Received: 24 September 2021 | Revised: 29 October 2021 | Accepted: 13 November 2021

DOI: 10.1002/sta4.438

ORIGINAL ARTICLE

WILEY

Multilevel varying coefficient spatiotemporal model

Yihao Li[1] | Danh V. Nguyen[2] | Esra Kürüm[3] | Connie M. Rhee[2,4] | Sudipto Banerjee[1] | Damla Şentürk[1]

[1]Department of Biostatistics, University of California, Los Angeles, CA, 90095, USA

[2]Department of Medicine, University of California Irvine, Orange, CA, 92868, USA

[3]Department of Statistics, University of California, Riverside, CA, 92521, USA

[4]Harold Simmons Center for Chronic Disease Research and Epidemiology, University of California Irvine School of Medicine, Orange, CA, 92868, USA

Correspondence
Damla Şentürk, Department of Biostatistics, University of California, Los Angeles, CA 90095, USA.
Email: dsenturk@ucla.edu

Funding information
National Institute of Diabetes and Digestive and Kidney Diseases, Grant/Award Numbers: R01 DK 122767, R01 DK092232

Over 785,000 individuals in the United States have end-stage renal disease (ESRD), with about 70% of patients on dialysis, a life-sustaining treatment. Dialysis patients experience frequent hospitalizations. In order to identify risk factors of hospitalizations, we utilize data from the large national database, United States Renal Data System (USRDS). To account for the hierarchical structure of the data, with longitudinal hospitalization rates nested in dialysis facilities and dialysis facilities nested in geographic regions across the United States, we propose a multilevel varying coefficient spatiotemporal model (M-VCSM) where region- and facility-specific random deviations are modelled through a multilevel Karhunen–Loéve (KL) expansion. The proposed M-VCSM includes time-varying effects of multilevel risk factors at the region- (e.g., urbanicity and area deprivation index) and facility-levels (e.g., patient demographic makeup) and incorporates spatial correlations across regions via a conditional autoregressive (CAR) structure. Efficient estimation and inference are achieved through the fusion of functional principal component analysis (FPCA) and Markov chain Monte Carlo (MCMC). Applications to the USRDS data highlight significant region- and facility-level risk factors of hospitalizations and characterize time periods and spatial locations with elevated hospitalization risk. Finite sample performance of the proposed methodology is studied through simulations.

KEYWORDS
conditional autoregressive model, end-stage renal disease, hospitalization risk, multilevel longitudinal data, United States Renal Data System

1 | INTRODUCTION

As of 2018, the number of patients with end-stage renal disease (ESRD) in the United States exceeded 785,000, with approximately 554,000 (70%) patients undergoing dialysis, a life-sustaining treatment (United States Renal Data System [USRDS], 2020). Patients on dialysis have a high burden of complex comorbid conditions and are typically hospitalized twice a year with hospitalization rates sharply elevated in the first year of dialysis. In addition, ESRD patients remain on dialysis for long periods of time (for the duration of their lives or until receiving a kidney transplant), and their needs may change as they stay on dialysis. Our own works and others (Li et al. 2018, 2020; USRDS, 2020) have also shown significant variation in hospitalizations among dialysis facilities contributing to spatial variation with regional clusters of high rates (e.g., 'hot spots'). More generally, there is a compelling need to more fully understand the region- (e.g., urbanicity and area deprivation index [ADI]) and facility-level factors (e.g., patient demographic makeup or comorbidity burden) that contribute to differences in longitudinal hospitalizations observed across the United States over the time course that patients receive dialysis. Comprehensive modelling of hospitalizations to identify the time-dynamic/time-varying effects of modifiable risk factors can better inform patient care and is essential in efforts aimed at reducing hospitalization in the dialysis population.

The United States Renal System (USRDS) is a large national database which contains data on almost all dialysis patients in the United States, with a hierarchical structure: longitudinal hospitalizations over time nested in dialysis facilities and dialysis facilities nested within regions across the United States. We note that rather than modelling aggregated rates at the region-level (see USRDS, 2020 for a conditional autoregressive [CAR] Bayesian modelling of aggregated rates), it is important to develop a multilevel model for hospitalization rates at the facility-level, incorporating the hierarchical structure of the data, to target a more granular estimate of the effects of multilevel risk factors, while also accounting for the spatiotemporal variation across dialysis facilities. The multilevel approach not only allows for incorporation of multilevel risk factors more naturally, but it also facilitates the study of variation in hospitalizations among facilities within a region. This is important since elevated hospitalization rates within a region may be due to a few outlying facilities with higher rates or consistently high hospitalization rates across a high percentage of the facilities within a region. The multilevel approach provides the ability to study the reasons behind elevated hospitalization rates within a region, whether it be region-level risk factors or variation among facilities within the region, which could also partially be driven by facility-level risk factors.

There is extensive literature in spatiotemporal modelling where data are viewed as time series observed on a lattice of spatial locations (Cressie & Wikle, 2011). The goals of typical spatiotemporal models mainly focus on prediction, either for time points in the future or for unmeasured spatial locations (Quick et al. 2013; Zhang et al. 2016). For the few works on *multilevel* spatiotemporal modelling, a functional data analysis (FDA) approach is taken to model structured functional trajectories where the dependencies are induced by spatial or temporal proximity (Crainiceanu et al. 2009; Di et al. 2009; Kundu et al. 2016; Morris et al. 2003; Morris & Carroll, 2006; Zipunnikov et al. 2011). Functional principal components analysis (FPCA) using the Karhunen–Loéve (KL) representation is utilized as an effective dimension reduction tool in modelling functional variability. For multilevel functional data that are spatially correlated, spatial correlations have typically been modelled across lower level units that are nested within independent subjects (Baladandayuthapani et al. 2008; Hasenstab et al. 2017; Staicu et al. 2010; Scheffler et al. 2020). For multilevel functional data where spatial correlation is at the highest level of the hierarchy (e.g., longitudinal hospitalizations nested in dialysis facilities and facilities nested in spatially correlated geographic regions), Li et al. (2021) considered a multilevel spatiotemporal functional model with a focus on drawing valid multilevel inference accounting for spatiotemporal correlations. However, these models do not include potentially time-varying effects of multilevel covariates.

To study time-varying effects of covariates, varying coefficient models are an effective tool and have been widely used to model longitudinal outcomes (Cleveland et al. 1991; Hastie & Tibshirani, 1993). For single-level spatiotemporal data, Serban (2011) proposed a space–time-varying coefficient model to examine the association of service accessibility and demographics. Zhang et al. (2016) considered a functional CAR approach for modelling spatially correlated genomic changes over areal regions of the bladder tissue (space) and genomic locations (time). However, both models do not include multilevel covariates or account for the hierarchical dependency structure in the data. For two-level longitudinal data, Li et al. (2018, 2020) proposed multilevel varying coefficient models (M-VCM) to study the time-varying effects of facility- and patient-level covariates on hospitalizations of dialysis patients. However, these works do not include modelling of spatiotemporal correlations in the data for valid inference.

Therefore, to quantify the time-varying effects of region- and facility-level risk factors on hospitalization rates in the dialysis cohort while accounting for both the hierarchical dependency structure and the spatiotemporal correlations in the data, we propose a novel multilevel varying coefficient spatiotemporal model (M-VCSM). Multilevel covariates are included in the model to study the time-varying effects of both region- and facility-level risk factors. Additional region- and facility-specific time-varying random deviations model the multilevel spatiotemporal correlation structure of the data. Spatial correlations are induced among the region-specific random deviations through a CAR model to account for dependencies across regions and to stabilize the estimates for small regions. Estimation of time-varying effects of multilevel covariates jointly with multilevel time-varying random deviations (with an additional spatial correlation at the region-level) poses a major computational challenge, especially in big data settings. To achieve computational efficiency in modelling large data sets, such as data from USRDS, the proposed estimation and inference rely on dimension reduction via Bayesian P-splines (Lang & Brezger, 2004) in targeting time-varying coefficient functions of multilevel covariates and FPCA in targeting the multilevel time-varying random deviations. Following dimension reduction, model parameters (including P-spline coefficients and variance components) are targeted via Markov chain Monte Carlo (MCMC) in a mixed effects modelling framework. Section 2 introduces the proposed M-VCSM and outlines the proposed estimation and inference procedures. Applications to USRDS data to model hospitalization risk among 5494 dialysis facilities across 423 health service areas (HSAs), regions with relatively self-contained infrastructure for the provision of hospital care, in the United States, are given in Section 3. Simulations to study the finite sample performance of the proposals, including comparisons with a M-VCM that ignores the spatial correlation in the data, are presented in Section 4, followed by a discussion given in Section 5.

2 | PROPOSED M-VCSM

2.1 | Model specification

Let $Y_{ijk} = Y_{ij}(t_{ijk})$ denote the hospitalization rate of the jth facility from region i at time (month) t_{ijk}, where $i = 1,2,...,n$ denotes regions, $j = 1,2,...,N_i$ denotes dialysis facilities within the ith region and $k = 1,2,...,T$ denotes the kth month after transition to dialysis. The hospitalization rate for the

*j*th facility in month *k* is defined as the ratio of the total number of patient hospitalizations to the total patient follow-up time in month *k*. For ease of interpretation, monthly rates are multiplied by 12 before modelling, to target the hospitalization rate PPY (consistent with USRDS reporting). Note that monthly hospitalization rates after transition to dialysis are targeted as a continuous outcome, similar to previous seminal works of (Quick et al. 2013; Short et al. 2002) and amenable for the use of functional data analysis tools. While the vast majority of yearly hospitalization rates are less than 5 (median 1.8), the range of the yearly rates extends from 0 to 9.5.

The proposed M-VCSM includes P region- ($\boldsymbol{X}_i = (X_{i1},...,X_{iP})^{\mathsf{T}}$) and Q facility-level ($\boldsymbol{Z}_{ij}(t) = (Z_{ij1}(t),...,Z_{ijQ}(t))^{\mathsf{T}}$) covariates as well as time-dynamic region- ($U_i(t)$) and facility-specific ($V_{ij}(t)$) deviations:

$$Y_{ij}(t) = \boldsymbol{X}_i^{\mathsf{T}}\boldsymbol{\beta}(t) + \boldsymbol{Z}_{ij}(t)^{\mathsf{T}}\boldsymbol{\theta}(t) + U_i(t) + V_{ij}(t) + \epsilon_{ij}(t), \tag{1}$$

where $\boldsymbol{\beta}(t) = \{\beta_1(t),...,\beta_P(t)\}^{\mathsf{T}}$ and $\boldsymbol{\theta}(t) = \{\theta_1(t),...,\theta_Q(t)\}^{\mathsf{T}}$ denote the time-varying coefficient functions of the region- and facility-level covariates, respectively, and $\epsilon_{ij}(t)$ denotes the measurement error. Note that in the formulation in Equation (1), the facility-level covariates are time-dependent and region-specific covariates are time-invariant, to mimic our data application. Facility-level covariates, such as patient demographic makeup or comorbidity burden, which characterize the patient cohort, change over time as the patient cohort at the facility changes over time, while region-level covariates, such as urbanicity or ADI, are taken to be time-static over the 2-year follow-up on dialysis. Nevertheless, the formulations below can easily accommodate both time-varying and time-invariant covariates at both the region- and facility-levels, with minor changes to the design matrix.

The time-varying region- and facility-specific deviations, capturing the remaining spatiotemporal variation in hospitalizations after adjusting for region- and facility-level covariates, are both modelled via the KL expansions,

$$U_i(t) = \sum_{\ell=1}^{\infty} \xi_{i\ell}\psi_{\ell}^{(1)}(t),\ V_{ij}(t) = \sum_{m=1}^{\infty} \zeta_{ijm}\psi_{m}^{(2)}(t). \tag{2}$$

where $\psi_{\ell}^{(1)}(t)$ and $\psi_{m}^{(2)}(t)$ denote the region- and facility-level (first- and second-level) orthonormal eigenfunctions (denoted by superscripts $^{(1)}$ and $^{(2)}$) and $\xi_{i\ell}$ and ζ_{ijm} denote respectively region- and the facility-level principal component (PC) scores. Note that while PC scores across levels are assumed to be uncorrelated, the eigenfunctions across levels are not assumed to be mutually orthogonal. In practice, only finite numbers of eigencomponents are selected in the KL expansions in Equation (2), denoted by L (region-level) and M (facility-level). The number of eigencomponents can be chosen by cross validation (Rice & Silverman, 1991), Akaike information criterion (Yao et al. 2005), or the estimated fraction of variance explained (FVE). We found FVE to be effective in the selection of the number of eigencomponents in applications given in this paper.

To capture dependencies due to potential region-level or dialysis-chain-specific practices, spatial correlations are induced among the region-specific PC scores $\xi_{i\ell}$ via a CAR model. More specifically, let $W = \{w_{ii'}\}$ denote the $n \times n$ adjacency matrix, describing the neighbourhood structure among the regions, with $w_{ii'} = 1$ if regions i and i' ($i \neq i'$) are neighbours (denoted by $i \sim i'$) and $w_{ii'} = 0$ otherwise. The diagonal elements of W are set to zero by convention. Further let $d_i = \sum_{i' \sim i} w_{ii'}$ denote the total number of neighbours of region i and $D = \mathrm{diag}(d)$ denote the $n \times n$ diagonal matrix with $d = (d_1,...,d_n)$. The full conditional distribution for the ℓth PC score for region i, $\xi_{i\ell}$, is specified by a Markov random field (MRF): $\xi_{i\ell}|\{\xi_{i'\ell}\}_{i' \neq i} \sim N\left(\nu\sum_{i' \sim i} w_{ii'}\xi_{i'\ell}/d_i, \alpha_\ell/d_i\right)$, with the variance component α_ℓ and the spatial correlation parameter ν. Hence, the conditional mean of $\xi_{i\ell}$ is a weighted average of the ℓth PC scores from neighbours of region i. Under this specification, the joint distribution of the ℓth PC scores $\boldsymbol{\xi}_\ell = (\xi_{1\ell},...,\xi_{n\ell})^{\mathsf{T}}$ can be derived as $\boldsymbol{\xi}_\ell \sim N\{\mathbf{0}, \alpha_\ell(D - \nu W)^{-1}\}$ through Brook's lemma. It has been shown that when the spatial correlation parameter ν is in $(0, 1)$, the precision matrix $(D - \nu W)/\alpha_\ell$ is guaranteed to be positive definite; see, for example, the discussion on p. 82 in Banerjee et al. (2014). The CAR model borrows spatial information across neighbours and therefore can be thought of as a smoother over neighbouring regions, stabilizing the estimates for small regions. Facility-specific PC scores, ζ_{ijm}, are assumed to be uncorrelated with $E(\zeta_{ijm}) = 0$ and $\mathrm{var}(\zeta_{ijm}) = \lambda_{im}$. Note that the facility-level variance λ_{im} is allowed to vary across regions. Finally, the measurement error, $\epsilon_{ij}(t)$, are assumed to be i.i.d with mean zero and variance σ^2, and uncorrelated with both region- and facility-specific PC scores.

2.2 | Estimation and inference

The proposed estimation algorithm is outlined in the table below, with key details highlighted in this section. The estimation algorithm utilizes P-splines and FPCA to reduce the dimension of the multilevel varying coefficient functions (VCFs) ($\boldsymbol{\beta}(t)$ and $\boldsymbol{\theta}(t)$) and the multilevel time-dynamic stochastic deviations ($U_i(t)$ and $V_{ij}(t)$), respectively. To achieve dimension reduction in the multilevel time-varying deviations, the region- and facility-level eigenfunctions, $\psi_{\ell}^{(1)}(t)$ and $\psi_{m}^{(2)}(t)$, are targeted using the between- and within-facility covariances (Steps 2–4 in the estimation algorithm below). The between- and within-facility covariances are obtained using residuals $E_{ij}(t)$ from a multilevel varying coefficient fit to the data in

a preliminary step under the working independence assumption (Step 1). The residuals $E_{ij}(t)$ represent the remaining deviation in the data after adjusting for time-dynamic effects of multilevel covariates. Keeping the estimated region- and facility-level eigenfunctions fixed, the remaining model parameters (including the P-spline coefficients from expansions of the multilevel VCFs and the variance components, α_ℓ, λ_{im}, ν, and σ^2) and the multilevel PC scores ($\xi_{i\ell}$ and ζ_{ijm}) are targeted in a mixed effects modelling framework using MCMC (Steps 5–6). Posterior distributions of the P-spline coefficients and the PC scores are utilized for inference on the multilevel VCFs.

Estimation algorithm

Step 1: Fit a M-VCM to the observed data under the working independence assumption and obtain the residuals $E_{ij}(t)$.

Step 2: Obtain the raw estimators of the total variance $\hat{G}_{Ti}(t,t') = \widehat{\text{cov}}\{E_{ij}(t), E_{ij}(t')\}$, and between-facility covariance of the residuals, $\hat{G}_{Bi}(t,t') = \widehat{\text{cov}}\{E_{ij}(t), E_{ij'}(t')\}$, within region i, directly using empirical covariances. The raw estimator of within-facility covariance in region i is obtained by the difference $\hat{G}_{Wi}(t,t') = \hat{G}_{Ti}(t,t') - \hat{G}_{Bi}(t,t')$.

Step 3: Estimate the between-facility ($\hat{G}_B(t,t')$) and within-facility ($\hat{G}_W(t,t')$) covariances by smoothing the average of the raw between- $\left(\sum_{i=1}^{n}\hat{G}_{Bi}(t,t')/n\right)$ and within-facility $\left(\sum_{i=1}^{n}\hat{G}_{Wi}(t,t')/n\right)$ covariances across all regions.

Step 4: Employ FPCA on $\hat{G}_B(t,t')$ and $\hat{G}_W(t,t')$ to obtain the region- and facility-level eigenfunction estimators: $\{\hat{\psi}_\ell^{(1)}(t) : \ell = 1,2,\dots,L\}$ and $\{\hat{\psi}_m^{(2)}(t) : m = 1,2,\dots,M\}$.

Step 5: Keeping the estimated eigenfunctions fixed, combine expansions of the VCFs ($\beta(t)$ and $\theta(t)$) onto B-spline basis and of the stochastic deviations ($U_i(t)$ and $V_{ij}(t)$) onto the estimated multilevel eigenfunctions under a mixed effects modelling framework.

Step 6: Estimate the VCFs $\beta(t)$ and $\theta(t)$, the variance components (α_ℓ, ν, λ_{im} and σ^2), and PC scores ($\xi_{i\ell}$ and ζ_{ijm}) via MCMC.

More specifically, the proposed estimation algorithm begins with an M-VCM fit to the data under the working independence assumption: $Y_{ij}(t) = \boldsymbol{X}_i^{\mathsf{T}}\boldsymbol{\beta}(t) + \boldsymbol{Z}_{ij}(t)^{\mathsf{T}}\boldsymbol{\theta}(t) + \epsilon_{ij}(t)$. The residuals from this initial fit, denoted by $E_{ij}(t)$, capture the remaining spatiotemporal variation in the data after adjusting for the time-varying effects of multilevel covariates and are used in targeting the between- ($G_B(t,t')$) and within-facility covariances ($G_W(t,t')$) in Steps 2 and 3. FPCA is employed on the between- and within-facility covariances to target the region- and facility-level eigenfunctions $\{\psi_\ell^{(1)}(t), \ell = 1,2,...,L\}$ and $\{\psi_m^{(2)}(t), m = 1,2,...,M\}$ in Step 4, similar to Li et al. (2021). The number of eigenfunctions (L and M) included in the decompositions of the region- ($U_i(t)$) and facility-specific deviations ($V_{ij}(t)$) are selected by FVE. To stabilize the estimation of the region-specific variance parameter $\text{var}(\zeta_{ijm}) = \lambda_{im}$, a two-compartment model is utilized with $\lambda_{im} = \lambda_{i1}\omega_m$, where λ_{i1} is allowed to change across regions, but ω_m, denoting the proportion of FVE of the mth second-level eigencomponent to FVE of the leading second-level eigencomponent (with $\omega_1 = 1$), is assumed to be fixed. In other words, while the ordering of the second-level eigencomponents, as well as their FVE, is assumed to be fixed across regions, regions are still allowed to be different in the variance of the second-level PC scores through multiplication by λ_{i1}. The term ω_m, which is constant across regions, is estimated as $\hat{\omega}_m = \hat{\Lambda}_m/\hat{\Lambda}_1, m = 2,...,M$, where $\{\Lambda_m : m = 1,2,...,M\}$ are the eigenvalues of the within-facility covariance $G_W(t,t')$.

Step 5 combines expansions of the VCFs $\boldsymbol{\beta}(t)$ and $\boldsymbol{\theta}(t)$ onto B-spline basis and of the region- and facility-specific deviations onto the estimated eigenfunctions from Step 4, in a mixed effects framework. The VCFs are modelled using Bayesian P-splines: $\beta_p(t) = \sum_{r=1}^{R} B_{pr}\phi_r(t), \theta_q(t) = \sum_{r=1}^{R} C_{qr}\phi_r(t)$, where $\phi_r(t), r = 1,...,R$, denote the B-spline basis functions and B_{pr} and C_{qr} denote the corresponding coefficients. As proposed by Lang and Brezger (2004), the priors for $\boldsymbol{B}_p = (B_{p1},...,B_{pR})^{\mathsf{T}}$ and $\boldsymbol{C}_q = (C_{q1},...,C_{qR})^{\mathsf{T}}$ are taken to be $\boldsymbol{B}_p|\tau_{Bp}^2 \propto \exp\{-\boldsymbol{B}_p^{\mathsf{T}}K_p\boldsymbol{B}_p/(2\tau_{Bp}^2)\}, \boldsymbol{C}_q|\tau_{Cq}^2 \propto \exp\{-\boldsymbol{C}_q^{\mathsf{T}}K_q\boldsymbol{C}_q/(2\tau_{Cq}^2)\}$, where K_p and K_q denote the $R \times R$ penalty matrices and τ_{Bp} and τ_{Cq} denote the hyperparameters. Let $\boldsymbol{Y}_{ij} = \{Y_{ij}(t_1),...,Y_{ij}(t_T)\}^{\mathsf{T}}$ denote the $T \times 1$ response vector; $\Phi = (\boldsymbol{\phi}_1,...,\boldsymbol{\phi}_R)$ denote the $T \times R$ matrix of B-spline basis functions, with $\boldsymbol{\phi}_r = \{\phi_r(t_1),...,\phi_r(t_T)\}^{\mathsf{T}}$ and $B = (B_1,...,B_P)^{\mathsf{T}}$; and $C = (C_1,...,C_Q)^{\mathsf{T}}$ denote the $P \times R$ and $Q \times R$ matrices of B-spline coefficients, respectively. Keeping the estimated $\{\hat{\psi}_\ell^{(1)}(t), \hat{\psi}_m^{(2)}(t), \hat{\omega}_m, L, M\}$ fixed and assuming the PC scores and measurement errors are normally distributed, the mixed effects model is given as follows:

$$\boldsymbol{Y}_{ij} = \Phi B^{\mathsf{T}}\boldsymbol{X}_i + (\Phi C^{\mathsf{T}}) \circ Z_{ij}\boldsymbol{1} + \Psi^{(1)}\tilde{\boldsymbol{\xi}}_i + \Psi^{(2)}\boldsymbol{\zeta}_{ij} + \boldsymbol{\epsilon}_{ij},$$
$$\boldsymbol{B}_p \propto \exp\left(-\frac{1}{2\tau_{Bp}^2}\boldsymbol{B}_p^{\mathsf{T}}K_p\boldsymbol{B}_p\right), \boldsymbol{C}_q \propto \exp\left(-\frac{1}{2\tau_{Cq}^2}\boldsymbol{C}_q^{\mathsf{T}}K_q\boldsymbol{C}_q\right),$$
$$\boldsymbol{\xi}_\ell \sim N(\boldsymbol{0}, \alpha_\ell(D - \nu W)^{-1}), \boldsymbol{\zeta}_{ijm} \sim N(0, \lambda_{i1}\omega_m), \epsilon_{ij}(t_k) \sim N(0, \sigma^2),$$

where '$\circ$' denotes element-wise matrix product, $\boldsymbol{X}_i = (X_{i1},...,X_{ip})^{\mathsf{T}}$ and $Z_{ij} = \{Z_{ij}(t_1),...,Z_{ij}(t_T)\}^{\mathsf{T}}$ denote the $P \times 1$ and $T \times Q$ region- and facility-level design vector and matrix, $\Psi^{(1)} = (\boldsymbol{\psi}_1^{(1)},...,\boldsymbol{\psi}_L^{(1)})$ and $\Psi^{(2)} = (\boldsymbol{\psi}_1^{(2)},...,\boldsymbol{\psi}_M^{(2)})$ denote the $T \times L$ and $T \times M$ matrices, respectively, made up of the region- and facility-level eigenfunctions $\boldsymbol{\psi}_\ell^{(1)} = \{\psi_\ell^{(1)}(t_1),...,\psi_\ell^{(1)}(t_T)\}^{\mathsf{T}}$ and $\boldsymbol{\psi}_m^{(2)} = \{\psi_m^{(2)}(t_1),...,\psi_m^{(2)}(t_T)\}^{\mathsf{T}}$, $\tilde{\boldsymbol{\xi}}_i = (\xi_{i1},...,\xi_{iL})^{\mathsf{T}}, \boldsymbol{\zeta}_{ij} = (\zeta_{ij1},...,\zeta_{ijM})^{\mathsf{T}}$ denote the $L \times 1$ and $M \times 1$ vectors containing the level 1 and level 2 PC scores, and $\boldsymbol{\epsilon}_{ij} = \{\epsilon_{ij}(t_1),...,\epsilon_{ij}(t_T)\}^{\mathsf{T}}$ denotes the $T \times 1$ error vector. In Step 6, model

parameters are targeted via MCMC sampling with inverse Gamma (IG) priors for variance components $\alpha_\ell \sim IG(a_{\alpha_\ell}, b_{\alpha_\ell}), \lambda_{i1} \sim IG(a_\lambda, b_\lambda)$ and $\sigma^2 \sim IG(a_{\sigma^2}, b_{\sigma^2})$, hyperparameters $\tau^2_{Bp} \sim IG(a_{\tau^2}, b_{\tau^2})$ and $\tau^2_{Cq} \sim IG(a_{\tau^2}, b_{\tau^2})$ and Beta priors for the spatial correlation parameter $\nu \sim \text{Beta}(a_\nu, b_\nu)$. Posterior distributions of the model parameters and details on sampling methods are deferred to Appendix S1. For inference on the VCFs $\boldsymbol{\beta}(t)$ and $\boldsymbol{\theta}(t)$, we form simultaneous credible bands as proposed in Crainiceanu et al. (2007). Let $f(t)$ denote a single varying coefficient function (in either $\boldsymbol{\beta}(t)$ or $\boldsymbol{\theta}(t)$), observed at time points t_k, for $k = 1, 2, ..., T$. Further let $\hat{f}(t)$ and $\text{SD}\{f(t)\}$ denote the mean and standard deviation of $f(t)$ based on a total of J MCMC samples $f^{(j)}(t)$ and let c_b denote the $(1-\alpha)$ sample quantile of $\max_{k=1,...,T} |\{f^{(j)}(t_k) - \hat{f}(t_k)\}/\text{SD}\{f(t_k)\}|, j = 1, 2, ..., J$. Then, the $(1-\alpha)$ simultaneous credible band for $f(t)$ is given by $[\hat{f}(t_k) \pm c_b \text{SD}\{f(t_k)\}]$.

3 | DATA ANALYSIS

3.1 | Description of the USRDS study cohort and multilevel risk factors

The facility hospitalization rates PPY, calculated monthly over the 2-year follow-up on dialysis, are modelled using the USRDS, a large database that collects data on nearly all patients on dialysis in the United States. The cohort includes dialysis patients of age 18 years or older who transitioned to dialysis between 1 January 2005 and 30 September 2013, followed up for 2 years with the last date of follow-up on 31 December 2015. Patients are followed up starting from the 91st day of dialysis treatment (after a 90-day period to establish stable treatment modality). Regional units are taken to be HSAs in the contiguous United States, consistent with national USRDS reporting. HSAs are merged to guarantee that each resulting region contains at least four facilities, for stable estimation and inference. After merging, the final study cohort includes a total of 5494 facilities and 423 regions/HSAs, with an average hospitalization rate of 1.8 per person-year (PPY). Detailed descriptions of the study cohort, exclusion rules and the region merging algorithm are deferred to Appendix S2.

The region-level covariates considered include urbanicity, ADI and medical underservice index (MUI). To capture urbanicity, HSAs are categorized into three classes: large metropolitan, small metropolitan or non-metropolitan. The classes are determined by the categorization assigned to the majority of the counties within each HSA, by the urban-rural classification scheme from National Center for Health Statistics (https://www.cdc.gov/nchs/data_access/urban_rural.htm). The HSA is assigned to the larger category to break occasional ties in the number of county designations. Non-metropolitan regions are taken as the reference group. The second region-level covariate, ADI, reflects the HSA's socio-economic status, consisting of 17 education, employment, housing-quality and poverty measures (Kind & Buckingham, 2018). The index, available at the level of census block groups (available at https://www.neighborhoodatlas.medicine.wisc.edu), is a rank-based index taking on values between 0 and 100, with higher values corresponding to lower socio-economic status (and higher deprivation). ADIs assigned to census block groups within each HSA are averaged to derive the HSA-level indices. The last region-level covariate, MUI, is used to reflect the medical service availability within each HSA. Medically underserved areas (MUA) are areas with too few primary care providers, high infant mortality, high poverty or a high elderly population, designated by Health Resources and Services Administration. The index is available at the census tract/county subdivision level at https://data.hrsa.gov/tools/shortage-area. The proportion of census tracts/county subdivisions that are designated as MUA is first targeted for each county, and county MUIs are then averaged within each HSA to arrive at the HSA-level MUI index, with higher indices corresponding to higher underservice. Finally, facility-level covariates summarize the patient demographic makeup and comorbidity burden at initiation of dialysis for each dialysis facility. Covariates considered include patients' average age and body mass index (BMI), percentage of female patients within a facility and percentage of patients with diabetes as the cause of ESRD and with the comorbidities of COPD, septicemia, other infections and psychological disorders. Note that even though patient characteristics from initiation of dialysis are summarized, the facility-level covariates are time-varying due to the dynamic nature of the patient cohort served by each dialysis facility (due to patients changing dialysis facilities or death). The covariates ADI, MUI, age, BMI and percentage of female patients are mean-centred for ease of interpretation in modelling.

3.2 | Results

3.2.1 | Time-varying effects of multilevel covariates

The estimated time-varying effects of region-level covariates including the time-varying y-intercept, representing the hospitalization rate of an 'average' facility in a non-metropolitan region (with median ADI of 53, MUI of 0.43, patient mean age of 63, BMI of 29 and female at 44%, patient comorbidities and diabetes as cause of ESRD at the reference level), are given in Figure 1. The estimated y-intercept displays an overall decreasing trend in hospitalization rates throughout the 2-year follow-up, with the highest rate of 1.77 PPY early after transitioning to dialysis (dropping down to 1.35 PPY at 2-year follow-up) (Figure 1a). Figure 1b–e displays the estimated effects of region-level covariates (solid) with simultaneous 95% credible bands (dashed), overlaying horizontal lines at zero (grey), included for reference. Both ADI and MUI are positively associated with hospitalization rates, suggesting that regions with a higher deprivation level or lower medical service availability have higher hospitalization rates.

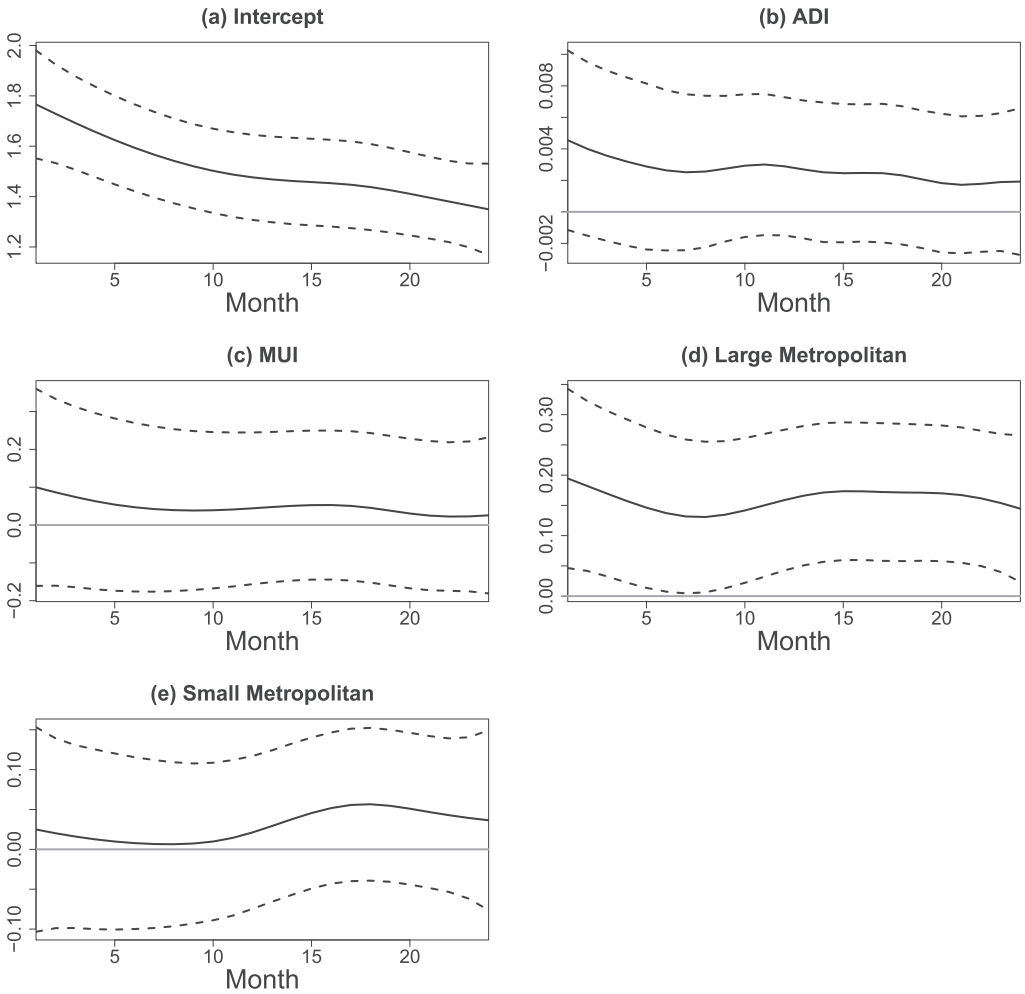

FIGURE 1 Estimated time-varying effects (solid) of region-level covariates: (a) *y*-intercept, (b) area deprivation index (ADI) (centred), (c) medical underservice index (MUI) (centred), (d) large metropolitan and (e) small metropolitan, along with their 95% simultaneous credible intervals (dashed), in modelling hospitalization rates among the US dialysis population. Horizontal lines at zero are included in grey for reference. Positive coefficients correspond to increased hospitalization rates

However, neither ADI nor MUI is found significant at the 0.05 significance level. Large metropolitan regions have significantly higher hospitalization rates than non-metropolitan regions (reference group), where the difference between hospitalization rates at small metropolitan and non-metropolitan regions is not found to be significant.

The estimated time-varying effects (solid) of the facility-level covariates, along with their simultaneous 95% credible bands (dashed), are given in Figure 2 (Horizontal lines are included in grey for reference). The estimated time-varying effect of age (adjusted for comorbidities) is positive at initiation of dialysis but becomes mostly nonsignificant for the 2-year follow-up. This can be explained by the mostly positive effects of comorbidities that partly explain away the effects of age (older patients typically have more comorbidities than younger patients). Higher BMI is associated with lower hospitalization risk for approximately the first 10 months on dialysis, where the protective effect gets weaker as patients stay longer on dialysis. The protective effect of BMI is a well-known phenomenon, as also documented in other studies on adverse events in the dialysis cohort, such as cardiovascular risk and mortality (Kalantar-Zadeh et al. 2005). While facilities with a higher percentage of female patients have higher hospitalization rates (consistent with previous findings), all comorbidities considered are mostly associated with higher hospitalization risk as expected. More specifically, while the effects of chronic conditions such as COPD and psychological disorders are getting stronger as patients stay longer on dialysis (COPD significant after 15 months and psychological disorders significant around 15 months on dialysis), effects of acute conditions such as septicemia weaken as patients stay longer on dialysis (significant stronger effects of septicemia observed within the first 10 months on dialysis). The effect of having diabetes as a cause of ESRD and other infections are mostly not found to be significant.

To further assess the effects of facility-level risk factors on hospitalization rates, Figure 3 displays the predicted hospitalization trajectories for two facilities with the percentage of patients with septicemia at 0% and 20% (while other covariates are kept at median or reference levels as specified above for the y-intercept). The 95% simultaneous credible bands (shaded in grey) for the predicted hospitalization trajectories are

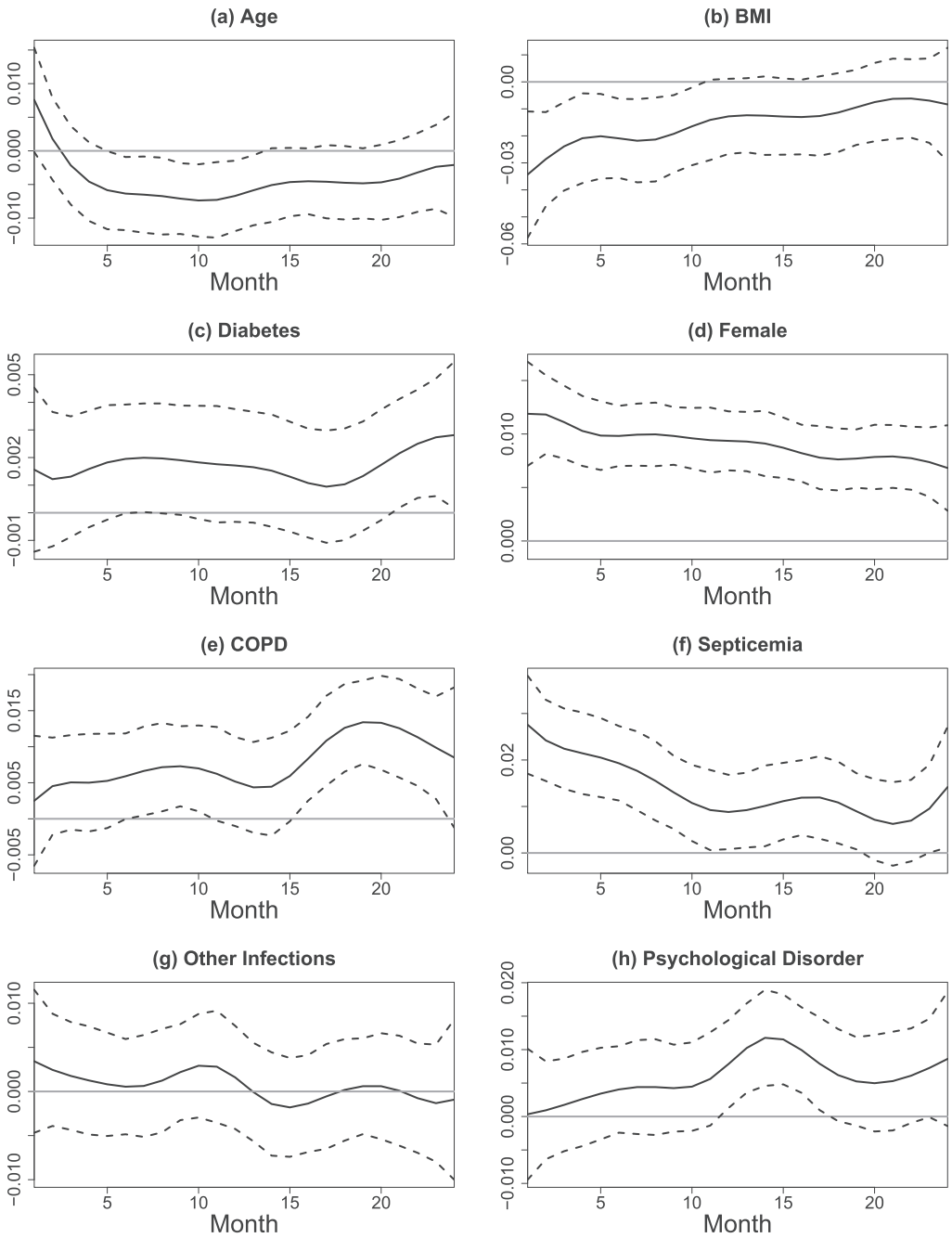

FIGURE 2 Estimated time-varying effects (solid) of facility-level covariates: (a) age (centred), (b) BMI (centred), (c) % of patients having diabetes as the cause of ESRD, (d) % of female patients (centred) and % patients with (e) COPD, (f) septicemia, (g) other infection and (h) psychological disorders, along with their 95% simultaneous credible intervals (dashed), in modelling hospitalization rates among the US dialysis population. Horizontal lines at zero are included in grey for reference. Positive coefficients correspond to increased hospitalization rates

formed based on the draws from $Y_{ij}(t) = \boldsymbol{X}_i^{\mathsf{T}}\beta(t) + \boldsymbol{Z}_{ij}(t)^{\mathsf{T}}\theta(t)$, corresponding to each posterior draw of the multilevel VCFs, and following the algorithm outlined at the end of Section 2.2 with $f(t) = Y_{ij}(t)$. A 20% increase in the number of patients with septicemia in a given facility (while all other covariates are kept fixed) corresponds to a significant difference in hospitalization risk trajectories in approximately the first 4 months on dialysis. This is consistent with the weakening effects of septicemia as patients stay longer on dialysis, displayed in Figure 2. More specifically, while a 20% increase in the number of patients with septicemia corresponds to an increase of 0.55 PPY in hospitalization rates at initiation of dialysis, the difference drops to 0.28 PPY at the end of the 2-year follow-up. In comparison, the difference in hospitalization rates of facilities in large metropolitan versus non-metropolitan areas is 0.19 PPY at initiation of dialysis, dropping down to 0.14 PPY at the end of the 2-year follow-up (Figure 1d). Hence, while the magnitude of the effects of significant region- and facility-level covariates are comparable, the effects of more

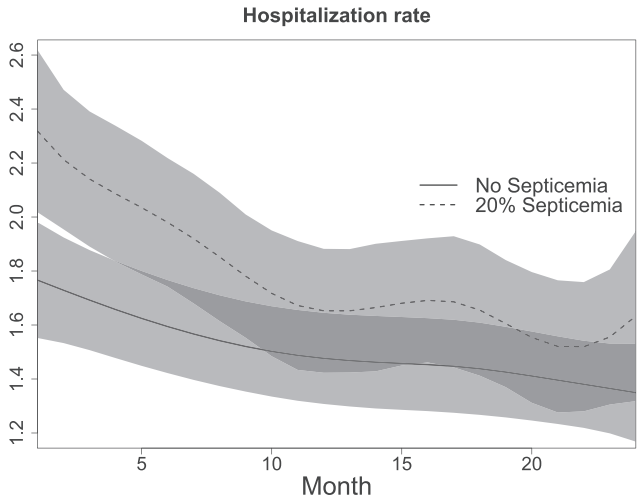

FIGURE 3 Predicted hospitalization trajectories for two facilities with per cent of patients with septicemia at 0% (solid) and 20% (dashed) (while other covariates are kept at median or reference levels) along with 95% simultaneous credible bands (shaded in grey)

facility-level covariates are time-dynamic (e.g., COPD, septicemia, BMI and female), compared to region-level covariates (where the effects are found to be relatively stable across follow-up).

3.2.2 | Spatiotemporal patterns of hospitalization rates

In addition to modelling time-varying effects of multilevel covariates, the proposed M-VCSM models the remaining spatiotemporal patterns in the data via multilevel stochastic deviations ($U_i(t)$ and $V_{ij}(t)$) with spatial correlations induced at the region-level. Multilevel FPCA is utilized in reducing the dimension of $U_i(t)$ and $V_{ij}(t)$, where spatial correlations are induced among the region-level eigenscores. The estimated multilevel eigenfunctions are displayed in Figure 4. At the region-level, the leading eigenfunction explains 95% of the variation, highlighting higher variation in hospitalization rates across regions during the first 5 to 10 months of dialysis. This is consistent with the elevated hospitalizations observed at initiation of dialysis. Figure 4b-d displays the three leading facility-level eigenfunctions, explaining more than 85% of the total facility-level variation. The leading facility-level eigenfunction (with 71% FVE) is relatively flat, capturing the constant variation across the 2-year follow-up among facilities. The second leading eigenfunction (with 11% FVE) highlights variation at beginning and end of follow-up, while the third leading eigenfunction (with 5% FVE) highlights variation also around the mid follow-up point, at 1 year on dialysis (in addition to the beginning and end of follow-up). The estimated spatial correlation parameter, $\hat{\nu}$, equals 0.94, inducing correlations between neighbouring HSAs ranging from 0.28 to 0.67. The findings affirm that there is still significant spatiotemporal variation remaining in the data, after adjusting for time-varying effects of multilevel covariates. To visualize the remaining spatiotemporal patterns in the data, Figure 5 displays the residuals $E_{ij}(t)$ (averaged across time) obtained from the initial M-VCM fit to the data under the working independence assumption (Step 1 of the estimation algorithm). Also displayed in Figure 5 are the predicted region-specific deviations $U_i(t)$ averaged across follow-up time. Both plots confirm the significant spatial variation remaining in the data after adjustment by covariates. In addition, the similarity between the two maps show that the predicted region-specific deviations are able to capture the remaining spatial variation in the data effectively. Note also that the remaining spatial patterns in the residuals $\epsilon_{ij}(t)$ (given in Figure S3), obtained after the proposed spatiotemporal modelling at the region-level, are much reduced. The spatial patterns in Figure 5 highlight higher hospitalizations in the 'band' from Massachusetts to southern Texas (dark blue), as well as Nevada, Arizona and Florida. In addition, spatial variation across regions is of the similar order to temporal variation observed across the 2-year patients stay on dialysis: the interquartile range of the predicted region-specific deviations is 0.35 PPY, while hospitalization risk decreases by 0.42 PPY throughout the 2-year follow-up.

To assess the spatial and temporal variation jointly while also taking into account time-varying effects of multilevel covariates on hospitalization risk, Figure 6 displays raw hospitalization rates as well as region-specific predictions from the full model at 1, 12 and 24 months after initiation of dialysis. Region-specific predictions are obtained via averaging predicted facility-level rates across all facilities within each region: $\hat{Y}_i(t) = \sum_{j=1}^{N_i} \{X_i^T \hat{\beta}(t) + Z_{ij}(t)^T \hat{\theta}(t) + \hat{U}_i(t) + \hat{V}_{ij}(t)\}/N_i$. Predicted maps correspond closely to raw maps indicating the satisfactory fit of the proposed M-VCSM model. Overall, both sets of maps highlight elevated hospitalization rates in Massachusetts to southern Texas, Nevada, Arizona and Florida, also accounting for effects of multilevel covariates, similar to the findings of Figure 5. Across the United States, rates are higher in the

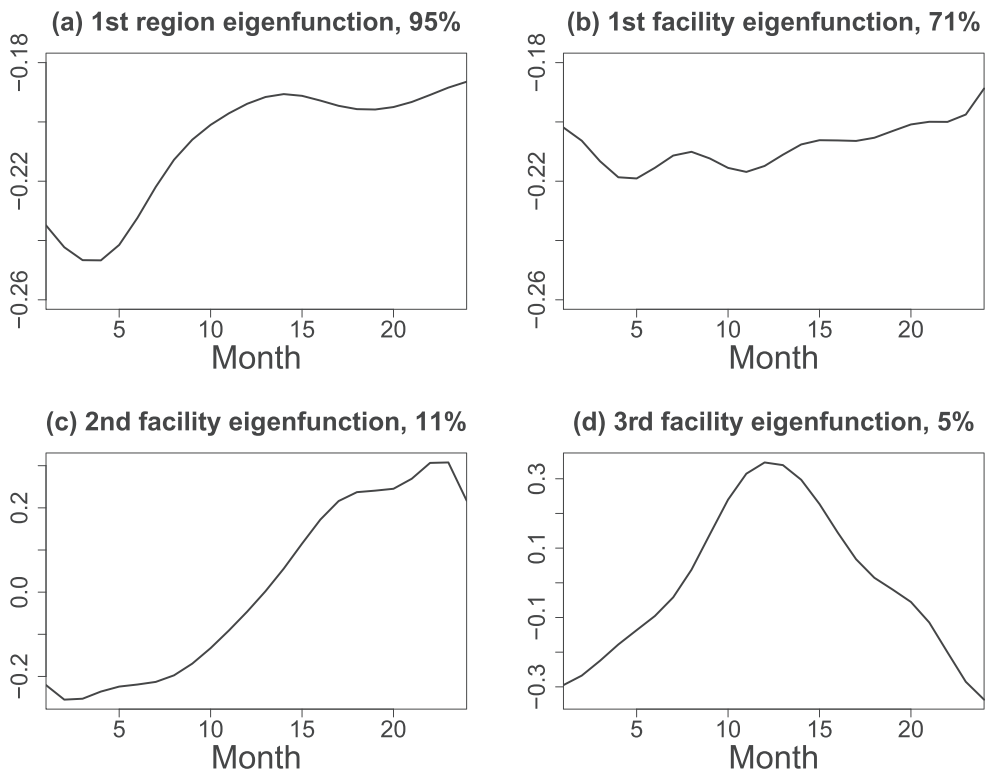

FIGURE 4 The estimated leading region-level eigenfunction (a) with 95% FVE and the leading (b), second leading (c) and third leading (d) facility-level eigenfunctions with FVE at 71%, 11% and 5%, respectively

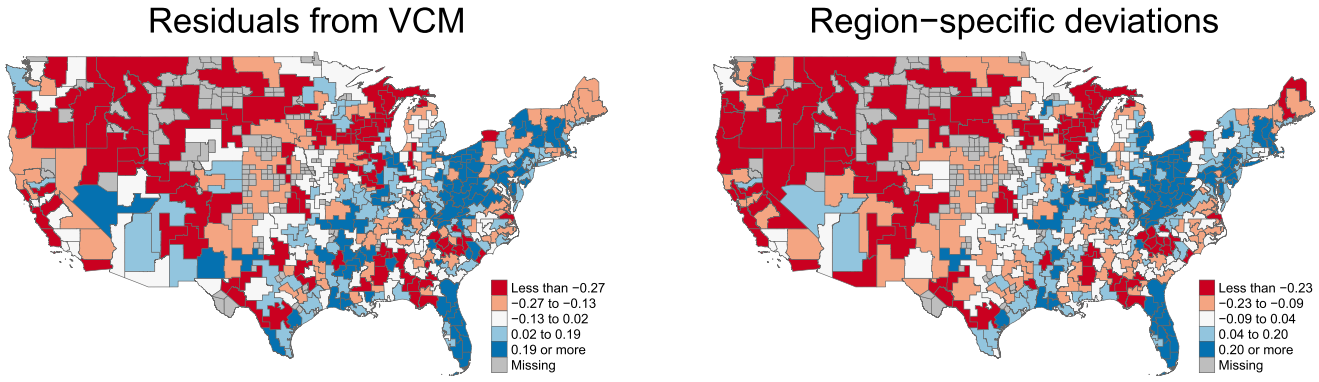

FIGURE 5 The residuals $E_{ij}(t)$ (averaged across time) obtained from the initial multilevel varying coefficient model fit to the data under the working independence assumption (left), predicted region-specific deviations $U_i(t)$ averaged across follow-up time (right)

northeast, northwest and central regions. Moreover, hospitalization rates are the highest in the early months after transitioning to dialysis, with an overall decreasing trend in almost all HSAs throughout the 2-year follow-up period.

4 | SIMULATION STUDIES

Finite sample properties of the proposed estimation algorithm and inference procedure are studied via simulation studies with a varying number of regions, number of facilities per region and measurement error variance. In addition, we compare the finite sample performance of the proposed M-VCSM with a multilevel varying coefficient (M-VCM) that ignores the spatial correlation at the region-level.

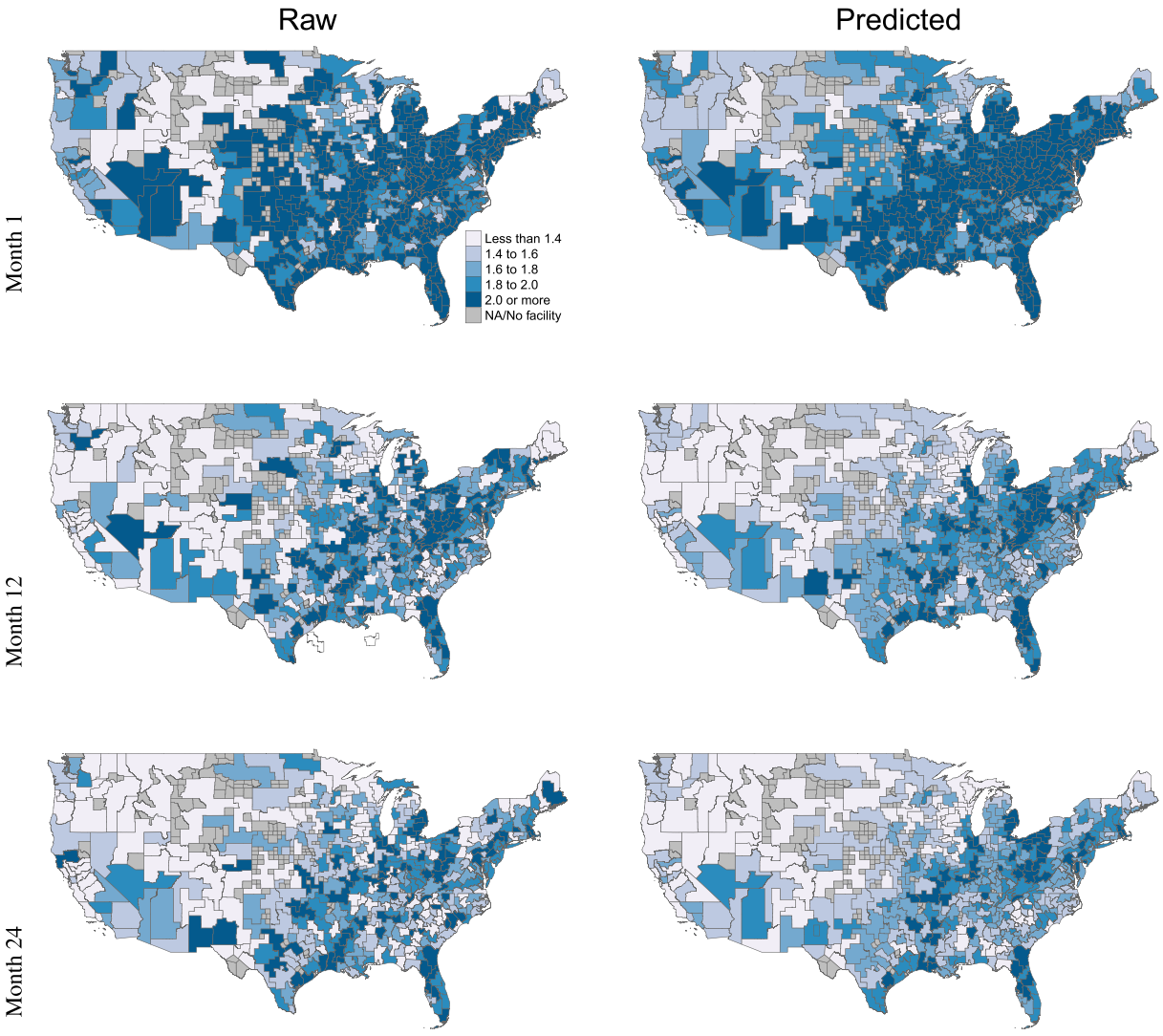

FIGURE 6 Raw (left column) and predicted (right column) hospitalization rates at 1, 12 and 24 months after initiation of dialysis for 423 HSAs across the United States. Predicted rates are obtained from the full model including effects of multilevel covariates and predictions for multilevel stochastic predictions

4.1 | Simulation design

We consider the simulation model $Y_{ij}(t_{ijk}) = X_i^{\mathsf{T}}\beta(t_{ijk}) + Z_{ij}(t_{ijk})^{\mathsf{T}}\theta(t_{ijk}) + \sum_{\ell=1}^{L}\xi_{i\ell}\psi_{\ell}^{(1)}(t_{ijk}) + \sum_{m=1}^{M}\zeta_{ijm}\psi_{m}^{(2)}(t_{ijk}) + \epsilon_{ij}(t_{ijk})$, evaluated at an equidistant grid of $k = 1,\ldots,24$ time points, t_{ijk}, between 0 and 1, mimicking the 24-month follow-up in our data application. The region-level covariates $X_i = (1, X_{i1}, X_{i2})^{\mathsf{T}}$ include a y-intercept term, and X_{i1}, X_{i2} are generated from $U(0,1)$, $N(0,1)$, respectively. To mimic the USRDS data where facility-level covariates are time-dependent, $Z_{ij}(t_{ijk}) = \{Z_{ij1}(t_{ijk}), Z_{ij2}(t_{ijk})\}^{\mathsf{T}}$, facility-level covariates are generated as $Z_{ij1}(t) = -t + \epsilon_1, Z_{ij2}(t) = t + \epsilon_2$ where ϵ_1 and ϵ_2 are generated independently from standard normal distributions. The region- and facility-level VCFs equal $\beta(t) = \{\beta_0(t), \beta_1(t), \beta_2(t)\}^{\mathsf{T}} = (1 - t, 2t, -4t^2 + 4t - 1)^{\mathsf{T}}$ and $\theta(t) = \{\theta_1(t), \theta_2(t)\}^{\mathsf{T}} = (2t^2 - 3t + 1, 10t^3 - 16t^2 + 7t - 1)^{\mathsf{T}}$, respectively. Region- and facility-level eigenfunctions equal $\psi_1^{(1)}(t) = \sqrt{2}\sin(2\pi t), \psi_2^{(1)}(t) = \sqrt{2}\cos(2\pi t)$ and $\psi_1^{(2)}(t) = \sqrt{3}(2t - 1), \psi_2^{(2)}(t) = \sqrt{5}(6t^2 - 6t + 1)$, respectively $(L = M = 2)$, where eigenfunctions are chosen not to be orthogonal across the two levels of the data. The region-specific PC scores $\boldsymbol{\xi}_{\ell} = (\xi_{1\ell}, \ldots, \xi_{n\ell})^{\mathsf{T}}$ are generated from a multivariate normal distribution with mean zero and covariance matrix $(1/\alpha_{\ell})(D - \nu W)^{-1}$, where W is the adjacency matrix and D is a diagonal matrix with diagonal elements equal to the number of neighbours for each region, as described in Section 2. The spatial correlation parameter ν equals 0.9, inducing spatial correlations between neighbouring regions in the range of 0.22 to 0.62, similar to the data application in Section 3.2.2. The spatial variance parameters α_1 and α_2 equal 1 and 0.25, respectively. The facility-specific PC scores ζ_{ijm} are generated from N

$(0, \lambda_{im})$, with λ_{i1} generated from a discrete uniform on $\{0.3, 0.2, 0.1\}$ and $\lambda_{i2} = 0.5\lambda_{i1}$, guaranteeing that FVE of the facility-level eigencomponents stay the same across regions while total variation still remains region-specific.

Eight simulation settings are considered with two sets of total region numbers, two sets of total facility numbers per region and two sets of measurement error variance. The two regional units considered are HSAs and states in the contiguous United States (including the District of Columbia), resulting in 423 (similar to the data application) and 49 total regions, respectively. In addition, two sets of total facility numbers per region are considered: $4 - 20$ and $10 - 30$. More specifically, the total number of facilities per region is generated as follows. Regions are first randomly designated into three groups: small (S), medium (M) and large (L), where the total number of facilities within each region is generated from a discrete uniform distribution on $S = \{4,5,6\}, M = \{7,8,9,10\}$, and $L = \{11,12,...,20\}$ in the first setting and $S = \{10,11,12,13\}, M = \{14,15,16,17\}$, and $L = \{18,19,...,30\}$ in the second. Finally, the measurement error, $\epsilon_{ij}(t_{ijk})$, is generated from $N(0, \sigma^2)$ with $\sigma^2 = 0.2$ and $\sigma^2 = 0.02$ in two separate simulation set-ups.

TABLE 1 The mean MSDE of VCFs, eigenfunctions and region- and facility-specific deviations and MSE for variance components and spatial correlation parameters from eight simulation settings with two sets of total region numbers, two sets of total facility numbers per region and two sets of measurement error variance σ^2

Number of regions	$n = 423$ regions				$n = 49$ regions			
Number of facilities	4-20		10-30		4-20		10-30	
Noise level, σ^2	0.2	1	0.2	1	0.2	1	0.2	1
	MSDE							
$\hat{\beta}_0(t)$	0.005	0.011	0.004	0.010	0.011	0.042	0.011	0.033
$\hat{\beta}_1(t)$	0.002	0.004	0.001	0.003	0.007	0.016	0.005	0.013
$\hat{\beta}_2(t)$	0.003	0.004	0.002	0.003	0.013	0.026	0.009	0.022
$\hat{\theta}_1(t)$	<0.001	0.001	<0.001	0.001	0.005	0.008	0.003	0.005
$\hat{\theta}_2(t)$	<0.001	0.001	<0.001	0.001	0.005	0.008	0.003	0.005
$\hat{\psi}_1^{(1)}(t)$	0.003	0.003	0.002	0.003	0.032	0.029	0.025	0.032
$\hat{\psi}_2^{(1)}(t)$	0.004	0.005	0.003	0.004	0.043	0.046	0.031	0.042
$\hat{\psi}_1^{(2)}(t)$	0.001	0.002	<0.001	0.001	0.013	0.019	0.004	0.007
$\hat{\psi}_2^{(2)}(t)$	0.001	0.003	0.001	0.001	0.016	0.026	0.005	0.010
$\hat{U}_i(t)$	0.129	0.246	0.089	0.180	0.122	0.286	0.084	0.222
Small	0.170	0.312	0.106	0.209	0.150	0.345	0.092	0.256
Medium	0.125	0.244	0.087	0.188	0.126	0.282	0.089	0.213
Large	0.092	0.184	0.074	0.145	0.088	0.231	0.071	0.197
$\hat{V}_{ij}(t)$	0.232	0.506	0.217	0.494	0.266	0.525	0.232	0.504
	MSE							
$\hat{\alpha}_1$	0.005	0.006	0.004	0.006	0.057	0.048	0.037	0.052
$\hat{\alpha}_2$	<0.001	0.001	<0.001	<0.001	0.004	0.005	0.003	0.004
$\hat{\nu}$	0.001	0.001	0.001	0.001	0.003	0.003	0.003	0.002
$\hat{\sigma}^2$	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
$\hat{\lambda}_{i1}$	0.005	0.006	0.003	0.004	0.006	0.007	0.003	0.004
Small	0.007	0.009	0.004	0.005	0.008	0.008	0.004	0.005
Medium	0.005	0.006	0.003	0.004	0.006	0.007	0.003	0.004
Large	0.003	0.004	0.002	0.003	0.003	0.004	0.002	0.003
$\hat{\lambda}_{i2}$	0.001	0.002	0.001	0.001	0.001	0.002	0.001	0.001
Small	0.002	0.002	0.001	0.001	0.002	0.002	0.001	0.001
Medium	0.001	0.002	0.001	0.001	0.001	0.002	0.001	0.001
Large	<0.001	0.001	<0.001	0.001	0.001	0.001	<0.001	0.001

Note: Results are based on 200 Monte Carlo runs.

4.2 | Results

The estimation of the time-varying and time-invariant parameters are assessed via the relative mean squared deviation error (MSDE), $\text{MSDE}_{\hat{f}} = \left[\int\left\{\hat{f}(t)-f(t)\right\}^2 dt\right] / \int f^2(t)dt$ for a generic function $f(t)$, and mean squared error (MSE), respectively. All VCFs are modelled with 20 cubic B-spline basis functions with equidistant knots and a second-order penalty. The mean MSDE and MSE values from eight simulation set-ups with two sets of total region numbers ($n = 423$ and $n = 49$), two sets of total facility numbers per region ($4-20$ and $10-30$) and two error variances ($\sigma^2 = 0.02$ and 0.2) are reported in Table 1. Reported results are based on a total of 200 Monte Carlo runs, with 12,000 iterations in each run (2000 for burn-in and 10,000 for estimation and inference) for the MCMC step. In addition to the time-varying model parameters, the MSDEs of the estimated region- and facility-specific deviations ($\hat{U}_i(t)$ and $\hat{V}_{ij}(t)$) are also reported. Reported summaries exclude outlier MSDE values (MSDE>5) for $\hat{U}_i(t)$ and $\hat{V}_{ij}(t)$ with denominator values close to zero (less than 4%).

MSDE for VCFs and region- and facility-specific deviations get smaller with decreasing noise level σ^2 as expected, while MSDE for eigenfunctions and variance components are comparable across the two noise levels, showing the efficacy in removal of the effects of measurement error via the proposed FPCA and the mixed effects modelling. The effects on VCFs (as part of the mean function) and the region- and facility-level deviation predictions are expected since it is harder to predict deviations under increasing measurement error noise. All error measures decrease with increasing the number of regions, except MSDE of region-specific deviations $U_i(t)$. MSE of region-specific variances λ_{im} is also

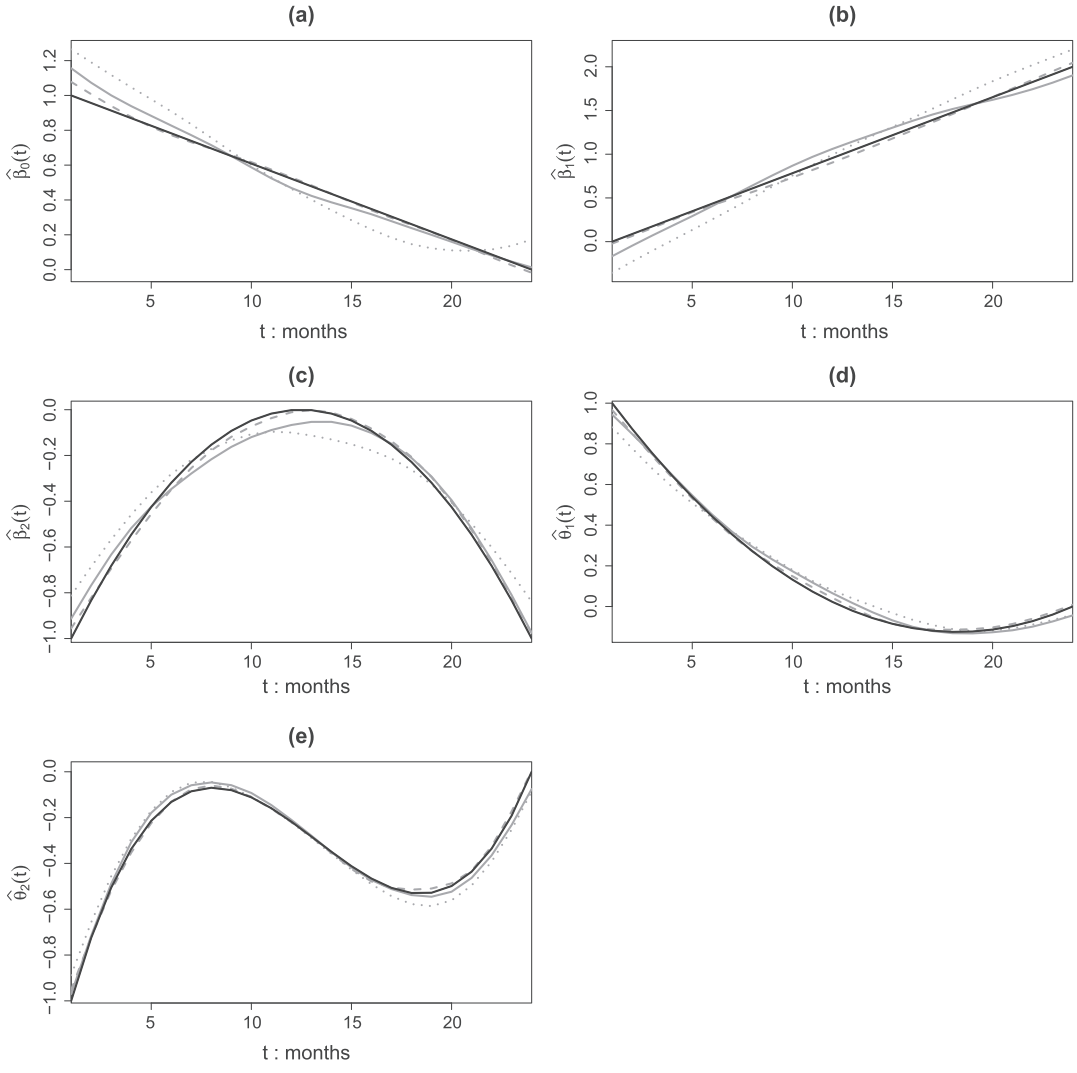

FIGURE 7 Estimated (a–c) region-level and (d–e) facility-level VCFs from runs with the 5th (dashed grey), 50th (solid grey) and 95th (dotted grey) percentile MSDEs using M-VCSM, from the set-up with 49 regions, $4-20$ number of facilities per region and $\sigma^2 = 0.2$. The true eigenfunctions are given in solid black

marginally effected by increasing the number of regions. This is as expected since the estimation of region-specific parameters mainly depends on within-region information. Indeed, estimation of region-specific parameters λ_{im} and deviation functions $U_i(t)$ are better in larger regions with increased region-specific information. All error measures decrease similarly with increasing the number of facilities. The effect of increasing the number of facilities is especially pronounced on facility-level modelling components including the facility-level eigenfunctions $\psi_m^{(2)}(t)$, facility-specific deviations $V_{ij}(t)$ and facility-level eigenscore variances λ_{im}, as expected.

The estimated time-varying coefficient functions of the multilevel risk factors $\beta(t)$ and $\theta(t)$ along with their simultaneous confidence bands from the simulation runs with the median MSDE based on $n = 49, \sigma^2 = 0.2$ and $4 - 20$ facilities per region are given in Figure S1. The estimated curves (solid black) track the true functions (solid grey), and the simultaneous confidence bands (dashed) cover the truth for all VCFs. To display the variability in the VCF estimates, we also display estimates from runs with the 5th, 50th and 95th percentile MSDEs in Figure 7. The estimated region- and facility-level eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs are displayed in Figure S2. The estimates also track the true eigenfunctions, indicating that the proposed estimation algorithm effectively targets the directions of variation at both region- and facility-levels adjusted for time-varying multilevel covariate effects.

The finite sample performance of the proposed M-VCSM is also compared to a M-VCM which ignores the spatial correlation at the highest level of the hierarchy (region-level). More specifically, we consider the M-VCM, $Y_{ij}(t) = X_i^{\mathsf{T}}\beta(t) + Z_{ij}(t)^{\mathsf{T}}\theta(t) + U_i(t) + V_{ij}(t) + \epsilon_{ij}(t)$, where $U_i(t) = \sum_{\ell=1}^{\infty}\xi_{i\ell}\psi_\ell^{(1)}(t), V_{ij}(t) = \sum_{m=1}^{\infty}\zeta_{ijm}\psi_m^{(2)}(t), \psi_\ell^{(1)}(t)$ and $\psi_m^{(2)}(t)$ are region- and facility-level eigenfunctions and $\xi_{i\ell}$ and ζ_{ijm} are the corresponding PC scores (Scheipl et al. 2015). Different from M-VCSM, the PC scores $\xi_{i\ell}$ in the M-VCM are assumed to be i.i.d across regions, ignoring the region-level spatial correlation. Similar to M-VCSM, the M-VCM also utilizes FPCA to reduce the dimension of the region- and facility-specific deviation and utilizes penalized B-splines to model VCFs. Following dimension reduction of the time-varying modelling components, estimation and inference for M-VCM is based on standard mixed effect modelling machinery and is implemented via `pffr` function in R package `refund` (Scheipl et al. 2015). The region- and facility-level eigenfunctions and eigenvalues that are estimated from FPCA analysis on between- and within-facility covariances described in estimation algorithm Steps 1–4 are utilized as input in `pffr`. The smoothing parameters associated with penalized B-spline fits are estimated using restricted maximum likelihood and inference is based on the approximate pointwise empirical Bayes' confidence intervals (Marra & Wood, 2012; Wood, 2017).

Model comparisons at four simulation set-ups (two total number of facilities per region and measurement error settings) for $n = 49$ regions are summarized in Table 2, including the MSDEs of VCFs and the region- and facility-specific deviations as well as the VCF coverage probabilities (CP). While MSDE of the facility-level VCFs are comparable between the two models, M-VCSM, modelling the spatial correlation at the region-level, leads to smaller MSDE for region-level VCFs than M-VCM, especially for $\hat{\beta}_0(t)$ and $\hat{\beta}_1(t)$. This difference is also apparent in Figures 7 and 8

TABLE 2 The mean MSDE of VCFs and region- and facility-specific deviations along with coverage probabilities of the 95% credible band/confidence interval of multilevel VCFs from M-VCSM and M-VCM, respectively, based on 200 Monte Carlo runs

	4–20 facilities per region				10–30 facilities per region			
Noise level, σ^2	**0.2**	**1**	**0.2**	**1**	**0.2**	**1**	**0.2**	**1**
MSDE	M-VCSM		M-VCM		M-VCSM		M-VCM	
$\hat{\beta}_0(t)$	0.011	0.042	0.104	0.130	0.011	0.033	0.106	0.109
$\hat{\beta}_1(t)$	0.007	0.016	0.044	0.031	0.005	0.013	0.032	0.035
$\hat{\beta}_2(t)$	0.013	0.026	0.020	0.027	0.009	0.022	0.019	0.025
$\hat{\theta}_1(t)$	0.005	0.008	0.006	0.008	0.003	0.005	0.003	0.005
$\hat{\theta}_2(t)$	0.005	0.008	0.006	0.008	0.003	0.005	0.003	0.005
$\hat{U}_i(t)$	0.122	0.286	0.739	0.706	0.084	0.222	0.775	0.649
Small	0.150	0.345	0.754	0.792	0.092	0.256	0.665	0.716
Medium	0.126	0.282	0.755	0.680	0.089	0.213	0.829	0.616
Large	0.088	0.231	0.708	0.648	0.071	0.197	0.830	0.616
$\hat{V}_{ij}(t)$	0.266	0.525	0.293	0.601	0.232	0.504	0.248	0.569
CP (%)								
$\hat{\beta}_0(t)$	100	97.5	49.1	55.9	100	97.5	44.8	50.9
$\hat{\beta}_1(t)$	100	97.0	73.2	83.6	100	99.0	77.8	82.0
$\hat{\beta}_2(t)$	100	97.0	98.3	96.7	100	94.0	97.7	96.5
$\hat{\theta}_1(t)$	100	97.0	94.5	96.3	100	98.0	94.0	94.9
$\hat{\theta}_2(t)$	100	97.0	93.9	95.8	100	97.0	95.6	94.4

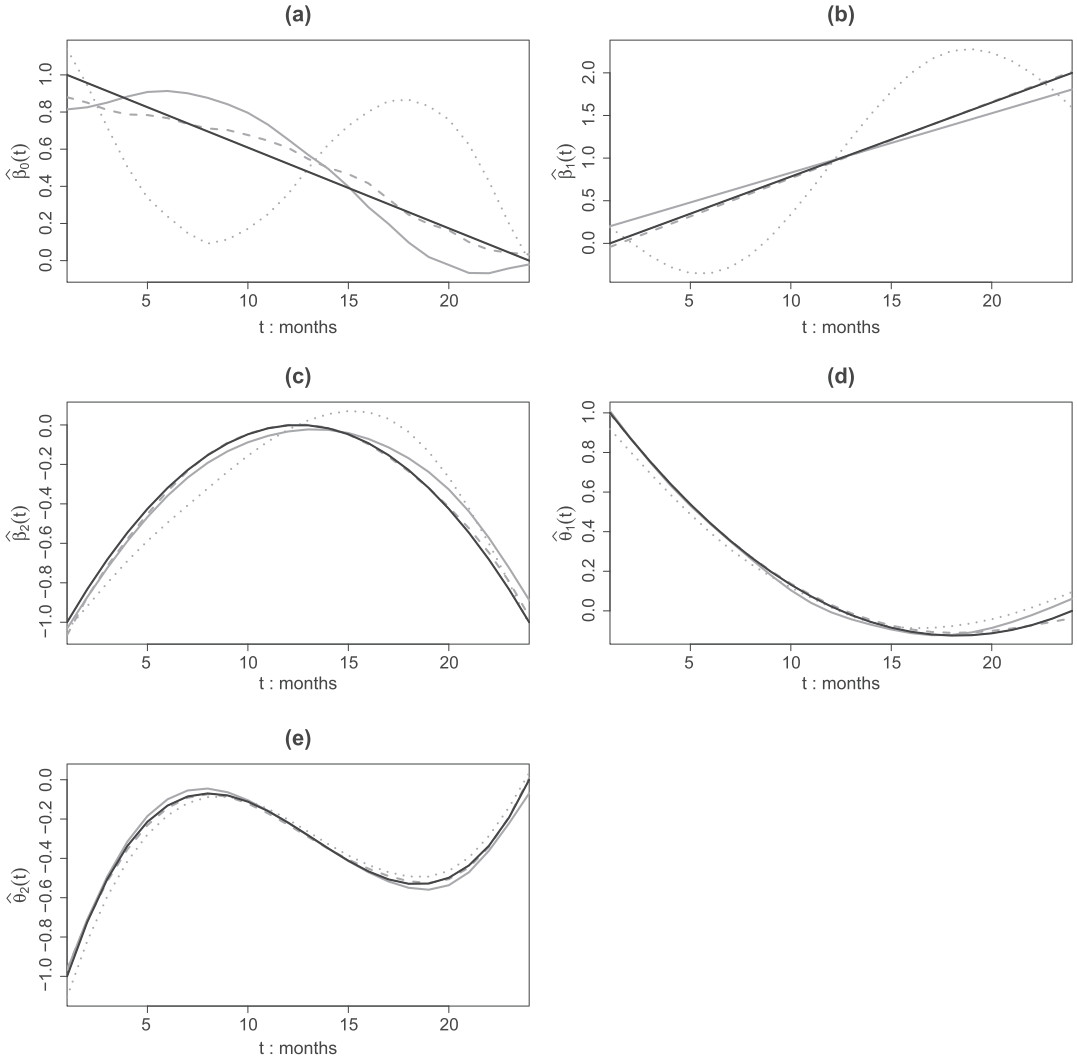

FIGURE 8 Estimated (a–c) region-level and (d–e) facility-level VCFs from runs with the 5th (dashed grey), 50th (solid grey) and 95th (dotted grey) percentile MSDEs using M-VCM, from the set-up with 49 regions, $4-20$ number of facilities per region and $\sigma^2 = 0.2$. The true eigenfunctions are given in solid black

which display VCF estimates from runs with the 5th, 50th and 95th percentile MSDEs for M-VCSM and M-VCM, respectively. The region-level VCF estimates from M-VCSM are less variable than M-VCM, for both linear and quadratic functions. In addition, CP for $\beta_0(t)$ and $\beta_1(t)$ from M-VCM is quite lower than the nominal at 95% (as low as 44.8%). Note that CPs reported from M-VCSM are higher than the nominal level at 95% most of the time. This is as expected since the Bayesian credible bands tend to be conservative (Cox, 1993; Krivobokova et al. 2010). While the MSDEs for facility-level deviation predictions ($\hat{V}_{ij}(t)$) are comparable between the two models, perhaps the largest MSDE difference between the two models is observed for region-level deviation predictions ($\hat{U}_i(t)$). Ignoring the spatial correlation at the region-level leads to poor estimation of region-level deviations, as expected. Finally, we want to point out that M-VCSM is computationally more efficient than M-VCM, with run times of 20and 80 min (for $n = 49$ and $4-20$ facilities per region) on a computer with 2.4 GHz CPU, 8 GB RAM, respectively.

5 | DISCUSSION

An M-VCSM is proposed to study time-varying effects of multilevel risk factors on hospitalization rates in the US dialysis population. The proposed modelling approach takes into account the multilevel structure of the USRDS data with longitudinal hospitalization rates nested in dialysis facilities and dialysis facilities nested in geographic regions (HSAs). Time-varying regression effects of both region- and facility-level covariates are characterized as functions of time patients stay on dialysis. To address the challenges in estimation and inference in large data applications, introduced by the multilevel dependency structure with spatial correlations at the highest level of the hierarchy (region-level), dimension reduction is

achieved by a combination of tools from FDA and Bayesian modelling. Bayesian P-splines are utilized in expansions of the VCFs, while FPCA reduces the dimension in modelling of the multilevel time-dynamic stochastic deviations. Resulting modelling parameters are targeted within a mixed effects framework using MCMC. Simultaneous credible bands are used for inference on the multilevel VCFs. Applications of the proposed methodology to USRDS data identifies significant region- and facility-level risk factors for hospitalization rates. Even after adjusting for the multilevel risk factors, there is still significant spatiotemporal variation detected in the data across United States, which is flexibly modelled via multilevel time-dynamic stochastic deviations at both the region- and facility-levels. Spatial correlation at the region-level are induced through a CAR model on the region-level PC scores.

M-VCSM can be extended to model further remaining spatial correlation at the facility-level. However, note that the use of a distance- or a neighbourhood-based metric (such as the one used in modelling region-level spatial correlation through CAR) for capturing potential facility-level spatial correlation may not be appropriate, especially in large metropolitan regions where characteristics of a facility neighbourhood or its patients may not be correlated with those of another facility simply based on geographic distance. Rather, a metric reflecting the socio-economic status of the patients receiving treatment at the facility may be more accurate in capturing similarities in facility-level hospitalizations within the same region. Hence, more work is needed to identify viable approaches to model remaining potential facility-level correlations.

ACKNOWLEDGEMENTS

This study was supported by research grants from the National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK 092232—DS, DVN, EK, SB, CMR, and YL, and R01 DK 122767—CMR and DVN). The data reported here have been supplied by the United States Renal Data System (USRDS). The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy or interpretation of the US government. We thank a reviewer and associate editor for helpful suggestions which improved the paper and Qi Qian for assistance with manuscript revision.

SUPPORTING INFORMATION

Additional supporting information (appendices referred to in Sections 2 and 3) may be found online in the supporting information tab for this article. The R code and documentation for implementing the M-VCSM on simulated datasets are provided on Github (https://github.com/dsenturk/M-VCSM).

DATA AVAILABILITY STATEMENT

The release of the data used in this paper is governed by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) through the USRDS Coordinating Center. The data can be requested from the USRDS through a data use agreement.

ORCID

Danh V. Nguyen https://orcid.org/0000-0002-4025-8239
Damla Şentürk https://orcid.org/0000-0002-7619-5428

REFERENCES

Baladandayuthapani, V., Mallick, B. K., Young Hong, M., Lupton, J. R., Turner, N. D., & Carroll, R. J. (2008). Bayesian hierarchical spatially correlated functional data analysis with application to colon carcinogenesis. *Biometrics*, *64*(1), 64–73.

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2014). *Hierarchical modeling and analysis for spatial data*. Boca Raton, FL: CRC Press.

Cleveland, W. S., Grosse, E., & Shyu, W. M. (1991). Local regression models. In Chambers, J. M., & Hastie, T. J (Eds.), *Statistical models in s*. Pacific Grove: Wadsworth & Brooks, pp. 309–376.

Cox, D. D. (1993). An analysis of Bayesian inference for nonparametric regression. *The Annals of Statistics*, *21*, 903–923.

Crainiceanu, C. M., Ruppert, D., Carroll, R. J., Joshi, A., & Goodner, B. (2007). Spatially adaptive Bayesian penalized splines with heteroscedastic errors. *Journal of Computational and Graphical Statistics*, *16*(2), 265–288.

Crainiceanu, C. M., Staicu, A.-M., & Di, C.-Z. (2009). Generalized multilevel functional regression. *Journal of the American Statistical Association*, *104*(488), 1550–1561.

Cressie, N., & Wikle, C. K. (2011). *Statistics for spatio-temporal data*. Hoboken, NJ: John Wiley and Sons.

Di, C.-Z., Crainiceanu, C. M., Caffo, B. S., & Punjabi, N. M. (2009). Multilevel functional principal component analysis. *The Annals of Applied Statistics*, *3*(1), 458.

Hasenstab, K., Scheffler, A., Telesca, D., Sugar, C. A., Jeste, S., DiStefano, C., & Şentürk, D. (2017). A multi-dimensional functional principal components analysis of EEG data. *Biometrics*, *73*(3), 999–1009.

Hastie, T., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society. Series B (Methodological)*, *55*(4), 757–796.

Kalantar-Zadeh, K., Abbott, K. C., Salahudeen, A. K., Kilpatrick, R. D., & Horwich, T. B. (2005). Survival advantages of obesity in dialysis patients. *The American Journal of Clinical Nutrition*, *81*(3), 543–554.

Kind, A. myJ. H., & Buckingham, W. R. (2018). Making neighborhood-disadvantage metrics accessible—The neighborhood atlas. *The New England Journal of Medicine*, *378*(26), 2456.

Krivobokova, T., Kneib, T., & Claeskens, G. (2010). Simultaneous confidence bands for penalized spline estimators. *Journal of the American Statistical Association*, *105*(490), 852–863.

Kundu, M. G., Harezlak, J., & Randolph, T. W. (2016). Longitudinal functional models with structured penalties. *Statistical Modelling*, *16*(2), 114–139.
Lang, S., & Brezger, A. (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics*, *13*(1), 183–212.
Li, Y., Nguyen, D. V., Banerjee, S., Rhee, C. M., Kalantar-Zadeh, K., Kürüm, E., & Şentürk, D. (2021). Multilevel modeling of spatially nested functional data: Spatiotemporal patterns of hospitalization rates in the us dialysis population. *Statistics in Medicine*, *40*(17), 3937–3952.
Li, Y., Nguyen, D. V., Chen, Y., Rhee, C. M., Kalantar-Zadeh, K., & Şentürk, D. (2018). Modeling time-varying effects of multilevel risk factors of hospitalizations in patients on dialysis. *Statistics in Medicine*, *37*(30), 4707–4720.
Li, Y., Nguyen, D. V., Kürüm, E., Rhee, C. M., Chen, Y., Kalantar-Zadeh, K., & Şentürk, D. (2020). A multilevel mixed effects varying coefficient model with multilevel predictors and random effects for modeling hospitalization risk in patients on dialysis. *Biometrics*, *76*(3), 924–938.
Marra, G., & Wood, S. N. (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, *39*(1), 53–74.
Morris, J. S., & Carroll, R. J. (2006). Wavelet-based functional mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *68*(2), 179–199.
Morris, J. S., Vannucci, M., Brown, P. J., & Carroll, R. J. (2003). Wavelet-based nonparametric modeling of hierarchical functions in colon carcinogenesis. *Journal of the American Statistical Association*, *98*(463), 573–583.
Quick, H., Banerjee, S., & Carlin, B. P. (2013). Modeling temporal gradients in regionally aggregated california asthma hospitalization data. *The Annals of Applied Statistics*, *7*(1), 154–176.
Rice, J. A., & Silverman, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society: Series B (Methodological)*, *53*(1), 233–243.
Scheffler, A., Telesca, D., Li, Q., Sugar, C. A., Distefano, C., Jeste, S., & Şentürk, D. (2020). Hybrid principal components analysis for region-referenced longitudinal functional EEG data. *Biostatistics*, *21*(1), 139–157.
Scheipl, F., Staicu, A.-M., & Greven, S. (2015). Functional additive mixed models. *Journal of Computational and Graphical Statistics*, *24*(2), 477–501.
Serban, N. (2011). A space-time varying coefficient model: The equity of service accessibility. *The Annals of Applied Statistics*, *5*, 2024–2051.
Short, M., Carlin, B. P., & Bushhouse, S. (2002). Using hierarchical spatial models for cancer control planning in Minnesota (United States). *Cancer Causes & Control*, *13*(10), 903–916.
Staicu, A.-M., Crainiceanu, C. M., & Carroll, R. J. (2010). Fast methods for spatially correlated multilevel functional data. *Biostatistics*, *11*(2), 177–194.
USRDS (2020). United States Renal Data System 2020 Annual Data Report: "Epidemiology of Kidney Disease in the United States": National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD.
Wood, S. N. (2017). *Generalized additive models: An introduction with r*: CRC press.
Yao, F., Müller, H.-G., & Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, *100*(470), 577–590.
Zhang, L., Baladandayuthapani, V., Zhu, H., Baggerly, K. A., Majewski, T., Czerniak, B. A., & Morris, J. S. (2016). Functional CAR models for large spatially correlated functional datasets. *Journal of the American Statistical Association*, *111*(514), 772–786.
Zipunnikov, V., Caffo, B., Yousem, D. M., Davatzikos, C., Schwartz, B. S., & Crainiceanu, C. (2011). Multilevel functional principal component analysis for high-dimensional data. *Journal of Computational and Graphical Statistics*, *20*(4), 852–873.

How to cite this article: Li, Y., Nguyen, D. V., Kürüm, E., Rhee, C. M., Banerjee, S., & Şentürk, D. (2022). Multilevel varying coefficient spatiotemporal model. *Stat*, *11*(1), e438. https://doi.org/10.1002/sta4.438