarios for both matched and no-matched scenarios: $n = 10, 30; \sigma_\alpha^2 = 90{,}500, 900$. We also examine the average matching percentage of CFMMP, which is the proportion of times when the new group is finally matched to one of the training groups based on the algorithm by CFMMP. In order to mimic a real-world situation where the number of observations usually differs between subjects, the data points are generated with a random missing probability of 0.2 across all observations.

CFMMP performs better than FRP under both matched and no-matched scenarios and the performance of CFMMP also improves with increasing sample size of the training set (Appendix Table S3). Under the matched scenario, the difference between the MSPE comparing CFMMP and FRP increases as the random effects become more dominant under the matched scenario. The matching percentage is also generally higher under the matched scenario comparing to the no-matched scenario, especially under larger component variance ($\sigma^2 = 500{,}900$), which suggests that our algorithm is effective in capturing the true matching status between the test group and the training set. An interesting thing to notice is that even when there is no exact match to the training set, CFMMP still consistently outperforms FRP. This additional prediction accuracy over FRP can be attributed to the curve-specific deviations captured by the functional random effects that enables the CFMMP to identify a group from the training set with characteristics as close to the test curve as possible.

4.3 | Performance of CFMEP confidence bands

We evaluated the performance of the confidence band of CFMEP under the similar scenarios of Section 4.2 based on the widths and the coverage probabilities (CP) of the confidence bands (CB). We first explored the overall average width and coverage probability across all grid points under matched scenario under the same simulation conditions as Table S1. The overall average coverage probabilities are generally satisfactory ($\geq 85\%$), and show increasing trend with increasing test sample size at a given σ_α^2 (Table S4). The overall average width of the CB remains almost stable with increasing test sample sizes. Comparing across σ_α^2 at a given test sample size, the average CP reduce with increasing σ_α^2. This could be due to that the relatively small increase in the average width is likely to be offset by the larger increase within-curve variations controlled by σ_α^2. is maintained and CP of the CBs show an increasing trend as the variance of the random effects goes larger. overall higher CPs and smaller width are observed in the matched scenario compared with those under the no-matched scenarios (Table S5). The point-wise average width and point-wise average CP yield consistent conclusions comparing the matched and no-matched scenarios (Figures S3 and S4). In generally, the confidence bands of CFMEP display relatively satisfactory and stable performance.

5 | APPLICATION

In this section, we illustrate the CFMMP method on two real-world datasets. We compare the CFMEP with FRP, and discuss about the flexibility of CFMMP with respect to longitudinal curves bearing different characteristics: one with distinct subject-level variations, the other one featuring densely observed data. The same approach of model fitting in Section 4 are used to implement the FMEMs in this section.

5.1 | Menstrual cycle data

5.1.1 | Data description

The first example we use comes from a study of the dynamics of women's hormone release during ovarian cycles.[22] This dataset was analyzed as an example for the application of non-parametric functional data analysis using smoothing splines (eg, Brumback and Rice[3]), and was also used by Selukar to illustrate the SSM fitting an FMEM.[22] The dataset contains daily levels of progesterone (PDG), a critical hormone for reproductive system, collected from 51 healthy women with normal reproductive histories in an artificial insemination clinic research. The daily hormone levels were measured over 22 conceptive and 69 non-conceptive menstrual cycles, and were aligned based on the day of ovulation which was taken to be day 0 and were truncated at day 9 and day 15 before and after day 0 for equal length of cycle duration. This yielded a total of 24 observations for each cycle. The initial research goal of this study was to characterize differences in PDG profiles between the conceptive and non-conceptive cycles. The log-transformed PDG levels show non-linear trends and distinct variations between cycles as well as subjects (Supplemental Information, Appendix Figure S1). It would be of interest to study the similarities of the curve characteristics at subject level using CFMMP, which may provide further implications to biomedical research on individualized therapy.

The test curve is selected from the 51 subjects in a leave-one-out manner. In other words, for each trial of the application, we use one subject as a test group, and apply CFMMP by matching it to one of the groups in the training set composed by the remaining 50 subjects. Since there are no predefined similarities among the subjects in this sample, we assume the unknown-matching scenario for the CFMMP algorithm.

5.1.2 | Model fitting and results

The menstrual data in Figure S1 feature generic group mean profile by conceptive conditions, with random deviations from the mean profile at individual level, and the time points of the menstrual cycle data are not too densely distributed. This is a typical example of longitudinal data that can be treated as functional data that fit into the general form of a FMEM (1) in Section 2. The PDG levels observed over days are averaged across cycles for each patient so that each subject has one PDG curve ($mPDG$). The average progesterone level is then normalized through logarithm transformation ($\log mPDG$) which is the outcome variable of the FMEM. Similar to the cortisol data example discussed in Guo,[13] we assume that the underlying variance-covariance structure of subject-level deviations are different between the conceptive and non-conceptive groups which is modeled as the fixed effect, and set up the following functional mixed models with one functional random intercept by group of conceptive methods:

$$Y_i^{(k)}(\boldsymbol{t}_i) = \boldsymbol{\beta}^{(k)}(\boldsymbol{t}_i) + \boldsymbol{\alpha}_i^{(k)}(\boldsymbol{t}_i) + \boldsymbol{e}_i^{(k)}, \quad i = 1, \ldots, 50; k = 1, 2, \tag{22}$$

where $k = 1, 2$ corresponds to the conceptive group and the non-conceptive group, respectively. $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{im}), m \leq 24$, contain the cycle days observed for curve i. $\boldsymbol{\beta}^{(k)}(\boldsymbol{t}_i)$ describes the functional mean profile for conceptive condition k, $\boldsymbol{\alpha}_i^{(k)}(\boldsymbol{t}_i) \sim N(\boldsymbol{0}, G_k \otimes \Sigma_i)$ with the variance components G_k and correlation matrix Σ_i defined similarly for model (1). $\boldsymbol{e}_i^{(k)} \sim N(\boldsymbol{0}, \sigma_e^2 I)$.

Figure S5 in the Supplemental Information shows the CFMMP results for the non-conceptive test group (Subject 1–29) and conceptive test group (Subject 30–51). The CFMEP of each test curve corresponds well with the group mean profiles, and well captures the observed curves. The CFMMP algorithm under an unknown-matched scenario yielded an average matching percentage of 98.0%. For those test curves at which CFMMP concludes a match, a decent amount of resemblance can be observed between the test curve and the training group the test curve is matched to, and CFMEP captures the test profile much better than FRP. The 95% prediction bands generally suggest consistent prediction accuracy over the observed grid points. There are relatively large standard errors of prediction at the ends of each test curve, compared with the middle part, which may be justified by more missing observations at the beginning and ending of the menstrual cycles. For subjects that display more variations over time, for example, test ID = 23 and test ID = 43, the matched training curve also display relatively large deviations from the mean. The CFMEPs though based on smoothed functions do not show as much variations as the observe average PDG levels of these test subjects, they do reveal larger fluctuations over the grid points where irregular peaks are observed for the mean PDG levels, compared with the predicted curves of other test subjects. Overall, our results suggest a relatively satisfactory prediction performance.

5.2 | Diffusion tensor imaging data

The second example is illustrated with the diffusion tensor imaging (DTI) data that were collected at Johns Hopkins University and the Kennedy-Krieger Institute available in the *refund* package of R software. DTI tractography is a magnetic resonance imaging (MRI) technique that studies the white-matter tract of diseases through measuring the diffusivity of waterin the brain.[23] The data we use to implement CFMMP are based on a subset of the source DTI dataset and contain the fractional anisotropy tract profiles for the corpus callosum (CCA) of 142 subjects including multiple sclerosis (MS) patients and controls. Each subject has the DTI tract data collected up to 93 locations of the brain.

We randomly select 82 subjects to form the training set, with each subject representing one unique group. The test set comprises the rest sample. our goal is to match the DTI tracts of the test groups to the training set through applying CFMMP algorithms. FMEM in (1) is used to fit the baseline DTI tract data, with the mean profile grouped by case of MS or control. One subject-specific functional intercept is used to model the deviation of subject from the mean trend over locations. The SSM approach is used to fit the FMEM and for parameter estimation. Assuming an unknown-matched scenario, CFMMP yields 83.3% matching between the test set and training set. For the majority test subjects, where the CFMMP identifies a match between the test and training curves, the CFMEP yield consistently closer capturing of the observed test curve compared with FRP (Supplemental Information, Appendix Figure S6). The matching percentage in this DTI example is lower than that of the hormone study, which is consistent with the simulation results which suggested that matching percentages increase as random effects take over. The DTI data show relatively small variations at individual level compared to the menstrual cycle data, and therefore a smaller estimated variance component of the functional random effects. Among the 50 test subjects that a match with the training curve was identified, all 18 controls and 32 MS cases were matched correctly to the training set in terms of case status, that is, no mismatch of case and control was identified. This not only showed the effectiveness of CFMMP in capturing the mean profile and its accurateness in making predictions, but also suggested there maybe a distinct difference in the CCA patterns over locations between the MS patients and the controls.

6 | DISCUSSION

We have shown that the CFMEP is a consistent estimator of the functional mixed effects of the new group, and outperforms FRP based on the simulation studies. For the proof of the consistency property of CFMEP under the matched case, we used the subset argument so that we can assume all training groups share the same grid points. To be more consistent with the real world application where the grid points usually differ between training groups, we may consider proving the theorem based on the case where the number of observations of each curve varies with groups. The concepts of future curve prediction and the confidence bands were briefly discussed in this paper, and we used simulations to and applications to evaluate and illustrate the performance of confidence bands of CFMEP. It would be worthwhile exploring the asymptotic properties of the confidence band of CFMEP.

As a prediction and classification tool, it would be worth comparing the CFMEP with other prediction and classification methods in longitudinal data analysis. We have done some preliminary explorations of another possible extension of CMMP, the Bayesian CMMP, through concept development and simulation studies. With artificial intelligence being a trending topic nowadays, comparing the CFMMP with other classification methods in machine learning field will be another interesting extension of our work. In addition to the FRP we compared the performance of CFMMP with in this paper, there are also other closely related models that would be interesting to evaluate against CFMMP, for example, linear mixed models (LMM) with splines. Though it is plausible that the LMM with splines may perform better in terms of model fitting or prediction, however, the application may not be as common as CFMMP, especially in medical studies where the applicable problems of interest for the former are expected to be less common.

The model fitting of the FMEM is a critical part of CFMMP. As indicated by its prediction mechanism, the CFMEP is robust to the methods of fitting FMEM, as the prediction mainly replies on the finally estimated smoothed curves of the functional fixed and random effects. Compared to the CFMEP, the confidence band tends to be impacted more by the implementation approach of the FMEM, which the estimation of variance components of the random effects depends on. It would be an insightful exercise to explore the performance and robustness of CFMMP under more complicated framework of FMEMs, such as the multivariate FMEM and the multilevel hierarchical FMEM with nested random effects,[16,17] which is commonly used in complex disease settings.

The grid points of the examples in the simulation studies and applications are relatively dense. The rate of missing observations in the simulated data in Section 4.2 allows flexibility in the point density of the curves and we did not observe an declining performance of the CFMMP comparing the no-missing scenarios with that under a missing rate of 0.2, and the CFMMP displayed certain robustness to increased density of the grid points. In real world medical research longitudinal data are more often sparsely observed. It would be worth investigating the performance and property of CFMMP for sparse data in further research.

ACKNOWLEDGEMENTS

This research is partially funded by the NSF grants DMS-1510219 and DMS-1713120.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Xiaoyan Liu https://orcid.org/0000-0002-8276-907X

REFERENCES

1. Morris JS, Carroll RJ. Wavelet-based functional mixed models. *J R Stat Soc Series B Stat Methodol*. 2006;68(2):179-199.
2. Ramsay SB. *Functional Data Analysis*. New York, NY: Springer; 2005.
3. Brumback BA, Rice JA. Smoothing spline models for the analysis of nested and crossed samples of curves. *J Am Stat Assoc*. 1998;93(443):961-994.
4. Jameson JL, Longo DL. Precision medicine–personalized, problematic, and promising. *N Engl J Med*. 2015;372(23):2229-2234.
5. Muller HG. Functional modeling and classification of longitudinal data. *Ścand J*. 2005;32:223-240.
6. Zhu H, Brown PJ, Morris JS. Robust classification of functional and quantitative image data using functional mixed models. *Biometrics*. 2012;68(4):1260-1268.
7. Jiang J, Rao JS, Fan J, Nguyen T. Classified mixed model prediction. *J Am Stat Assoc*. 2018;113(521):269-279.
8. Sun H, Luan Y, Jiang J. A new classified mixed model predictor. *J Stat Plan Inference*. 2020;207:45-54.
9. Ma H, Jiang J. Classified generalized linear mixed model prediction incorporating pseudo-prior information. *Can J Stat*. 2023;51(2):580-595.
10. Ma H, Jiang J. Pseudo-Bayesian classified mixed model prediction. *J Am Stat Assoc*. 2023;118(543):1747-1759.
11. Wang Y. Smoothing spline models with correlated random errors. *J Am Stat Assoc*. 1998;93(441):341-348.
12. Zhang D, Lin X, Raz J, Sowers M. Semiparametric stochastic mixed models for longitudinal data. *J Am Stat Assoc*. 1998;93(442):710-719.
13. Guo W. Functional mixed effects models. *Biometrics*. 2002;58:121-181.
14. Wahba G. Improper priors, spline smoothing and the problem of guarding against model errors in regression. *J R Stat Soc*. 1978;40(3):364-372.
15. Antoniadisa A, Sapatinas T. Estimation and inference in functional mixed-effects. *Comput Stat Data*. 2007;51:4793-4813.
16. Di CZ, Crainiceanu CM, Caffo BS, Punjabi NM. Multilevel functional principal component analysis. *Ann Appl Stat*. 2009;3(1):458.
17. Li C, Xiao L, Luo S. Joint model for survival and multivariate sparse functional data with application to a study of Alzheimer's disease. *Biometrics*. 2022;78(2):435-447.
18. Quinonero-Candela J, Rasmussen CE. A unifying view of sparse approximate Gaussian process regression. *J Mach Learn Res*. 2005;6:1939-1959.
19. Jiang J. The subset argument and consistency of MLE in GLMM: answer to an open problem and beyond. *Ann Stat*. 2003;41(1):177-195.
20. Liu Z, Guo W. Fmixed: a SAS macro for smoothing-spline-based functional mixed effects models. *J Stat Softw*. 2011;43:1-13.
21. Gu C. *Smoothing Spline ANOVA Models*. Vol 297. New York: Springer; 2013.
22. Selukar R. Functional modeling of longitudinal data with the SSM procedure. SAS Global Forum Proceedings. 2015 1–19.
23. Goldsmith J, Bobb J, Crainiceanu CM, Caffo B, Reich D. Penalized functional regression. *J Comput Graph Stat*. 2011;20(4):830-851.

SUPPORTING INFORMATION

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

How to cite this article: Liu X, Jiang J. Classified functional mixed effects model prediction. *Statistics in Medicine*. 2024;43(7):1329-1340. doi: 10.1002/sim.10007