Multivariate varying coefficient spatiotemporal model

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

[1]Department of Biostatistics, University of California, Los Angeles, CA, USA.
[2]Department of Medicine, University of California, Irvine, CA, USA.
[3]Department of Statistics, University of California, Riverside, CA, USA.
[4]Harold Simmons Center for Chronic Disease Research and Epidemiology, University of California Irvine, School of Medicine, Irvine, CA, US.

*Corresponding author(s). E-mail(s): dsenturk@ucla.edu;

Abstract

As of 2020, 807,920 individuals in the U.S. had end-stage kidney disease (ESKD) with about 70% of patients on dialysis, a life-sustaining treatment. Dialysis patients experience high mortality rates where frequent hospitalizations are a major contributor to morbidity and mortality. There is growing interest in identifying the risk factors for the correlated outcomes of hospitalization and mortality among dialysis patients across the U.S. Utilizing national data from the United States Renal Data System (USRDS), we propose a novel multivariate varying coefficient spatiotemporal model to study the time dynamic effects of risk factors (e.g., urbanicity and area deprivation index) on the multivariate outcome of hospitalization and mortality rates, as a function of time on dialysis. While capturing time-varying effects of risk factors on the mean, the proposed model also incorporates spatiotemporal patterns of the residuals for efficient inference. Estimation is based on the fusion of functional principal component analysis and Markov Chain Monte Carlo techniques, following basis expansions of the varying coefficient functions and multivariate Karhunen-Loéve expansion of region-specific random deviations. The finite sample performance of the proposed method is studied through extensive simulations. Novel applications to the USRDS data highlight significant risk factors of hospitalizations and mortality as well as characterizing time periods on dialysis and spatial locations across U.S. with elevated hospitalization and mortality risks.

Keywords: End-stage kidney disease · Multivariate functional data · Multivariate varying-coefficient models · Conditional autoregressive model · United States Renal Data System

1 Introduction

As of 2020, the number of individuals with prevalent end-stage kidney disease (ESKD) in the U.S. was 807,920, where about 70% of patients with ESKD were on the life-sustaining treatment of dialysis [1]. Patients on dialysis have a high burden of complex comorbid conditions, and the overall mortality rates are 10-20 times higher than the general population [1, 2]. Moreover, they experience frequent hospitalizations (around 2 per person-year (PPY)), which is a major source of morbidity and mortality [1]. Hence, modeling of the correlated multivariate outcome of hospitalization and survival rates in the U.S. dialysis cohort is of interest for more targeted patient monitoring.

Prior studies have shown important spatiotemporal variations in both hospitalization and survival rates. ESKD patients typically remain on dialysis for the duration of their lives or until kidney transplantation and their needs change as they stay longer on dialysis. Hence an important time index along which rates have been shown to vary is time since initiation of dialysis, where elevated hospitalization and mortality rates have been reported within the first 3 to 6 months of initiation of dialysis [3–8]. In addition to time-dynamic variation, there has been significant spatial variation reported in hospitalization and mortality rates across U.S., with the highest rates observed in health service areas (HSAs) from the North East and lowest in HSAs from the West [1, 9–11]. Hence, there is a compelling need for multivariate modeling of hospitalization and mortality rates to identify time-dynamic/time-varying effects of region-specific risk factors (e.g., urbanicity and area deprivation index) that contribute to differences in longitudinal hospitalization and mortality risks observed across the U.S., while adjusting for spatiotemporal correlations in the data.

Multivariate varying coefficient modeling is an effective tool in studying time-varying effects of covariates on multivariate longitudinal response. Estimation typically relies on effective decompositions of the varying coefficient functions (VCFs) via local polynomial [15–17, 19], or basis expansions [20–24], similar to the univariate case. Literature on multivariate varying coefficient models that include space-time varying coefficients or incorporate spatiotemporal correlations in the data are limited. Gelfand et al. [25] considered multivariate space-time varying coefficient models, while Congdon et al. [26] and Cheng et al. [27] incorporated spatial correlations in the residuals. However, all three approaches model time as discrete. In our application of multivariate modeling of hospitalization and mortality risk across HSAs in U.S., space is modeled as discrete, since HSAs (regions with relatively self-contained infrastructure for the provision of hospital care) across U.S. are fixed

and time is modeled as continuous, since we are interested in drawing inference along the continuous time index of duration on dialysis.

While there is extensive literature on multivariate spatiotemporal modeling where space and time are both modeled as discrete, the literature on discrete space and continuous time modeling is limited. For multivariate spatiotemporal models with discrete space and continuous time covariances, Zhang et al. [12] proposed a separable structure for the multivariate space-time covariance, while Hepler et al. [13] and Baer et al. [14] considered nonseparable space-time covariances with quite restricted linear or user-defined parametric forms in the temporal modeling. Recently, Qian et al. [29] proposed a multivariate spatiotemporal functional principal analysis model (MST-FPCA) with nonseparable space-time covariances, which couples dimension reduction in modeling nonparametric time trends through data-driven lower dimensional multivariate eigenfunctions with modeling of spatial correlations via a parametric conditional autoregressive (CAR) correlation structure. However, all these works do not include time-varying effects of covariates.

To quantify the time-varying effects of risk factors on the correlated multivariate outcome of hospitalization and mortality rates in the dialysis cohort while incorporating spatiotemporal correlations in the data, where space is taken to be discrete and time is modeled as continuous, we propose a novel multivariate varying coefficient spatiotemporal model (MV-VCSTM). In addition to modeling time-varying effects of covariates on the multivariate outcome, MV-VCSTM also includes region-specific time-varying multivariate random deviations in the model to account for the remaining multivariate spatiotemporal correlations in the data that are not explained by covariates. Similar to developments in Qian et al. [29], time-varying trends of the random deviations are modeled nonparametrically based on data-driven lower dimensional multivariate eigenfunction bases. Spatial correlations are induced through a CAR model assumed among region-specific scores, leading to a nonseparable structure on the multivariate space-time covariance. Estimation and inference in the proposed MV-VCSTM, simultaneously targeting multivariate time-varying coefficient functions and a multivariate spatiotemporal process poses a major computational challenge, especially in higher dimensional data applications. To achieve computational efficiency, the proposed estimation and inference relies on multiple computational savings. First, while the multivariate varying coefficient functions are expanded using thin-plate spline bases, region-specific random deviations are expanded via data-driven multivariate eigenfunctions, which significantly reduces the number of parameters estimated associated with basis expansions. The multivariate eigenfunctions used in the expansion of the region-specific random deviations are derived based on only univariate functional principal component analysis (FPCA), avoiding smoothing over higher than 2-dimensional covariance surfaces, leading to major computational savings as the dimension of the multivariate outcome increases. The estimation of the multivariate eigenfunctions relies on univariate eigenfunctions and the estimated spatial correlation structure among the region-specific scores

obtained from univariate FPCA. Estimation of this spatial correlation structure requires multivariate CAR (MCAR) modeling, which also poses challenges in higher dimensions. To handle this challenge, the MCAR structure is modeled via a linear combination of independent latent Gaussian processes with a CAR correlation structure, effectively lowering the dimension in estimation. Once the multivariate eigenfunctions for bases expansions of the region-specific random deviations are obtained, we target the multivariate time-varying coefficient functions and region-specific deviation scores as well as the needed spatial parameters via Markov Chain Monte Carlo (MCMC) in a Bayesian hierarchical modeling framework.

The paper is organized as follows. Section 2 introduces the proposed MV-VCSTM, and outlines the proposed estimation and inference procedures. The application to modeling of hospitalization and mortality rates in the U.S. dialysis cohort is based on data from the national USRDS database, containing longitudinal hospitalization and mortality rates from 367 HSAs across U.S. and is outlined in Section 3. Simulations to study the finite sample performance of the proposed methodology, including comparisons with multivariate varying coefficient models that do not incorporate spatiotemporal dependencies observed in the residuals are presented in Section 4, followed by a discussion given in Section 5.

2 Proposed Multivariate Varying Coefficient Spatiotemporal Model (MV-VCSTM)

2.1 Model specification

Let $i = 1, \ldots, n$ index regions, $k = 1, \ldots, T$ index time (months) after transition to dialysis and $j = 1, \ldots, J$ $(J \geq 2)$ index the dimensions of the J-dimensional multivariate outcome vector, $\boldsymbol{Y}_i(t_k) = \{Y_i^{(1)}(t_k), \ldots, Y_i^{(J)}(t_k)\}^\top$. In our data application, the multivariate outcome considered contains region-specific monthly hospitalization and mortality rates (i.e. $J = 2$). The region-specific rates are obtained as averages of dialysis facility-specific rates within regions (i.e., HSAs). Rates at the dialysis facility level are defined as the ratio of the total number of patient hospitalizations or deaths to the total patient follow-up time for that specific facility at month k. Even though they are obtained monthly, rates are multiplied by 12 to allow interpretations as per person-year (PPY), consistent with annual national reporting of USRDS. Finally, region-specific rates are analyzed for a total of 24 months (i.e., 2 year follow-up) after transitioning to dialysis. Note that we opt to model the multivariate response as continuous data similar to the works of [9, 28, 29] and is amenable for functional data modeling.

To study the time-varying effects of risk factors on the multivariate response (of hospitalization and mortality rates), the proposed MV-VCSTM includes P covariates $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{iP})^\mathsf{T}$ and P time-varying coefficient functions

$\boldsymbol{\beta}^{(j)}(t) = \{\beta_1^{(j)}(t), \ldots, \beta_P^{(j)}(t)\}^\mathsf{T}$ associated with each covariate,

$$Y_i^{(j)}(t) = \boldsymbol{X}_i^\mathsf{T}\boldsymbol{\beta}^{(j)}(t) + U_i^{(j)}(t) + \epsilon_i^{(j)}(t), \quad \text{for} \quad j = 1, \ldots, J. \tag{1}$$

In (1), the remaining spatiotemporal variations in the data after adjusting for time-varying effects of risk factors are captured by the latent region-specific deviations, denoted by $U_i^{(j)}(t)$, while $\boldsymbol{\epsilon}_i(t) = \{\epsilon_i^{(1)}(t), \ldots, \epsilon_i^{(J)}(t)\}^\mathsf{T}$ denotes the i.i.d. measurement error with different variances along different dimensions of the multivariate outcome ($\epsilon_i^{(j)}(t) \sim_{ind} N(0, \sigma_j^2)$). Note that in the formulation in (1), the covariates are taken to be time-invariant, to mimic our data application. However, the developments can be extended easily to include time-varying covariates by changing the design matrix.

The latent region-specific deviations $\boldsymbol{U}_i(t) = \{U_i^{(1)}(t), \ldots, U_i^{(J)}(t)\}^\mathsf{T}$ are further decomposed via

$$\boldsymbol{U}_i(t) = \sum_{\ell=1}^{\infty} \rho_{i\ell}\boldsymbol{\psi}_\ell(t),$$

where $\rho_{i\ell}$ denotes the region-specific principal component (PC) scores and $\boldsymbol{\psi}_\ell(t) = \{\psi_\ell^{(1)}(t), \ldots, \psi_\ell^{(J)}(t)\}^\mathsf{T}$ denote multivariate eigenfunctions which form an orthonormal system,

$$\langle\langle \boldsymbol{\psi}_\ell, \boldsymbol{\psi}_{\ell'} \rangle\rangle := \sum_{j=1}^{J} \langle \psi_\ell^{(j)}, \psi_{\ell'}^{(j)} \rangle_2 = \sum_{j=1}^{J} \int_{\mathcal{T}} \psi_\ell^{(j)}(t)\psi_{\ell'}^{(j)}(t)\mathrm{d}t = \delta_{\ell\ell'},$$

with $\delta_{\ell\ell'} = 1$ if $\ell = \ell'$, and $\delta_{\ell\ell'} = 0$ otherwise. Under the classic multivariate FPCA framework [30, 32], the multivariate PC scores $\{\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\mathsf{T} : \ell = 1, 2, \ldots\}$ are assumed to be uncorrelated, with zero means and $\mathrm{Var}(\boldsymbol{\rho}_{i\ell}) = \lambda_\ell$, where λ_ℓ's denote the eigenvalues. Similar to univariate FPCA, the multivariate eigenfunctions describe directions of leading modes of variation in the different dimensions of the multivariate functional response, while the eigenvalues quantify the amount of variation explained along the identified modes of variation. In most applications the expansion can be truncated to include only a few number (denoted below with L) of eigencomponents explaining most of the variation in the data, leading to effective dimension reduction of the high-dimensional data.

To capture the remaining spatiotemporal correlation among regions, that is not explained by the time varying effects of regional risk factors, we induce a parametric CAR structure on the region-specific PC scores $\rho_{i\ell}$, deviating from the classic FPCA framework. Note that there are no repetitions of data over space, necessitating the adoption of a parametric spatial modeling approach. More specifically, let the $n \times n$ adjacency matrix $W = \{w_{ii'}\}$ describe the neighborhood structure of the regions, where $w_{ii'} = 1$ if regions i and i' ($i \neq i'$) are neighbors, denoted by $i \sim i'$, and $w_{ii'} = 0$ otherwise. By convention, the diagonal elements of W are set to zero. Further, let D denote the diagonal matrix with elements $d_i = \sum_{i \sim i'} w_{ii'}$, representing the total

number of neighbors of regions. Then a Markov Random Field [33] (Chapter 4.2) specifies the full conditional distribution for the ℓth PC score $\rho_{i\ell}$ for region i as a weighted average of the ℓth PC scores from neighboring regions: $\rho_{i\ell} \mid \{\rho_{i'\ell}\}_{i \neq i'} \sim N(\nu \sum_{i' \sim i} w_{ii'}\rho_{i'\ell}/d_i, \alpha_\ell/d_i)$, where α_ℓ denotes the variance component and ν the spatial correlation parameter. Through Brook's lemma [42], the joint distribution of the ℓth PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\mathsf{T}$ takes the form $\boldsymbol{\rho}_\ell \sim N\left(\mathbf{0}, \alpha_\ell(D - \nu W)^{-1}\right)$, where the scalar ν is constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues of the matrix $D^{-1/2}WD^{-1/2}$ for the precision matrix $(D - \nu W)/\alpha_\ell$ to be positive definite. Note that CAR models are commonly used in modeling latent components of hierarchical models to capture possible correlations among the observations of interest, such as the spatial correlations within a neighborhood structure. In our MV-VCSTM modeling, the assumed CAR structure among the region-specific scores facilitates borrowing of spatial information across neighboring regions, acting similar to a spatial smoother.

In the proposed modeling, the latent process $\boldsymbol{U}_i(t)$ capture remaining spatiotemporal variations in the outcome (after adjusting for time-varying effects of covariates). Since the multivariate eigenfunctions vary across time, the spatial correlation induced also varies across time (through the product of the region-specific scores and the eigenfunctions) and since the eigenfunctions are allowed to vary across different eigencomponents and different dimensions of $\boldsymbol{U}_i(t)$, the proposed modeling is able to induce different spatiotemporal correlations across ℓ and j (via the linear combination of different scores multiplying different multivariate eigenfunctions). Note also that the proposed modeling implies a nonseparable space-time correlation structure on $\boldsymbol{U}_i(t)$, since the covariance of $\boldsymbol{U}_i(t)$ is a weighted linear combination (weighted by α_ℓ) of the spatial correlation induced across regions through the adjacency (W) and neighborhood matrices (D) multiplying different eigenfunctions. Hence the assumed spatiotemporal structure is flexible and does not make restrictive assumptions such as separability, as is common in spatiotemporal modeling.

Finally, the constant measurement error variance assumed across time does not imply that the variance of the region-specific deviations (after adjusting for time-varying effects of covariates) or the variance of the outcome itself is not allowed to change across time. After accounting for the time-varying effects of covariates, the remaining spatiotemporal variation of the multivariate outcome is modeled via multivariate FPCA with spatially correlated region-specific scores. Hence, the leading directions of temporal variation in the multivariate outcomes (after adjusting for time-varying effects of covariates) are modeled via the multivariate eigenfunctions. A measurement error component with a constant variance in time is typical in multivariate FPCA [30, 32], since time-varying variance structure on the measurement error would not be identifiable after the nonparametric modeling of the temporal variation via multivariate eigenfunctions.

2.2 Estimation and inference

The proposed algorithm begins by fitting a multivariate varying coefficient model to the observed data under the working independence assumption (Step 1). The multivariate residuals obtained from this initial fit, denoted by $\boldsymbol{E}_i(t) = \{E_i^{(1)}(t), \ldots, E_i^{(J)}(t)\}^\mathsf{T}$, are used as a proxy of the remaining spatiotemporal deviations $\boldsymbol{U}_i(t)$ in the data. Steps 2-4 target the lower dimensional data-driven multivariate eigenfunction bases ($\boldsymbol{\psi}_\ell(t)$) to be used in the expansions of the spatiotemporal deviations $\boldsymbol{U}_i(t)$ in the data. Multivariate eigenfunctions are targeted using univariate eigenfunctions of the different dimensions of the multivariate residual $\boldsymbol{E}_i(t)$ and the between eigencomponent dependencies of the univariate PC scores $\widehat{\boldsymbol{\xi}}$ of $\boldsymbol{E}_i(t)$ (reflecting the dependency structure between the different dimensions of the multivariate outcome) (Steps 2-3). Hence, the multivariate eigenfunctions ($\boldsymbol{\psi}_\ell(t)$) are derived based only on univariate FPCA employed on the different dimensions of $\boldsymbol{E}_i(t)$. Multivariate FPCA involves smoothing over four or more dimensional covariance surfaces, which can be quite cumbersome in terms of computational costs. By comparison, the proposed MV-VCSTM relies only on univariate FPCA, which involves smoothing over only two-dimensional covariance surfaces (even for $J > 2$), leading to greater computational gains for higher dimensional outcome (i.e. as J increases).

Note that the multivariate conditional autoregressive (MCAR) modeling needed for targeting the between eigencomponents covariance of $\widehat{\boldsymbol{\xi}}$ is fitted via a linear combination of independent latent Gaussian processes with a CAR correlation structure, achieving further dimension reduction and computational savings (Step 3). Once the data-driven multivariate eigenfunction bases for $\boldsymbol{U}_i(t)$ are estimated and the multivariate varying coefficient functions (VCFs) $\boldsymbol{\beta}^{(j)}(t)$ are expanded onto thin-plate spline bases, spline coefficients needed for estimation of the multivariate VCFs, multivariate PC scores $\rho_{i\ell}$ and variance components (α_ℓ, ν and σ^2) are targeted via a Bayesian hierarchical model (Step 5), followed by inference based on posterior distributions of the penalized spline coefficients and the multivariate PC scores.

Additional computational gains achieved in the proposed algorithm are based on the choice of multivariate eigenfunctions as a basis set for the expansion of $\boldsymbol{U}_i(t)$. Since they are estimated based on the data, a lower number of them are enough typically to explain leading directions of variation in the data (compared to a pre-determined basis set), leading to a parsimonious model and additional computational savings in Step 5 of the proposed algorithm. Moreover, because eigenfunctions are estimated from the data, they carry additional interpretations as dominant modes of variation in time. While the proposed estimation procedure is outlined in the algorithm below, key features in each step are detailed further in this section. The R codes and a tutorial for implementing the MV-VCSTM algorithm are made publicly available on Github (https://github.com/dsenturk/MV-VCSTM).

Estimation Algorithm

Step 1: Fit a multivariate varying coefficient model to the observed data under the working independence assumption and obtain the multivariate residuals $\boldsymbol{E}_i(t) = \{E_i^{(1)}(t), \ldots, E_i^{(J)}(t)\}^\mathsf{T}$.

Step 2: Employ univariate FPCA on $E_i^{(j)}(t)$ to target univariate eigenfunctions and the univariate PC score vector $\widehat{\boldsymbol{\xi}}$.

Step 3: Decompose dependencies among $\widehat{\boldsymbol{\xi}}$ into a between eigencomponent Σ_b and a within eigencomponent Σ_w (between region) variation via an MCAR correlation structure.

Step 4: Target the multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t)$ using the estimated univariate eigenfunctions and eigenvectors of the estimated Σ_b.

Step 5: Expanding the stochastic deviations $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions and the VCFs $\boldsymbol{\beta}^{(j)}(t)$ on thin-plate spline bases, target the VCFs $\boldsymbol{\beta}^{(j)}(t)$, multivariate PC scores ($\rho_{i\ell}$) and the variance components (α_ℓ, ν, and σ^2) within a Bayesian hierarchical modeling framework via MCMC.

Step 1: A multivariate varying-coefficient model is fit to the observed data under the working independence assumption (referred to as MV-VCM): $Y_i^{(j)}(t) = \boldsymbol{X}_i^\mathsf{T}\boldsymbol{\beta}^{(j)}(t) + \epsilon_i^{(j)}(t)$, where $\epsilon_i^{(j)}(t)$ is assumed to be i.i.d. MV-VCM is implemented using the built-in `gam` function in the R package `mgcv`, where thin-plate splines are used in estimation of the VCFs and the smoothing parameters are selected using restricted maximum likelihood (REML). The multivariate residuals from this initial fit, denoted by $\boldsymbol{E}_i(t) = \{E_i^{(1)}(t), \ldots, E_i^{(J)}(t)\}^\mathsf{T}$, capture the remaining spatiotemporal variation in the data that is not explained by the time varying effects of risk factors and are used as a proxy for the region-specific deviations $\boldsymbol{U}_i(t)$.

Step 2: Univariate FPCA is employed on $E_i^{(j)}(t)$ separately along each dimension of the multivariate residuals. More specifically, in jth dimension, the empirical covariance is calculated as $\widehat{G}^{(j)}(t,t') = \sum_{i=1}^n E_i^{(j)}(t)E_i^{(j)}(t')/n$. Next, a two-dimensional penalized smoothing is applied to the off-diagonal elements of $\widehat{G}^{(j)}(t,t')$ via the built-in `gam` function in the R package `mgcv`. The diagonal entries of $\widehat{G}^{(j)}(t,t')$ are removed in smoothing, since they are prone to measurement error. The two-dimensional penalized smoothing yields the univariate eigenfunction $\{\widehat{\phi}_m^{(j)}(t) : m = 1, \ldots, M_j\}$ and PC score estimates $\{\widehat{\xi}_{im}^{(j)} : i = 1, \ldots, n;\ m = 1, \ldots, M_j\}$. The smoothing parameters are selected by REML and the number of eigencomponents M_j is selected using the fraction of variance explained (FVE), where FVE $> 99\%$ is used in applications to retain enough information at this initial step of the algorithm.

Step 3: Next, the univariate PC score vector $\widehat{\boldsymbol{\xi}}_i^\mathsf{T} = (\widehat{\xi}_{i1}^{(1)}, \ldots, \widehat{\xi}_{iM_1}^{(1)}, \ldots, \widehat{\xi}_{i1}^{(J)}, \ldots, \widehat{\xi}_{iM_J}^{(J)})$ is modeled using an MCAR correlation

structure. Let $\Xi = [\widehat{\boldsymbol{\xi}}_1, \ldots, \widehat{\boldsymbol{\xi}}_n]^\mathsf{T}$ denote the $n \times M^+$ score matrix, combining the score vectors from all subjects, where $M^+ = M_1 + M_2 + \ldots + M_J$ denotes the total number of eigencomponents retained across all J dimensions of $\boldsymbol{E}_i(t)$. Stacking the columns of Ξ leads to the $nM^+ \times 1$ score vector $\widehat{\boldsymbol{\xi}} = \text{vec}(\Xi)$ that is modeled as a linear combination of M^+ independent latent Gaussian processes with a CAR correlation structure [34],

$$\underset{nM^+\times 1}{\widehat{\boldsymbol{\xi}}} = \underset{nM^+\times 1}{\boldsymbol{\xi}} + \underset{nM^+\times 1}{\mathbf{e}} = (\underset{M^+\times M^+}{A} \otimes \underset{n\times n}{I}) \underset{nM^+\times 1}{\mathbf{f}} + \underset{nM^+\times 1}{\mathbf{e}}. \tag{2}$$

In (2), $\mathbf{f} = (\mathbf{f}_1^\mathsf{T}, \ldots, \mathbf{f}_{M^+}^\mathsf{T})^\mathsf{T}$ denotes the vector of stacked $n \times 1$ latent Gaussian processes $\mathbf{f}_\ell \sim N(\mathbf{0}, (D - \nu W)^{-1})$, $\ell = 1, \ldots, M^+$, where D and W are as defined in Section 2.1 and ν denotes the common spatial smoothing parameter, $A = \{a_{\ell\ell'}\}$, $1 \leq \ell' \leq \ell \leq M^+$, denotes an $M^+ \times M^+$ full rank lower triangular matrix and $\mathbf{e} \sim N(\mathbf{0}, \tau^2 I_{nM^+})$ denotes the vector of measurement errors, assumed to be uncorrelated with $\boldsymbol{\xi}$. Note that the distribution of $\mathbf{f}$ can be given as $N\left(\mathbf{0}, I_{M^+} \otimes (D - \nu W)^{-1}\right)$, leading to the decomposition of the covariance Σ of $\boldsymbol{\xi}$ into between eigencomponent ($\Sigma_b \equiv AA^\mathsf{T}$) and within eigencomponent ($\Sigma_w \equiv (D - \nu W)^{-1}$) variation

$$\begin{aligned}\underset{nM^+\times nM^+}{\Sigma} &= (\underset{M^+\times M^+}{A} \otimes \underset{n\times n}{I}) \text{Bdiag}\{\underset{n\times n}{(D - \nu W)^{-1}}, \ldots, (D - \nu W)^{-1}\}(A^\mathsf{T} \otimes I) \\ &= \underset{M^+\times M^+}{AA^\mathsf{T}} \otimes \underset{n\times n}{(D - \nu W)^{-1}} \equiv \underset{M^+\times M^+}{\Sigma_b} \otimes \underset{n\times n}{\Sigma_w},\end{aligned}$$

where Bdiag$(\cdot)$ denotes a block-diagonal matrix and $\boldsymbol{\xi}$ is MCAR (ν, Σ_b). While the between eigencomponent variation captures the dependency between the eigencomponents from univariate expansions, the within eigencomponent (between region) variation captures the spatial dependency among the regional units. Note that representation of $\boldsymbol{\xi}$ as a linear combination of independent spatial Gaussian processes modeled with a CAR correlation structure, via the lower triangle matrix A, allows for easy implementation of the MCAR model in WinBUGS, without the need for operations with large covariance matrices and without the need to check for positive-definiteness of covariance matrices in each iteration of the algorithm.

The parameters of the MCAR model are targeted via MCMC. Elementwise priors are imposed on the lower-triangular matrix A: $a_{\ell\ell} \sim$ Log-normal$(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}})$ and $a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}})$ for $1 \leq \ell' < \ell \leq M^+$. In addition, an Inverse Gamma (IG) (a_{τ^2}, b_{τ^2}) prior is imposed on the measurement error variance τ^2, and a Uniform prior is used for the spatial correlation parameter ν with the parameters constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues of the matrix $D^{-1/2}WD^{-1/2}$ (denoted by a_ν, b_ν, respectively). The posterior distribution we seek can be expressed as

$$\pi(a_{\ell\ell}, a_{\ell\ell'}, \nu, \tau^2 \mid \widehat{\boldsymbol{\xi}}) \propto N(\widehat{\boldsymbol{\xi}} \mid \boldsymbol{\xi}, \tau^2) \times N\left(\boldsymbol{\xi} \mid \mathbf{0}, AA^\mathsf{T} \otimes (D - \nu W)^{-1}\right)$$

$$\times \prod_{\ell=1}^{M^+} \text{Log-normal}(a_{\ell\ell} \mid \mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}}) \times \prod_{\ell=1}^{M^+} \prod_{\ell'=1}^{\ell-1} N(a_{\ell\ell'} \mid \mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}})$$
$$\times \text{Unif}(\nu \mid a_\nu, b_\nu) \times \text{IG}(\tau^2 \mid a_{\tau^2}, b_{\tau^2}).$$

Model parameters are sampled from the posterior distributions using MCMC with Gibbs sampling and random walk metropolis as implemented in WinBUGS.

Step 4: The multivariate eigenfunctions are targeted using the univariate eigenfunctions $\widehat{\phi}_m^{(j)}(t)$ of $E_i^{(j)}(t)$ estimated in Step 2 and the eigenvectors $\widehat{\mathbf{c}}_\ell$, $\ell = 1, \ldots, M^+$, of the between eigencomponent variation $\widehat{\Sigma}_b$ targeted in Step 3 as

$$\widehat{\psi}_\ell^{(j)}(t) = \sum_{m=1}^{M_j} [\widehat{\mathbf{c}}_\ell]_m^{(j)} \widehat{\phi}_m^{(j)}(t), \quad \ell = 1, \ldots, M^+, \quad j = 1, \ldots, J. \tag{3}$$

For justification of (3), we refer the reader to Web Appendix A of [29], where the authors show that the variance components α_ℓ's are also the eigenvalues of the between-eigencomponent covariance matrix Σ_b in the context of a spatiotemporal model. This link enables the targeting of each dimension of the multivariate eigenfunction $\widehat{\psi}_\ell^{(j)}(t)$ as a linear combination of the estimated univariate eigenfunctions $\widehat{\phi}_m^{(j)}(t)$ for that dimension with weights equal to the corresponding entries in the eigenvector of the estimated between-eigencomponent covariance matrix Σ_b of the univariate scores. More specifically, the weights equal the mth entry of $[\widehat{\mathbf{c}}_\ell]^{(j)} \in \mathbb{R}^{M_j}$, the jth block of the ℓth eigenvector $\widehat{\mathbf{c}}_\ell$ of $\widehat{\Sigma}_b$ (equal to the posterior mean of Σ_b), where they enable incorporation of the dependence among different eigencomponents into the multivariate eigenfunctions.

Step 5: Expanding the stochastic deviations $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions from Step 4 and the VCFs $\boldsymbol{\beta}^{(j)}(t)$ on thin-plate spline bases, the VCFs, multivariate PC scores ($\rho_{i\ell}$) and the variance components (α_ℓ, ν, and σ^2) are targeted via a hierarchical Bayesian modeling framework. Note that once the multivariate eigenfunctions are targeted in Steps 1-4, the parametric CAR assumption on the multivariate PC scores is pivotal in allowing likelihood-based inference via a Bayesian framework.

Consider the expansion of the VCFs $\boldsymbol{\beta}^{(j)}(t) = \{\beta_1^{(j)}(t), \ldots, \beta_P^{(j)}(t)\}^\mathsf{T}$ onto penalized low-rank thin-plate spline bases:

$$\beta_p^{(j)}(t) \approx u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} v_{pq}^{(j)} |t - \kappa_q|^3, \quad \text{for} \quad p = 1, \ldots, P, \tag{4}$$

where $\kappa_1 < \kappa_2 < \ldots < \kappa_Q$ denote the fixed knots and $u_{p0}^{(j)}$, $u_{p1}^{(j)}$ and $\boldsymbol{v}_p^{(j)} = (v_{p1}^{(j)}, \ldots, v_{pQ}^{(j)})^\mathsf{T}$ denote the expansion coefficients. The number of knots

Q utilized in applications is typically selected between $3-20$ to ensure the desired flexibility [35, 36], and κ_q is taken to be the sample quantile of the time points corresponding to probability $q/(Q+1)$.

The P-splines, including a penalty matrix Ω_Q with (q', q)th entry $|\kappa_{q'} - \kappa_q|^3$, penalizing the coefficients of $|t-\kappa_q|^3$, can be fitted using the transformed model

$$\beta_p^{(j)}(t_k) \approx u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq}, \quad \text{for} \quad p = 1, \ldots, P, \tag{5}$$

where $\widetilde{\boldsymbol{v}}_p^{(j)} \equiv (\widetilde{v}_{p1}^{(j)}, \ldots, \widetilde{v}_{pQ}^{(j)})^{\mathsf{T}} = \Omega_Q^{1/2} \boldsymbol{v}_p^{(j)}$ with $\text{cov}(\widetilde{\boldsymbol{v}}_p^{(j)}) = \sigma^2_{\widetilde{v}_{pj}} I_Q$ and z_{kq} denotes the (k,q)th entry of the transformation $Z = Z_Q \Omega_Q^{-1/2}$ of the design matrix Z_Q with kth row equaling $\{|t_k - \kappa_1|^3, \ldots, |t_k - \kappa_Q|^3\}$ [36].

The hierarchical model combining expansions of the VCFs on thin-plate spline bases and expansion of $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions $\widehat{\psi}_\ell^{(j)}(t)$ can be given as

$$Y_i^{(j)}(t_k) = \sum_{p=1}^{P} X_{ip} \left(u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq} \right) + \sum_{\ell=1}^{M^+} \rho_{i\ell} \widehat{\psi}_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k),$$

for $j = 1, \ldots, J$ and $k = 1, \ldots, T$, where $\widetilde{v}_{pq}^{(j)} \sim N(0, \sigma^2_{\widetilde{v}_{pj}})$, $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^{\mathsf{T}} \sim N(\mathbf{0}, \alpha_\ell (D - \nu W)^{-1})$ and $\epsilon_i^{(j)}(t_k) \sim N(0, \sigma_j^2)$. Using normal priors for $u_{p0}^{(j)} \sim N(0, \sigma^2_{u_{p0j}})$ and $u_{p1}^{(j)} \sim N(0, \sigma^2_{u_{p1j}})$, inverse Gamma priors for the variance components $\sigma^2_{\widetilde{v}_{pj}} \sim \text{IG}(a_{\sigma^2_{\widetilde{v}_{pj}}}, b_{\sigma^2_{\widetilde{v}_{pj}}})$, $\sigma_j^2 \sim \text{IG}(a_{\sigma_j^2}, b_{\sigma_j^2})$, $\alpha_\ell \sim \text{IG}(a_{\alpha_\ell}, b_{\alpha_\ell})$ and a Uniform prior for $\nu \sim \text{Unif}(a_\nu, b_\nu)$, the posterior distribution we seek can then be expressed as

$$\begin{aligned}
&\pi(u_{p0}^{(j)}, u_{p1}^{(j)}, \widetilde{v}_p^{(j)}, \boldsymbol{\rho}_\ell, \nu, \alpha_\ell, \sigma_j^2 \mid Y_i^{(j)}(t), \boldsymbol{X}_i, Z, \widehat{\psi}_\ell^{(j)}(t), M^+) \\
&\propto \prod_{i=1}^{n} \prod_{k=1}^{T} \prod_{j=1}^{J} N\left(\boldsymbol{Y}_i^{(j)}(t_k) \mid \{ \sum_{p=1}^{P} X_{ip} (u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq}) + \sum_{\ell=1}^{M^+} \rho_{i\ell} \psi_\ell^{(j)}(t_k) \}, \sigma_j^2 \right) \\
&\times \prod_{j=1}^{J} \text{IG}(\sigma_j^2 \mid a_{\sigma_j^2}, b_{\sigma_j^2}) \times \prod_{j=1}^{J} \prod_{p=1}^{P} N\left(u_{p0}^{(j)} \mid 0, \sigma^2_{u_{p0j}} \right) \times \prod_{j=1}^{J} \prod_{p=1}^{P} N\left(u_{p1j}^{(j)} \mid 0, \sigma^2_{u_{p1j}} \right) \\
&\times \prod_{j=1}^{J} \prod_{p=1}^{P} \prod_{q=1}^{Q} N\left(\widetilde{v}_{pq}^{(j)} \mid 0, \sigma^2_{\widetilde{v}_{pj}} \right) \times \prod_{j=1}^{J} \prod_{p=1}^{P} \prod_{q=1}^{Q} \text{IG}\left(\sigma^2_{\widetilde{v}_{pj}} \mid a_{\sigma^2_{\widetilde{v}_{pj}}}, b_{\sigma^2_{\widetilde{v}_{pj}}} \right) \\
&\times \prod_{\ell=1}^{M^+} N\left(\boldsymbol{\rho}_\ell \mid \mathbf{0}, \alpha_\ell (\text{D} - \nu \text{W})^{-1} \right) \times \prod_{\ell=1}^{M^+} \text{IG}(\alpha_\ell \mid a_{\alpha_\ell}, b_{\alpha_\ell}) \times \text{Unif}(\nu \mid a_\nu, b_\nu).
\end{aligned}$$

Model parameters are sampled from the posterior distributions using MCMC as implemented in WinBUGS. Since M^+ is the total number of univariate eigencomponents retained in the FPCA decompositions of $E_i^{(j)}(t)$

across the J dimensions, the total number (L) of multivariate eigencomponents retained in expansions of $\boldsymbol{U}_i(t)$ could be chosen to be smaller than M^+ in applications where M^+ may be large. Following [29], we recommend the use of the estimated variance components $\widehat{\alpha}_\ell$, $\ell = 1, \ldots, M^+$ in formulating the fraction of variance explained (FVE) in the choice of L, since they are established as the eigenvalues of the between eigencomponent covariance of the univariate PC scores (see Web Appendix A of [29]).

VCFs $\beta_p^{(j)}(t)$ are targeted by mean of the posterior draws for $u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq}$, based on posterior draws for $u_{p0}^{(j)}$, $u_{p1}^{(j)}$ and $\widetilde{\boldsymbol{v}}_p^{(j)}$. Simultaneous credible bands for the VCFs are obtained along each dimension j as follows. Let $f(t)$ denote a single VCF observed at time points t_k for $k = 1, 2, \ldots, T$ in one dimension. Further let $\widehat{f}(t)$ and $\mathrm{SD}\{f(t)\}$ denote the mean and standard deviation of $f(t)$ based on a total of B MCMC samples $f^{(b)}(t)$ and let c_δ denote the $(1-\delta)$ sample quantile of $\max_{k=1,\ldots,T} \mid \{f^{(b)}(t_k) - \widehat{f}(t_k)\}/\mathrm{SD}\{f(t_k)\} \mid$, $b = 1, 2, \ldots, B$. Then the $(1-\delta)$ simultaneous credible band (CB) for $f(t)$ in that dimension is given by $[\widehat{f}(t_k) \pm c_\delta \mathrm{SD}\{f(t_k)\}]$ [31]. In addition, using the estimated multivariate eigenfunctions from Step 4 and taking $\widehat{\boldsymbol{\beta}}^{(j)}(t)$ and $\widehat{\rho}_{i\ell}$ as the mean of the posterior samples on the VCFs and the multivariate eigenscores, respectively, region-specific prediction of the multivariate outcome is targeted by

$$\widehat{Y}_i^{(j)}(t) = \boldsymbol{X}_i^{\mathsf{T}} \widehat{\boldsymbol{\beta}}^{(j)}(t) + \sum_{\ell=1}^{L} \widehat{\rho}_{i\ell} \widehat{\psi}_\ell^{(j)}(t_k) \quad \text{for} \quad j = 1, \ldots, J.$$

When different dimensions of the residuals $\boldsymbol{E}_i(t)$ obtained in Step 1 have large differences in their range or if they exhibit different amounts of variation, standardization of the residuals may be considered along each dimension $E_i^{(j)}(t)$, to obtain interpretable multivariate functional principal components [32, 37, 38]. The standardization can be carried out by $\widetilde{E}^{(j)}(t) = s_j^{1/2} E^{(j)}(t)$ where $s_j = \left\{ \int_{\mathcal{T}} \widehat{\mathrm{Var}} \left[E^{(j)}(t) \right] \mathrm{d}t \right\}^{-1}$, to guarantee that the integrated variance along the rescaled residuals equals one [32]. In this way all dimensions of the multivariate residuals contribute equal amounts of variation to the analysis, similar to the standardization in classical multivariate principal component analysis (PCA).

3 Data Analysis

3.1 Description of the USRDS study cohort and region-specific risk factors

The USRDS is a large national database which collects data on nearly all patients with ESKD in the U.S. We include patients aged 18 years or older who transitioned to dialysis between January 1, 2005, and September 30, 2013 in

our study cohort and include a maximum follow-up of two years (beginning on day 91 of dialysis, after 90 days to establish a stable treatment modality, and with the last date of follow-up on December 31, 2015). Regional units are taken to be Health Service Areas (HSAs) with relatively self-contained infrastructure for the provision of hospital care in the contiguous U.S., including the District of Columbia. Region-specific multivariate outcome is obtained as the average of facility-specific hospitalization and mortality rates, calculated monthly per person-year (PPY) over the two year follow-up. In obtaining facility-specific rates, we exclude facilities with less than 10 patients in any one month follow-up interval to obtain stable facility-specific hospitalization and mortality rates over time (17.4% facilities excluded). At the region level, we merge HSAs to guarantee that each region contains at least 5 facilities (reducing available HSAs from 719 to 402) and merge HSAs to ensure that the total number of months with zero mortality rate is limited to one-third of the two-year follow-up, to further stabilize region-specific inference. The final study cohort contains 367 regions/HSAs after the second merging, where majority (61.6%) of the resulting 367 regions are unmerged, and 15.0%, 8.4%, and 15.0% of the regions consist of merging 2, 3, and > 3 original HSAs, respectively. The mean region-specific hospitalization and mortality rate in the final cohort are 1.810 and 0.073 PPY, respectively. Detailed descriptions of the study cohort, exclusion rules and region merging algorithms are deferred to Supplementary Information Appendix A.

The proposed modeling includes urbanicity, area deprivation index (ADI) and medical underservice index (MUI) as region-specific covariates. Urbanicity is assigned by categorizing HSAs into the three classes of large metropolitan, small metropolitan or non-metropolitan regions. Categorization of HSAs are assigned according to the class that majority of the counties within each HSA fall into, determined by the urban-rural classification scheme of the National Center for Health Statistics (https://www.cdc.gov/nchs/data_access/urban_rural.htm). Non-metropolitan regions are taken as the reference group. ADI, consisting of 17 education, employment, housing-quality, and poverty measures, captures the socio-economic status of HSAs [39]. The rank-based index, taking on values between 0 and 100 (higher values correspond to lower socio-economic status and higher deprivation), is available at the census block group level (available at https://www.neighborhoodatlas.medicine.wisc.edu), and is averaged over census block groups within each HSA to arrive at the HSA level index. The last covariate, MUI, takes on values between 0 and 1 to reflect the medical service availability within each HSA (higher indices corresponding to higher underservice). The index is released by the Health Resources and Services Administration at the census tract/county subdivision level at https://data.hrsa.gov/tools/shortage-area, where medically underserved tracts/subdivisions are areas with too few primary care providers, high infant mortality, high poverty or a high elderly population. The proportion of census tracts/county subdivisions that are designated as underserved is first targeted for each county, and county MUIs are then averaged within each HSA

to arrive at the HSA-level MUI index. ADI and MUI are mean-centered for ease of interpretation in modeling.

3.2 Results

3.2.1 Time-varying effects of region-specific covariates

The estimated time-varying effects of the region-specific covariates are given in Figure 1. The time-varying y-intercepts represent the hospitalization (Figure 1 (a)) and mortality (Figure 1 (b)) rates of an "average" non-metropolitan region (reference group) with mean ADI of 58.948 and MUI of 0.502. The highest rates in both outcomes (2.023 PPY for hospitalization and 0.083 PPY for mortality) are observed within the first three months after transitioning to dialysis where both rates steadily decline during the first year on dialysis. While hospitalization rates continue to decline after the first year, mortality rates remain relatively stable in the second year on dialysis. Figures 1 (c)-(j) display the estimated effects of risk factors (solid) with simultaneous 95% credible bands (dashed), overlaying horizontal lines at zero (gray) included for reference. Large metropolitan regions have significantly higher hospitalization rates than non-metropolitan regions (reference group), especially at initiation of dialysis, where the effect gradually weakens in the second year on dialysis. Similarly, ADI is significantly positively associated with hospitalizations, especially during the first and a half years on dialysis, suggesting that regions with higher deprivation level have higher hospitalization risk. The effects of small metropolitan regions and MUI are not found significant on hospitalizations as the credible bands include zero, indicating that small metropolitan regions have similar hospitalization rates to non-metropolitan regions, and medical service availability is not found to be significantly associated with hospitalization rates. Note that none of the covariates have significant effects on mortality, indicating that hospitalization rates are more susceptible to urbanicity and socioeconomic status of a region than mortality in the dialysis population.

To further assess the effects of the identified risk factors (being a large metropolitan region with a high ADI index) on hospitalization and mortality rates, Supplementary Information Figure S1 displays predicted hospitalization and mortality trajectories for two hypothetical regions, one that is a large metropolitan with 95% percentile ADI, and the other a non-metropolitan region with 5% percentile ADI (while MUI kept at the mean level of 0.502). The 95% simultaneous credible bands (shaded) for the predicted hospitalization and mortality trajectories are formed based on the draws from $\widetilde{Y}_i^{(j)}(t) = \boldsymbol{X}_i^\top \widehat{\boldsymbol{\beta}}^{(j)}(t)$ without the region-specific random effects, corresponding to each posterior draw of the multivariate varying coefficient functions, and following the algorithm outlined at the end of Section 2.2 for forming 95% simultaneous credible bands with $f(t) = Y_i^{(j)}(t)$. A large metropolitan with 95% percentile ADI is predicted to have higher hospitalization and mortality rates, compared to a non-metropolitan region with 5% percentile ADI, although the difference is not found significant with credible bands overlapping throughout the

two-year follow-up for both outcomes. Moreover, while the difference of hospitalization rates between the two regions is relatively stable across the two-year follow-up, the difference of morality rates between these two regions is larger during the first year on dialysis.

3.2.2 Spatiotemporal patterns of hospitalization and mortality rates

In addition to modeling the time-varying effects of the above risk factors, the proposed MV-VCSTM also models the remaining spatiotemporal patterns in the data via time-varying region-specific stochastic deviations $\boldsymbol{U}_i(t)$. Multivariate FPCA is utilized in reducing the dimension of $\boldsymbol{U}_i(t)$, where spatial correlations are induced among the region-specific eigenscores. A total of four and three eigencomponents are retained in univariate FPCA decompositions of hospitalization and mortality, respectively, explaining 99.02% and 99.17% of the total variation. The estimated multivariate eigenfunctions based on the univariate FPCA decompositions are given in Figure 2 with estimated spatial variance parameters $\widehat{\alpha}_1 = 6.964$, $\widehat{\alpha}_2 = 0.259$, $\widehat{\alpha}_3 = 0.151$, $\widehat{\alpha}_4 = 0.149$, $\widehat{\alpha}_5 = 0.021$, $\widehat{\alpha}_6 = 0.017$, and $\widehat{\alpha}_7 = 0.016$. The leading four multivariate eigenfunctions (given in Figures 2 (a)-(b)) mainly explain variation in hospitalization rates, while the last three (given in Figure 2 (c)-(d)) explain variation in mortality. Among the multivariate functions that mostly explain variability in hospitalization, the leading eigenfunction describes constant variation throughout the two-year follow-up, while the remaining three highlight variation in the first and last six months on dialysis (second eigenfunction), variation at 18 months and at end of the two year follow-up (third eigenfunction) and finally at one year and 18 months on dialysis (fourth eigenfunction). Among the multivariate eigenfunctions that mostly explain variation in mortality, the leading one (5th eigenfunction) explains constant variation over the two-year follow-up (similar to the leading eigenfunction explaining mostly constant variation in hospitalization) (given in Figure 2 (c)-(d)). The sixth multivariate eigenfunction explains variation in mortality at one year and 18 months on dialysis, while the seventh highlighting variation in mortality at 5 and 10 months on dialysis. The leading time-varying variation in hospitalization observed within the first six months of dialysis is consistent with higher hospitalization rates observed at initiation of dialysis, while higher variation in the last six months of follow-up may be related to the decrease in sample size of the study cohort at the end of the two year follow-up.

To assess the spatial and temporal variation jointly in hospitalization and mortality risk while accounting for the time-varying effects of region-specific covariates on the multivariate outcome, Figure 3 displays raw hospitalization (a) and mortality (b) rates as well as region-specific predictions from MV-VCSTM at 3 month, 12 months and 18 months after initiation of dialysis. Overall both set of maps highlight elevated hospitalization and mortality rates in the "band" from Massachusetts to southern Texas (darker blue), and a decreasing trend in both rates for longer follow-up times on dialysis (consistent

with the decreasing trends of the estimated y-intercepts for both outcomes). There are also outcome-specific trends with different spatial patterns among hospitalization and mortality risk in some regions, such as some HSAs in northern California, Oregon, Montana, and Idaho have relatively low hospitalization rates but elevated mortality rates, whereas some HSAs in Florida and Arizona with relatively high hospitalization rates have low mortality rates. Note that the predicted maps correspond closely to the raw maps, especially for hospitalization risk, indicating the satisfactory fit of the proposed MV-VCSTM model. Mortality risk is harder to predict than hospitalization risk, however the proposed MV-VCSTM leads to the smallest prediction error also for mortality, compared to reduced models that ignore the spatiotemporal correlations in the region-specific deviations.

The estimated spatial correlation parameter of $\widehat{\nu} = 0.875$, inducing correlations between neighboring HSAs ranging from 0.19 to 0.60, confirm that there is still significant spatiotemporal variation remaining in the data, after adjusting for time-varying effects of covariates. To visualize the remaining spatial patterns in the data, Figure 4 (a)-(b) display the residuals $\boldsymbol{E}_i(t)$ (averaged across time) obtained from the initial multivariate varying coefficient model (MV-VCM) fit to the data in Step 1 under the working independence assumption. To assess the model fit of MV-VCSTM, Figure 4 also displays the predicted region-specific deviations $\boldsymbol{U}_i(t)$ (Figure 4 (c)-(d)) and residuals from MV-VCSTM (Figure 4 (e)-(f)) averaged across follow-up time. Maps of the residuals $\boldsymbol{E}_i(t)$ from MV-VCM and the predicted $\boldsymbol{U}_i(t)$ both confirm the significant spatial variation remaining in the data after adjusting for time-varying effects of covariates. In addition, the similarity between these two maps show that the predicted region-specific deviations are able to capture the remaining spatial variation in the data effectively. By comparison, the residuals from MV-VCSTM are quite small without any obvious spatial correlation pattern, implying that the spatiotemporal correlation has been effectively modeled by MV-VCSTM.

Next, the proposed MV-VCSTM fits are compared to fits from two reduced multivariate varying coefficient models that ignore the spatiotemporal correlations in the region-specific deviations using prediction error. The first comparative model is the multivariate varying-coefficient model (MV-VCM) from Step 1 of the MV-VCSTM estimation algorithm, fitted under the working independence assumption. MV-VCM, while accommodating time-varying effects of covariates on the multivariate outcome, ignores the spatiotemporal correlations remaining in the data after adjusting for time-varying effects of covariates. Different from MV-VCM, the second comparative model, referred to as the multivariate varying-coefficient temporal model (MV-VCTM) incorporates temporal trends in the error modeling after adjusting for time-varying effects of covariates: $Y_i^{(j)}(t_k) = \boldsymbol{X}_i^{\top}\boldsymbol{\beta}^{(j)}(t_k) + U_i^{(j)}(t) + \epsilon_i^{(j)}(t_k)$, where $U_i^{(j)}(t) = \sum_{\ell=1}^{L} \rho_{i\ell}\psi_{\ell}^{(j)}(t_k)$ with $\psi_{\ell}^{(j)}(t)$ denoting the multivariate eigenfunctions and $\rho_{i\ell}$ denoting the corresponding multivariate PC scores. Different from the proposed MV-VCSTM, the PC scores $\rho_{i\ell}$ in MV-VCTM are assumed to be i.i.d

across regions, ignoring the spatial correlations in the region-specific deviations. Similar to MV-VCSTM, MV-VCTM is also fitted using Bayesian hierarchical modeling (implemented via WinBUGS), where region-specific deviations are expanded on multivariate eigenfunctions and VCFs are expanded using penalized thin-plate splines. The estimation algorithm for MV-VCTM also starts with an initial MV-VCM fit to the data, but targets the multivariate eigenfunctions by the multivariate functional principal components of [32]. We compare the three model fits using relative mean squared deviation error (MSDE), i.e., $\text{MSDE}_{\widehat{Y}^{(j)}(t)} = \int\{Y_i^{(j)}(t) - \widehat{Y}_i^{(j)}(t)\}^2 \mathrm{d}t / \int\{Y_i^{(j)}(t)\}^2 \mathrm{d}t$. Note that MSDE is reported from the training data, since in case of a data split or cross-validation, the PC scores for the left out regions cannot be predicted based on the proposed CAR structure with a fixed neighborhood structure. The MSDE results are reported in Table 1. As expected, MV-VCM, ignoring the spatiotemporal correlation of the error leads to the largest MSDEs, followed by MV-VCTM (which ignores the spatial correlation of the error). By comparison, the proposed MV-VCSTM leads to the smallest MSDEs in modeling both hospitalizations and mortality. As will be shown through the simulation studies summarized in the next section, modeling of the spatiotemporal correlations in the errors via MV-VCSTM, not only leads to smaller prediction error for the outcome trajectories, but it also improves efficiency in inference on the VCFs (evaluated via coverage probability and length of the associated credible bands).

4 Simulation Studies

Finite sample properties of the proposed estimation and inference algorithm are studied via simulation studies with varying number of regions and measurement error variance for multivariate response from two and three dimensions under aligned and misaligned designs where data from different regions are observed at aligned and misaligned time points, respectively. In addition, we compare the finite sample performance of the proposed MV-VCSTM to that of the two comparative models MV-VCTM and MV-VCM introduced above, which ignore the spatial and spatiotemporal correlations of the error term, respectively.

4.1 Simulation design

In the simulation case with two-dimensional multivariate response with aligned time points, the bivariate response is simulated using

$$Y_i^{(j)}(t_k) = \boldsymbol{X}_i^{\mathsf{T}} \boldsymbol{\beta}^{(j)}(t_k) + \sum_{\ell=1}^{L} \rho_{i\ell} \psi_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k) \quad \text{for} \quad j = 1, 2,$$

where t_k, $k = 1, \ldots, 24$, denote the equidistant grid of time points between 0 and 1, mimicking the 24 month aligned follow-up in our data application.

In the misaligned design, 50% random missingness is induced for each region to achieve region-specific misaligned time points where t_k, $k = 1, \ldots, 48$, denote the equidistant grid of time points between 0 and 1, and each region is observed at a total of 24 time points after the induced missingness. The covariates $\boldsymbol{X}_i = \{1, X_{i1}, X_{i2}\}^\mathsf{T}$, $i = 1, \ldots, n$, include a y-intercept term, where X_{i1}, X_{i2} are generated from a bivariate normal distribution with mean $(0, 0)^\mathsf{T}$, variance 1 and covariance $2^{-0.5}$. The time-varying coefficient function $\boldsymbol{\beta}^{(j)}(t_k) = \{\beta_1^{(j)}(t_k), \beta_2^{(j)}(t_k), \beta_3^{(j)}(t_k)\}$, $j = 1, 2$, is set to:

$$\beta_0^{(1)}(t) = \frac{1}{5}\exp(2t-1), \quad \beta_1^{(1)}(t) = \frac{1}{2}t(1-t), \quad \beta_2^{(1)}(t) = \frac{1}{8}(1-t)^2,$$

$$\beta_0^{(2)}(t) = \frac{3}{10}\sqrt{2t} + \frac{1}{10}, \quad \beta_1^{(2)}(t) = \frac{1}{2}(t-0.5)^2, \quad \beta_2^{(2)}(t) = \frac{1}{8}t^2.$$

The multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t_k) = \{\psi_\ell^{(1)}(t_k), \psi_\ell^{(2)}(t_k)\}^\mathsf{T}$, $\ell = 1, 2, 3$ (for $L = 3$), are generated using 11 Fourier basis functions on the interval $\mathcal{T} = [0, 2]$. While $\psi_\ell^{(1)}(t)$, $\ell = 1, 2, 3$, are taken to be the 4th, 5th and 8th Fourier basis functions evaluated on $[0, 1]$, respectively, $\psi_\ell^{(2)}(t)$, $\ell = 1, 2, 3$, are set to equal 4th, 9th and 8th basis functions evaluated on $[1, 2]$, shifted to the left by 1 and multiplied by a random sign. The sign of $\psi_\ell^{(2)}(t)$, $\ell = 1, 2, 3$, are selected to guarantee that the multivariate eigenfunctions are orthonormal on $[0, 1]$. A CAR structure is induced on the multivariate PC scores $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \rho_{2\ell}, \ldots, \rho_{n\ell})^\mathsf{T}$. More specifically, the score vector $\boldsymbol{\rho}_\ell$ is generated from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell(D - \nu W)^{-1}$, where W denotes the adjacency matrix and D denotes the diagonal matrix with the total number of neighbors for each region, as described in Section 2.1. The matrices W and D are specified using the map of $n = 49$ states in the contiguous U.S. (including the District of Columbia) and the map of the $n = 367$ HSAs from our data application. The spatial variance parameters α_ℓ's are taken to be $\alpha_1 = 2.4$, $\alpha_2 = 1.2$ and $\alpha_3 = 0.6$, and the spatial correlation parameter ν is set to 0.97. Finally, the measurement errors $\boldsymbol{\epsilon_i}(t_k) = \{\epsilon_i^{(1)}(t_k), \epsilon_i^{(2)}(t_k)\}^\mathsf{T} \sim N\left((0, 0)^\mathsf{T}, \mathrm{diag}(\sigma_1^2, \sigma_2^2)\right)$, where $\sigma_1^2 = \sigma_2^2 = \sigma^2$ equal 0.2 and 2 corresponding to low and high noise simulation cases, respectively. Details on generation of data for the simulation with three-dimensional multivariate response is deferred to Supplementary Information Appendix B.

4.2 Results

The relative mean squared deviation error (MSDE), $\mathrm{MSDE}_{\widehat{\theta}(t)} = [\int\{\widehat{\theta}(t) - \theta(t)\}^2 \mathrm{d}t] / \int\{\theta(t)\}^2 \mathrm{d}t$ (for a generic function $\theta(t)$), and the mean squared error (MSE), $\mathrm{MSE}_{\widehat{\theta}} = (\theta - \widehat{\theta})^2/\theta^2$ (for a generic parameter θ), are utilized to assess estimation of the time-varying and time-invariant parameters, respectively. In addition to the time-varying model parameters, MSDEs of the estimated region-specific deviations $\widehat{U}_i^{(j)}(t)$ and predicted outcome trajectories $\widehat{Y}_i^{(j)}(t)$ are also reported. Reported summaries exclude outlier MSDE values

(MSDE> 5) for $\widehat{U}_i^{(j)}(t)$ with denominator values close to zero (less than 2%). All VCFs are modeled with penalized thin-plate spline basis functions using $Q = 3$ number of knots. Results were found to be robust to varying choice of number of knots in a sensitivity analysis. Details on the specific prior settings used in the Bayesian modeling of Steps 3 and 5 of the estimation algorithm are deferred to Supplementary Information Appendix C. Reported results are based on a total of 200 Monte Carlo runs. While the MCMC in Step 3 of the estimation algorithm uses two parallel chains with 15000 iterations (5000 burn-in), the MCMC in Step 5 uses one chain with 15000 iterations (5000 burn-in) and returns one out of every ten samples. All Markov chains are verified to have good mixing and convergence properties and the trace plots are given in Supplementary Information (Figure S2 and S3).

The mean MSDE and MSE values from four simulation set-ups with aligned time points and two sets of total number of regions ($n = 49$ and $n = 367$) and two error variances ($\sigma^2 = .2$ and $\sigma^2 = 2.0$) are reported in Tables 2 for the two-dimensional multivariate response. All error measures decrease with increasing number of regions, as expected. In addition, MSDEs for all quantities also decrease with decreasing noise level σ^2, while MSEs for variance components are comparable across the two noise levels, showing that the proposed MV-VCSTM can remove effects of measurement error under both error settings. Even under $n = 49$ and the higher error variance, error measures are small, signaling a good fit. Error measures for $n = 49$ and $\sigma^2 = .2$ and 2.0 from two-dimensional multivariate outcome with misaligned time points (reported in Table S1) and from three-dimensional multivariate outcome with aligned time points (reported in Table S2) follow similar trends, demonstrating that the proposed MV-VCSTM methodology is still effective when time points may be misaligned across regions and with response in higher dimensions. The estimated VCFs from the simulation runs with the 5th, 50th and 95th percentile MSDEs based on $n = 49$ and $n = 367$ regions are given in Figures S4 and Figure S5, respectively. In addition, the estimated multivariate eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs based on $n = 49$ and $n = 367$ regions are displayed in Figures S6 and Figure S7, respectively. The estimates track the true functions, where quantities are estimated better with increasing number of regions and decreasing error variance, as expected.

Finally, the finite sample performance of the proposed MV-VCSTM is compared to two multivariate varying coefficient models, MV-VCTM and MV-VCM, which ignore the spatial and spatiotemporal correlations of the error term, respectively. Model comparisons at varying error variance for $n = 49$ and $n = 367$ regions are summarized in Tables S3 and Table 3, respectively, for the two-dimensional multivariate response (results for the three-dimensional response from $n = 49$ are deferred to Table S4). Models are compared with respect to MSDEs of VCFs, region-specific deviations $\widehat{U}_i^{(j)}(t)$, outcome predictions $\widehat{Y}_i^{(j)}(t)$, as well as the VCF coverage probabilities (CP) and the corresponding length of 95% credible bands (CBs). Intervals reported for MV-VCM are approximate 95% pointwise empirical confidence intervals obtained

by the `gam` function in the R package `mgcv`. Simultaneous CBs for MV-VCTM are formed in the same way as MV-VCSTM, based on posterior samples of the VCFs. MV-VCM, ignoring the spatiotemporal correlations in the errors, has the worst performance among the three models, with higher MSDE for the y-intercept VCF and the predicted response trajectories $\widehat{Y}_i^{(j)}(t)$ and significant under coverage for the y-intercept VCF in both dimensions. The proposed MV-VCSTM, incorporating spatiotemporal correlations in the error leads to the lowest MSDEs out of all three methods where length of the CBs are also smaller compared to those from MV-VCTM as well, signaling improved efficiency in inference. The length of CBs for MV-VCSTM tend to be larger in the misaligned case, which is expected with misalignment adding additional variation to estimation. Nevertheless, the advantages of MV-VCSTM stay the same over the comparator methods of MV-VCM and MV-VCTM, in terms of leading to the smallest MSDE for varying coefficient functions, region-specific deviations and multivariate predictions, as well as CBs with the highest coverage probability and the smallest length, also in the misaligned time points case. Results given in Table S4 for the three-dimensional multivariate outcome follow similar trends. Note that CPs reported from MV-VCTM and MV-VCSTM are higher than the nominal level, as expected, as Bayesian CBs tend to be conservative [40, 41].

5 Discussion

A novel multivariate varying coefficient spatiotemporal model (MV-VCSTM) is proposed to study time-varying effects of risk factors jointly on hospitalization and mortality rates in the U.S. dialysis population. In addition to adjusting for time-varying regression effects of covariates which are characterized as functions of time that patients stay on dialysis, the proposed modeling approach accounts for the remaining joint spatiotemporal patterns of hospitalization and mortality rates among geographic regions (HSAs). To address challenges in estimation and inference in high-dimensional multivariate outcome settings, the proposed estimation relies on multiple dimension reduction techniques. The region-specific random deviations are expanded on multivariate eigenfunctions (which leads to representations in lower dimensions, involving only a few multivariate eigenfunctions), which are targeted using only univariate FPCA expansions, leading to significant computational savings by avoiding smoothing over high dimensional covariance surfaces. Expansions using multivariate eigenfunctions not only help achieve computational savings, they also lead to extra interpretations in modeling, via capturing the leading directions of temporal variation in the errors (after adjustment for time-varying effects of risk factors). Additional computational savings are achieved in MCAR modeling using independent latent Gaussian processes with a CAR correlation structure. Hence, the proposed modeling, while leading to a nonseparable space and time covariance structure in the outcomes, can still easily scale up to multivariate response in higher dimensions. Modeling

the spatiotemporal correlations in the errors via MST-VCSTM is shown to improve the efficiency of inference on the VCFs through simulation studies. Finally, applications of the proposed methodology to USRDS data lead to the discovery of significant risk factors (large metropolitan and ADI) for hospitalization rates. In addition, the proposed MV-VCSTM leads to identification of "hot" spots (with high rates in both hospitalization and mortality or with differing patterns in the outcomes) and time periods with elevated rates after transitioning to dialysis.

Data Availability 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.

Supplementary Information The supplementary material for this article, including referenced appendices, is available online. The R code and documentation for implementing the proposed MV-VCSTM on simulated datasets are provided on Github at https://github.com/dsenturk/MV-VCSTM.

Acknowledgements This study was supported by a grant from the National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK092232- DS, DVN, EK, SB, CMR, QQ and YL). 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 U.S. government.

References

[1] USRDS (2022) United States Renal Data System 2022 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, Maryland.

[2] Bello AK, Okpechi IG, Osman MA, Cho Y, Htay H, Jha V, Wainstein M, Johnson DW (2022) Epidemiology of haemodialysis outcomes. Nature Reviews Nephrology 18(6):378–395. https://doi.org/10.1038/s41581-022-00542-7.

[3] Hickson LJ, Thorsteinsdottir B, Ramar P, Reinalda MS, Crowson CS, Williams AW, Albright RC, Onuigbo MA, Rule AD, Shah ND (2018) Hospital readmission among new dialysis patients associated with young age and poor functional status. Nephron 139(1):1–2. https://doi.org/10.1159/000485985.

[4] Estes JP, Nguyen DV, Chen Y, Dalrymple LS, Rhee CM, Kalantar-Zadeh K, et al (2018) Time-dynamic profiling with application to hospital readmission among patients on dialysis. Biometrics 74(4):1383–94. https://doi.org/10.1111/biom.12908.

[5] Li Y, Nguyen DV, Chen Y, Rhee CM, 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. https://doi.org/10.1002/sim.7950.

[6] Estes JP, Nguyen DV, Dalrymple LS, Mu Y, Şentürk D (2016) Time-varying effect modeling with longitudinal data truncated by death: conditional models, interpretations, and inference. Statistics in Medicine 35(11):1834–1847. https://doi.org/10.1002/sim.6836.

[7] Noordzij M, Jager KJ (2014) Increased mortality early after dialysis initiation: a universal phenomenon. Kidney International 85(1):12–14. https://doi.org/10.1038/ki.2013.316.

[8] de Jager DJ, Grootendorst DC, Jager KJ, van Dijk PC, Tomas LM, Ansell D, et al (2009) Cardiovascular and noncardiovascular mortality among patients starting dialysis. Jama 302(16):1782–1789. https://doi.org/10.1001/jama.2009.1488.

[9] Li Y, Nguyen DV, Banerjee S, Rhee CM, Kalantar-Zadeh K, Kürüm E, et al (2021) Multilevel modeling of spatially nested functional data: Spatiotemporal patterns of hospitalization rates in the US dialysis population. Statistics in Medicine 40(17):3937–3752. https://doi.org/10.1002/sim.9007.

[10] Li Y, Nguyen DV, Kürüm E, Rhee CM, Banerjee S, Şentürk D (2022) Multilevel varying coefficient spatiotemporal model. Stat 11(1):e438. https://doi.org/10.1002/sta4.438.

[11] Erickson KF, Zhao B, Niu J, Winkelmayer WC, Bhattacharya J, Chertow GM, et al (2019) Association of hospitalization and mortality among patients initiating dialysis with hemodialysis facility ownership and acquisitions. JAMA Network Open 2(5):e193987. https://doi.org/10.1001/jamanetworkopen.2019.3987.

[12] Zhang S, Sun D, He CZ, Schootman M (2006) A Bayesian semi-parametric model for colorectal cancer incidences. Statistics in Medicine 25(2):285–309. https://doi.org/10.1002/sim.2221.

[13] Hepler SA, Waller LA, Kline DM (2021) A multivariate spatiotemporal change-point model of opioid overdose deaths in Ohio. The Annals of Applied Statistics 15(3):1329–1342. https://doi.org/10.1214/

20-aoas1415.

[14] Baer DR, Lawson AB, Joseph JE (2021) Joint space-time Bayesian disease mapping via quantification of disease risk association. Statistical Methods in Medical Research 30(1):35–61. https://doi.org/10.1177/0962280220938975.

[15] Zhu H, Li R, Kong L (2012) Multivariate varying coefficient model for functional responses. Annals of Statistics 40(5):2634–2666. https://doi.org/10.1214/12-AOS1045SUPP.

[16] Kürüm E, Li R, Shiffman S, Yao W (2016) Time-varying coefficient models for joint modeling binary and continuous outcomes in longitudinal data. Statistica Sinica 26(3):979–1000. https://doi.org/10.5705/ss.2014.213.

[17] Zhang F, Li R, Lian H, Bandyopadhyay D (2021) Sparse reduced-rank regression for multivariate varying-coefficient models. Journal of Statistical Computation and Simulation 91(4):752–767. https://doi.org/10.1080/00949655.2020.1829622.

[18] Morris JS, Carroll RJ (2006) Wavelet-based functional mixed models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68(2):179–199. https://doi.org/10.1111/j.1467-9868.2006.00539.x.

[19] Cai J, Fan J, Zhou H, Zhou Y (2007) Hazard models with varying coefficients for multivariate failure time data. Annals of Statistics 35(1):324–354. https://doi.org/10.1214/009053606000001145.

[20] He K, Lian H, Ma S, Huang JZ (2018) Dimensionality reduction and variable selection in multivariate varying-coefficient models with a large number of covariates. Journal of the American Statistical Association 113(522):746–754. https://doi.org/10.1080/01621459.2017.1285774.

[21] Yee TW, Wild CJ (1996) Vector generalized additive models. Journal of the Royal Statistical Society: Series B (Methodological) 58(3):481–493. https://doi.org/10.1111/j.2517-6161.1996.tb02095.x.

[22] Wild CJ, Yee TW (1996) Additive extensions to generalized estimating equation methods. Journal of the Royal Statistical Society: Series B (Methodological) 58(4):711–725. https://doi.org/10.1111/j.2517-6161.1996.tb02110.x.

[23] Yee TW, Mackenzie M (2002) Vector generalized additive models in plant ecology. Ecological Modelling 157(2-3):141–156. https://doi.org/10.1016/S0304-3800(02)00192-8.

[24] Guo Y, Sun D, Sun J (2022) Inference of a time-varying coefficient regression model for multivariate panel count data. Journal of Multivariate Analysis 192:105047. https://doi.org/10.1016/j.jmva.2022.105047.

[25] Gelfand AE, Banerjee S, Gamerman D (2005) Spatial process modelling for univariate and multivariate dynamic spatial data. Environmetrics 16(5):465–479. https://doi.org/10.1002/env.715.

[26] Congdon P (2004) A multivariate model for spatio-temporal health outcomes with an application to suicide mortality. Geographical Analysis 36(3):234–258. https://doi.org/10.1111/j.1538-4632.2004.tb01134.x.

[27] Cheng W, Gill GS, Dasu R, Xie M, Jia X, Zhou J (2017) Comparison of multivariate Poisson lognormal spatial and temporal crash models to identify hot spots of intersections based on crash types. Accident Analysis & Prevention 99:330–341. https://doi.org/10.1016/j.aap.2016.11.022.

[28] Quick H, Banerjee S, Carlin BP (2013) Modeling temporal gradients in regionally aggregated California asthma hospitalization data. The Annals of Applied Statistics 7(1):154–176. https://doi.org/10.1214/12-AOAS600.

[29] Qian Q, Nguyen DV, Telesca D, Kürüm E, Rhee CM, Banerjee S, et al (2023) Multivariate spatiotemporal functional principal component analysis for modeling hospitalization and mortality rates in the dialysis population. Biostatistics: In press. https://doi.org/10.1093/biostatistics/kxad013.

[30] Ramsay JO, Silverman BW (2005) Functional data analysis. Springer, New York.

[31] Crainiceanu CM, Ruppert D, Carroll RJ, Joshi A, Goodner B (2007) Spatially adaptive Bayesian penalized splines with heteroscedastic errors. Journal of Computational and Graphical Statistics 16(2):265-88. https://doi.org/10.1198/106186007X208768.

[32] Happ C, Greven S (2018) Multivariate functional principal component analysis for data observed on different (dimensional) domains. Journal of the American Statistical Association 113(522):649–659. https://doi.org/10.1080/01621459.2016.1273115.

[33] Banerjee S, Carlin BP, Gelfand AE (2003) Hierarchical modeling and analysis for spatial data. Chapman and Hall/CRC, Boca Raton, Florida.

[34] Jin X, Banerjee S, Carlin BP (2007) Order-free co-regionalized areal data models with application to multiple-disease mapping. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69(5):817–838. https://doi.org/10.1111/j.1467-9868.2007.00612.x.

[35] Ruppert D, Wand MP, Carroll RJ (2003) Semiparametric regression. Cambridge University Press, Cambridge.

[36] Crainiceanu C, Ruppert D, Wand MP (2005) Bayesian analysis for penalized spline regression using WinBUGS. Journal of Statistical Software 14(14):1–24. https://doi.org/10.18637/jss.v014.i14.

[37] Jacques J, Preda C (2014) Model-based clustering for multivariate functional data. Computational Statistics & Data Analysis 71:92–106. https://doi.org/10.1016/j.csda.2012.12.004.

[38] Chiou JM, Chen YT, Yang YF (2014) Multivariate functional principal component analysis: a normalization approach. Statistica Sinica 24(4):1571–1596. https://doi.org/10.5705/ss.2013.305.

[39] Kind AJ, Buckingham WR (2018) Making neighborhood-disadvantage metrics accessible the neighborhood atlas. The New England journal of medicine 378(26):2456–2458. https://doi.org/10.1056/NEJMp1802313.

[40] Cox DD (1993) An analysis of Bayesian inference for nonparametric regression. The Annals of Statistics 21(2):903–923. https://doi.org/10.1214/aos/1176349157.

[41] Krivobokova T, Kneib T, Claeskens G (2010) Simultaneous confidence bands for penalized spline estimators. Journal of the American Statistical Association 105(490):852–863. https://doi.org/10.1198/jasa.2010.tm09165.

[42] Brook, D (1964) On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. JBiometrika 51(3/4):481–483. https://www.jstor.org/stable/2334154.

Table 1: The mean MSDEs from predicted multivariate response trajectories using MV-VCM, MV-VCTM, MV-VCSTM fits to the USRDS data.

	MV-VCM	**MV-VCTM**	**MV-VCSTM**
Hospitalization	0.048	0.045	0.015
Mortality	0.321	0.314	0.236

Table 2: The mean MSDE of VCFs, eigenfunctions, region-specific deviations and multivariate response trajectories, and MSE for variance components and spatial correlation parameters from four simulation settings with aligned time points, two sets of total number of regions and two sets of measurement error variance σ^2. Results are based on two-dimensional multivariate response and 200 Monte Carlo runs.

Number of regions, n	**49 regions**		**367 regions**	
Noise level, σ^2	0.2	2	0.2	2
	MSDE			
$\widehat{\beta}_0^{(1)}(t)$	0.0033	0.0178	0.0003	0.0015
$\widehat{\beta}_0^{(2)}(t)$	0.0031	0.0236	0.0003	0.0011
$\widehat{\beta}_1^{(1)}(t)$	0.0017	0.0096	0.0004	0.0020
$\widehat{\beta}_1^{(2)}(t)$	0.0013	0.0093	0.0002	0.0015
$\widehat{\beta}_2^{(1)}(t)$	0.0011	0.0126	0.0001	0.0011
$\widehat{\beta}_2^{(2)}(t)$	0.0009	0.0093	0.0001	0.0012
$\widehat{\psi}_1^{(1)}(t)$	0.0795	0.0872	0.0105	0.0179
$\widehat{\psi}_1^{(2)}(t)$	0.0821	0.0846	0.0104	0.0176
$\widehat{\psi}_2^{(1)}(t)$	0.0960	0.1222	0.0164	0.0265
$\widehat{\psi}_2^{(2)}(t)$	0.1060	0.1208	0.0168	0.0281
$\widehat{\psi}_3^{(1)}(t)$	0.0349	0.1697	0.0048	0.0157
$\widehat{\psi}_3^{(2)}(t)$	0.0333	0.1769	0.0050	0.0172
$\widehat{U}_i^{(1)}(t)$	0.0465	0.3447	0.0416	0.2775
$\widehat{U}_i^{(2)}(t)$	0.0470	0.3516	0.0418	0.2789
$\widehat{Y}_i^{(1)}(t)$	0.0251	0.3499	0.0244	0.2010
$\widehat{Y}_i^{(2)}(t)$	0.0205	0.2152	0.0196	0.1594
	MSE			
$\widehat{\alpha}_1$	0.0985	0.1067	0.0284	0.0300
$\widehat{\alpha}_2$	0.0260	0.0347	0.0091	0.0156
$\widehat{\alpha}_3$	0.0188	0.0696	0.0150	0.0246
$\widehat{\nu}$	0.0067	0.0070	0.0066	0.0067
$\widehat{\sigma}^{2(1)}$	0.0022	0.0023	0.0005	0.0005
$\widehat{\sigma}^{2(2)}$	0.0023	0.0024	0.0004	0.0005

Table 3: The mean MSDE of VCFs, deviations $\widehat{U}_i^{(j)}(t)$ and predictions $\widehat{Y}_i^{(j)}(t)$, along with coverage probabilities (CPs, %) and length of the 95% credible bands for $n = 367$ regions from MV-VCM, MV-VCTM, and MV-VCSTM, respectively. Results are based on two-dimensional multivariate response with aligned time points and 200 Monte Carlo runs.

# regions	**367 regions**					
Noise level	0.2			2		
Model	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM
	MSDE					
$\widehat{\beta}_0^{(1)}(t)$	0.0691	0.0513	0.0003	0.0828	0.0606	0.0015
$\widehat{\beta}_0^{(2)}(t)$	0.0690	0.0514	0.0003	0.0829	0.0613	0.0011
$\widehat{\beta}_1^{(1)}(t)$	0.0005	0.0009	0.0004	0.0039	0.0031	0.0020
$\widehat{\beta}_1^{(2)}(t)$	0.0006	0.0008	0.0002	0.0026	0.0026	0.0015
$\widehat{\beta}_2^{(1)}(t)$	0.0012	0.0009	0.0001	0.0017	0.0024	0.0011
$\widehat{\beta}_2^{(2)}(t)$	0.0011	0.0008	0.0001	0.0026	0.0022	0.0012
$\widehat{U}_i^{(1)}(t)$	/	0.2045	0.0416	/	0.4150	0.2775
$\widehat{U}_i^{(2)}(t)$	/	0.2049	0.0418	/	0.4161	0.2789
$\widehat{Y}_i^{(1)}(t)$	0.6748	0.0282	0.0244	0.6777	0.2168	0.2010
$\widehat{Y}_i^{(2)}(t)$	0.5763	0.0222	0.0196	0.5853	0.1734	0.1594
	CP (%)					
$\widehat{\beta}_0^{(1)}(t)$	33.71	100	100	30.06	100	100
$\widehat{\beta}_0^{(2)}(t)$	31.48	100	100	28.08	100	100
$\widehat{\beta}_1^{(1)}(t)$	100	100	100	98.96	97.5	99.5
$\widehat{\beta}_1^{(2)}(t)$	100	100	100	99.98	100	100
$\widehat{\beta}_2^{(1)}(t)$	100	100	100	100	100	100
$\widehat{\beta}_2^{(2)}(t)$	100	100	100	100	100	100
	Length					
$\widehat{\beta}_0^{(1)}(t)$	0.192	1.136	0.955	0.339	1.158	0.920
$\widehat{\beta}_0^{(2)}(t)$	0.196	1.341	0.897	0.349	1.347	0.784
$\widehat{\beta}_1^{(1)}(t)$	0.227	0.412	0.265	0.285	0.557	0.334
$\widehat{\beta}_1^{(2)}(t)$	0.221	0.553	0.428	0.355	0.666	0.427
$\widehat{\beta}_2^{(1)}(t)$	0.213	0.363	0.278	0.277	0.519	0.333
$\widehat{\beta}_2^{(2)}(t)$	0.237	0.472	0.318	0.280	0.601	0.443

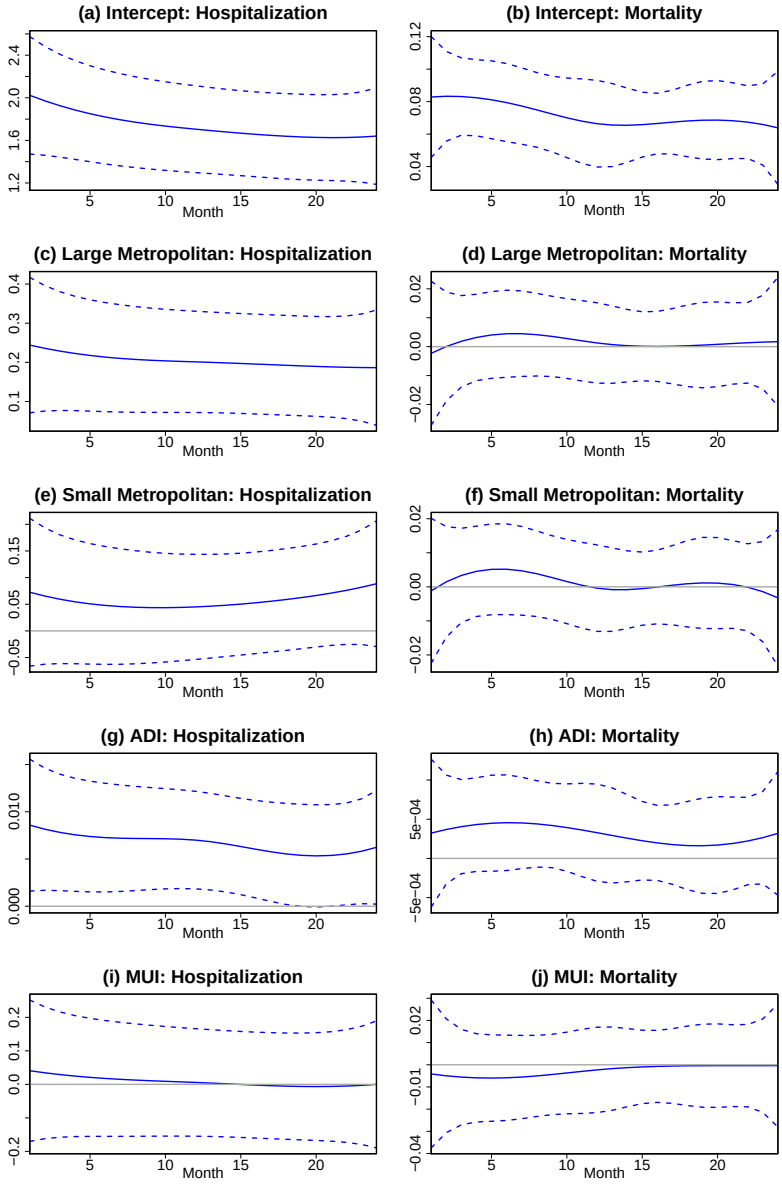

Fig. 1: Estimated time-varying effects (solid) of covariates: y-intercept, large metropolitan, small metropolitan, area deprivation index (ADI) (centered) and medical underservice index (MUI) (centered), along with their 95% simultaneous credible bands (dashed), in modeling hospitalization (a, c, e, g, i) and mortality (b, d, f, h, j) rates among the U.S. dialysis population. Horizontal lines at zero are included in gray for reference.

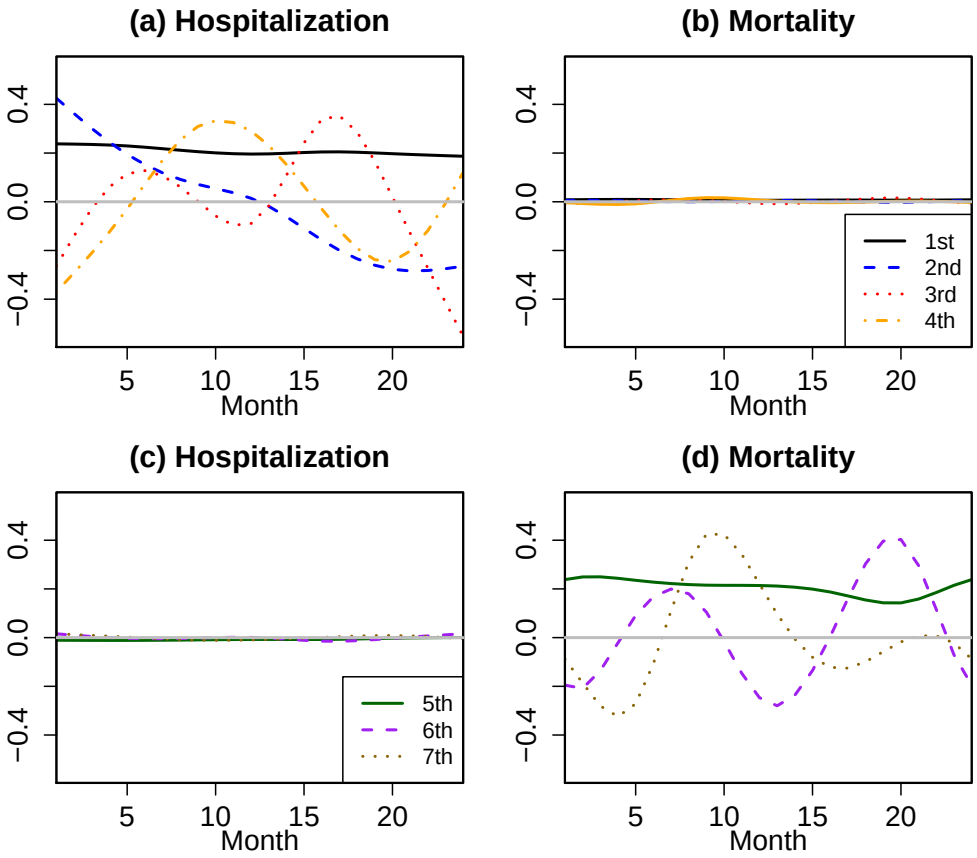

Fig. 2: Estimated multivariate eigenfunctions for hospitalization (a, c) and mortality (b, d). The leading four multivariate eigenfunctions (a, b) mainly explain variation in hospitalization, while the last three multivariate eigenfunctions (c, d) mainly explain variation in mortality. The zero horizontal lines are included in gray for reference.

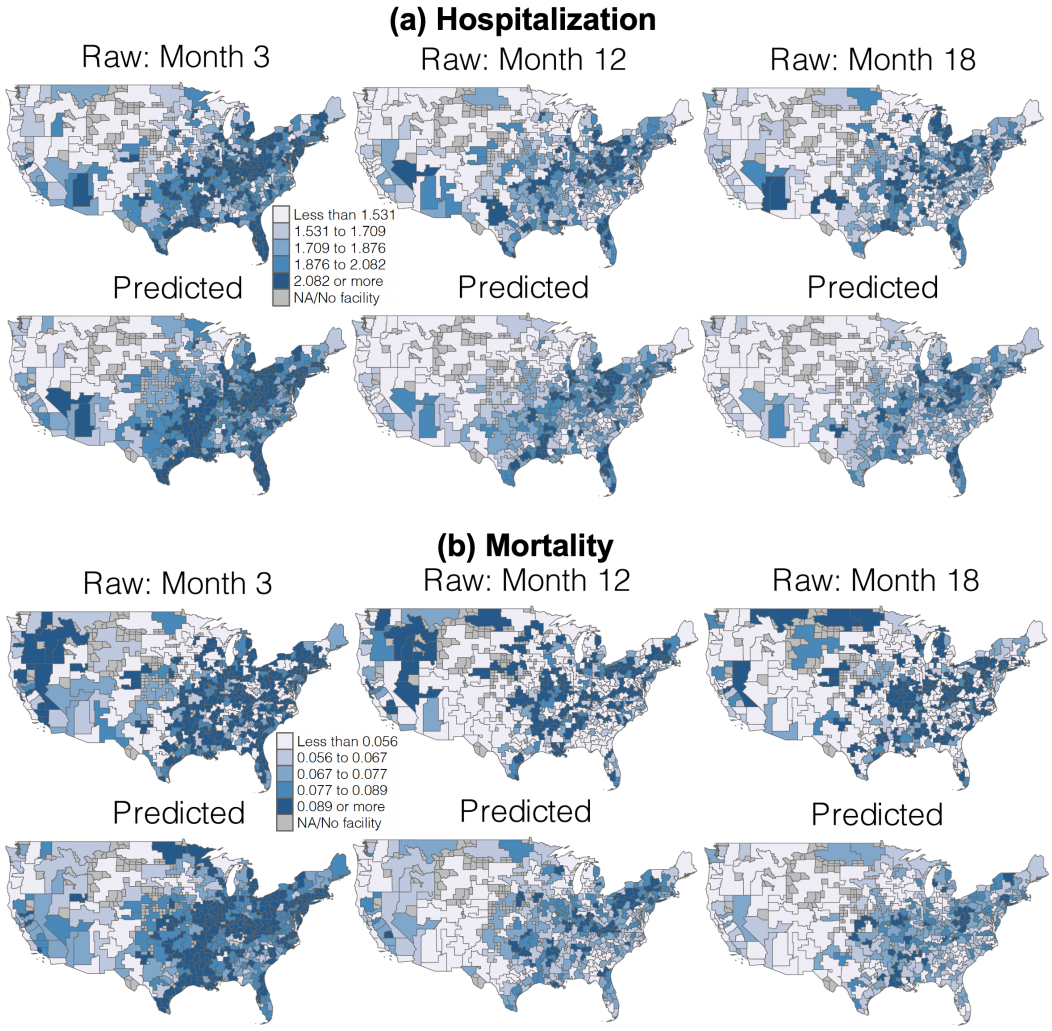

Fig. 3: The raw and predicted hospitalization (a) and mortality (b) rates from the 3rd, 12th, and 18th months on dialysis for all HSAs.

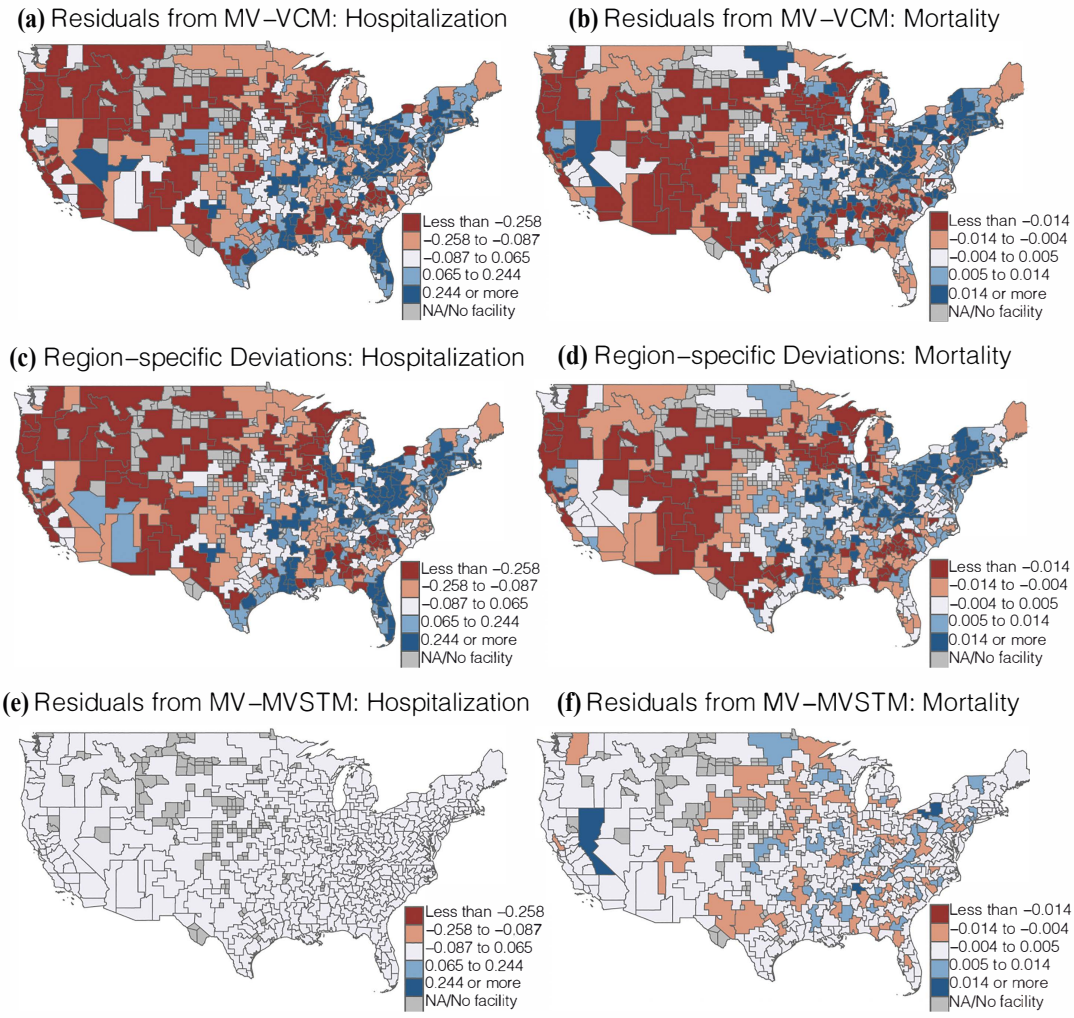

Fig. 4: The residuals $\boldsymbol{E}_i(t)$ (averaged across time) obtained from the initial multivariate varying coefficient model fit to the data under the working independence assumption (a-b), the predicted region-specific deviations $\boldsymbol{U}_i(t)$ averaged across follow-up time (c-d), and the residuals (averaged across time) obtained from the proposed MV-VCSTM (e-f) for hospitalization (a, c, e) and mortality (b, d, f).

Supplementary Materials for 'Multivariate varying coefficient spatiotemporal model'

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

[1]*Department of Biostatistics, University of California, Los Angeles, CA, USA*
[2]*Department of Medicine, University of California, Irvine, CA, USA*
[3]*Department of Statistics, University of California, Riverside, CA, USA*
[4]*Harold Simmons Center for Chronic Disease Research and Epidemiology, University of California Irvine, School of Medicine, Irvine, CA, US*

Appendices

Online Appendix A. Details on the USRDS Study Cohort

The main eligibility criteria for the study cohort are that (a) patients survive the first 90 days of dialysis and do not recover renal function or receive kidney transplantation and (b) that they have Medicare as their primary payer on day 91 of dialysis. To obtain stable region-specific estimation and inference, HSAs are merged first to guarantee that they contain at least 5 facilities and for a second time to guarantee that the total number of months with zero mortality rate does not exceed one-third of the two year follow-up. Both merging algorithms have four steps: Step 1) start with HSAs as regions; Step 2) among regions with the lowest number of facilities (or with greater than eight months of zero mortality), randomly select one region (call it R); Step 3) among neighbors of the selected region R, randomly select a neighbor, among those with the lowest number of facilities (or among those with the largest number of months with zero mortality rate), to merge with region R; Step 4) repeat Steps 2 and 3 until the lowest number of facilities within a region is 5 (or the largest number of months with zero mortality rate during the two-year follow up within a region is eight).

Online Appendix B. Simulation Set-up for Multivariate Response in Three Dimensions

In the simulation setting with three-dimensional multivariate response, the trivariate response is simulated using

$$Y_i^{(j)}(t_k) = \boldsymbol{X}_i^{\mathsf{T}}\boldsymbol{\beta}^{(j)}(t_k) + \sum_{\ell=1}^{L}\rho_{i\ell}\psi_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k), \text{ for } j = 1, 2, 3,$$

where t_k, $k = 1, \ldots, 24$, denote the equidistant grid of time points between 0 and 1, mimicking the 24 month follow-up in our data application. The covariates $\boldsymbol{X}_i = \{1, X_{i1}, X_{i2}\}^{\mathsf{T}}$ for $i = 1, \ldots, n$ include a y-intercept term, and X_{i1}, X_{i2} generated from a bivariate normal with mean $(0, 0)^{\mathsf{T}}$, variance 1 and covariance $2^{-0.5}$. The time-varying coefficient function, $\boldsymbol{\beta}^{(j)}(t_k) = \{\beta_1^{(j)}(t_k), \beta_2^{(j)}(t_k), \beta_3^{(j)}(t_k)\}$, $j = 1, 2, 3$, is set as follows:

$$\beta_0^{(1)}(t) = 0.1\exp(2t-1), \quad \beta_1^{(1)}(t) = 0.25t(1-t), \quad \beta_2^{(1)}(t) = 0.06(1-t)^2,$$

$$\beta_0^{(2)}(t) = 0.15\sqrt{2t} + 0.05, \quad \beta_1^{(2)}(t) = 0.25(t-0.5)^2, \quad \beta_2^{(2)}(t) = 0.06t^2,$$

$$\beta_0^{(3)}(t) = 0.15t + 0.05, \quad \beta_1^{(3)}(t) = 0.06t^2, \quad \beta_2^{(3)}(t) = 0.05\sqrt{0.5t}.$$

The multivariate eigenfunctions $\boldsymbol{\psi}_\ell(t_k) = \{\psi_\ell^{(1)}(t_k), \psi_\ell^{(2)}(t_k), \psi_\ell^{(3)}(t_k)\}^{\mathsf{T}}$, $\ell = 1, 2, 3$ (for $L = 3$), are generated using 13 Fourier basis functions on the interval $\mathcal{T} = [0, 3]$. While $\psi_\ell^{(1)}(t)$ is taken to be the 12th, 6th and 7th basis functions evaluated on $[0, 1]$ (i.e., $L = 3$), $\psi_\ell^{(2)}(t)$ equals the 7th, 6th and 12th basis functions

evaluated on $[1,2]$, shifted to the left by one and multiplied by a random sign, and $\psi_\ell^{(3)}(t)$ equals the 6th, 12th and 7th basis functions evaluated on $[2,3]$, shifted to the left by two and multiplied by a random sign. Similar to the simulation with the two-dimensional multivariate response, the multivariate PC scores, $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \rho_{2\ell}, \ldots, \rho_{n\ell})^\top$ are generated from a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\alpha_\ell(D - \nu W)^{-1}$, where W is the adjacency matrix and D is the diagonal matrix utilizing the map of $n = 49$ states in the contiguous U.S. (including the District of Columbia). The variance parameters α_ℓ's are taken to be $\alpha_1 = 2.4$, $\alpha_2 = 1.2$, $\alpha_3 = 0.6$, while the spatial correlation parameter is set to $\nu = 0.97$. Finally, the measurement errors, $\boldsymbol{\epsilon_i}(t_k) = \{\epsilon_i^{(1)}(t_k), \epsilon_i^{(2)}(t_k), \epsilon_i^{(3)}(t_k)\}^\mathsf{T} \sim N\left((0,0,0)^\mathsf{T}, \mathrm{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)\right)$, where $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ equal 0.2 and 2 corresponding to low and high noise simulation cases, respectively.

Online Appendix C. Priors Used in the Data Analysis and Simulation

In Step 3 of the proposed MV-VCSTM estimation algorithm, the between-score covariance matrix Σ_b is targeted through the modeling framework:

$$\widehat{\boldsymbol{\xi}} = \boldsymbol{\xi} + \mathbf{e}$$

$$\mathbf{e} \sim N(\mathbf{0}, \tau^2 I), \quad \boldsymbol{\xi} \sim N\left(\mathbf{0}, AA^\mathsf{T} \otimes (D - \nu W)^{-1}\right),$$

$$a_{\ell\ell} \sim \text{Lognormal}(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}}), \quad a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}}), \quad \nu \sim \text{Unif}(a_\nu, b_\nu), \quad \tau^2 \sim \text{IG}(a_{\tau^2}, b_{\tau^2}),$$

where elements of $\widehat{\boldsymbol{\xi}}$ are the scores estimated from univariate FPCA. Elementwise priors are imposed on A where $a_{\ell\ell} \sim$ lognormal$(\mu_{a_{\ell\ell}}, \sigma^2_{a_{\ell\ell}})$, and $a_{\ell\ell'} \sim N(\mu_{a_{\ell\ell'}}, \sigma^2_{a_{\ell\ell'}})$ for $1 \leq \ell' < \ell$, with $\mu_{a_{\ell\ell'}} = -3$, $\sigma^2_{a_{\ell\ell'}} = 1$, $\mu_{a_{\ell\ell'}} = 0$, and $\sigma^2_{a_{\ell\ell'}} = 5$. We apply an Inverse Gamma (IG)(a_{τ^2}, b_{τ^2}) prior for τ^2 ($a_{\tau^2} = b_{\tau^2} = 1$ in our application), and a Uniform prior for the spatial parameter ν with parameters constrained to lie between bounds given by the inverse of the minimum and maximum eigenvalues (denoted by a_ν, b_ν, respectively) of the matrix $D^{-1/2}WD^{-1/2}$.

In Step 5 of the MV-VCSTM estimation algorithm, we target the varying coefficient functions (VCFs) $\boldsymbol{\beta}^{(j)}(t)$, multivariate PC scores ($\rho_{i\ell}$), spatial parameters (α_ℓ and ν), and measurement error variance (σ^2) of MV-VCSTM via expanding the VCFs on thin-plate splines and the stochastic deviations $\boldsymbol{U}_i(t)$ on the estimated multivariate eigenfunctions $\widehat{\boldsymbol{\psi}}_\ell(t)$ as follows:

$$Y_i^{(j)}(t_k) = \sum_{p=1}^{P} X_{ip}\left(u_{p0}^{(j)} + u_{p1}^{(j)} t + \sum_{q=1}^{Q} \widetilde{v}_{pq}^{(j)} z_{kq}\right) + \sum_{\ell=1}^{M^+} \rho_{i\ell}\widehat{\psi}_\ell^{(j)}(t_k) + \epsilon_i^{(j)}(t_k) \quad \text{for } \ j = 1, \ldots, J,$$

$$\epsilon_i^{(j)}(t_k) \sim N(0, \sigma_j^2), \quad \boldsymbol{\rho}_\ell \sim N(\mathbf{0}, \alpha_\ell(D - \nu W)^{-1}), \quad \alpha_\ell \sim \text{IG}(a_{\alpha_\ell}, b_{\alpha_\ell}), \quad \nu \sim \text{Unif}(a_\nu, b_\nu),$$

$$u_{p0}^{(j)} \sim N(0, \sigma^2_{u_{p0j}}), \quad u_{p1}^{(j)} \sim N(0, \sigma^2_{u_{p1j}}), \quad \widetilde{v}_{pq}^{(j)} \sim N(0, \sigma^2_{\widetilde{v}_{pj}}),$$

$$\sigma^2_{\widetilde{v}_{pj}} \sim \text{IG}(a_{\sigma^2_{\widetilde{v}_{pj}}}, b_{\sigma^2_{\widetilde{v}_{pj}}}), \quad \sigma_j^2 \sim \text{IG}(a_{\sigma_j^2}, b_{\sigma_j^2}),$$

where $\boldsymbol{\rho}_\ell = (\rho_{1\ell}, \ldots, \rho_{n\ell})^\mathsf{T}$ denotes the $n \times 1$ score vector for ℓth eigencomponent. Noninformative or weak informative priors are used for $u_{p0}^{(j)}$, $u_{p1}^{(j)}$, $\widetilde{v}_{pq}^{(j)}$, $\epsilon_i^{(j)}(t)$ and α_ℓ, where $\sigma^2_{u_{p0j}} = \sigma^2_{u_{p1j}} = 10^6$, $a_{\sigma^2_{\widetilde{v}_{pj}}} = b_{\sigma^2_{\widetilde{v}_{pj}}} = 10^{-6}$, and $a_{\sigma_j^2} = b_{\sigma_j^2} = a_{\alpha_\ell} = b_{\alpha_\ell} = 1$. Finally, the prior imposed on ν is the same as in Step 3.

Table S1: The mean MSDE of VCFs, eigenfunctions, region-specific deviations and multivariate response trajectories, and MSE for variance components and spatial correlation parameters from four simulation settings with two sets of time points (aligned and misaligned) and two sets of measurement error variance σ^2. Results are based on two-dimensional multivariate response, $n = 49$ regions and 200 Monte Carlo runs.

Time points	**Aligned**		**Misaligned**	
Noise level, σ^2	0.2	2	0.2	2
	MSDE			
$\widehat{\beta}_0^{(1)}(t)$	0.0033	0.0178	0.0130	0.0590
$\widehat{\beta}_0^{(2)}(t)$	0.0031	0.0236	0.0107	0.0551
$\widehat{\beta}_1^{(1)}(t)$	0.0017	0.0096	0.0038	0.0095
$\widehat{\beta}_1^{(2)}(t)$	0.0013	0.0093	0.0036	0.0096
$\widehat{\beta}_2^{(1)}(t)$	0.0011	0.0126	0.0038	0.0089
$\widehat{\beta}_2^{(2)}(t)$	0.0009	0.0093	0.0034	0.0079
$\widehat{\psi}_1^{(1)}(t)$	0.0795	0.0872	0.0373	0.0513
$\widehat{\psi}_1^{(2)}(t)$	0.0821	0.0846	0.0330	0.0507
$\widehat{\psi}_2^{(1)}(t)$	0.0960	0.1222	0.0292	0.1013
$\widehat{\psi}_2^{(2)}(t)$	0.1060	0.1208	0.0272	0.1099
$\widehat{\psi}_3^{(1)}(t)$	0.0349	0.1697	0.0245	0.3649
$\widehat{\psi}_3^{(2)}(t)$	0.0333	0.1769	0.0283	0.3602
$\widehat{U}_i^{(1)}(t)$	0.0465	0.3447	0.2234	0.3773
$\widehat{U}_i^{(2)}(t)$	0.0470	0.3516	0.2207	0.3623
$\widehat{Y}_i^{(1)}(t)$	0.0251	0.3499	0.0398	0.2693
$\widehat{Y}_i^{(2)}(t)$	0.0205	0.2152	0.0278	0.2241
	MSE			
$\widehat{\alpha}_1$	0.0985	0.1067	0.0511	0.0993
$\widehat{\alpha}_2$	0.0260	0.0347	0.0168	0.0602
$\widehat{\alpha}_3$	0.0188	0.0696	0.0078	0.1126
$\widehat{\nu}$	0.0067	0.0070	0.0163	0.0836
$\widehat{\sigma}^{2(1)}$	0.0022	0.0023	0.0032	0.0034
$\widehat{\sigma}^{2(2)}$	0.0023	0.0024	0.0019	0.0021

Table S2: The mean MSDE of VCFs, eigenfunctions, region-specific deviations and multivariate response trajectories, and MSE for variance components and spatial correlation parameters from three-dimensional multivariate response with aligned time points, $n = 49$ regions and varying measurement error variance σ^2. Results are based on 200 Monte Carlo runs.

Number of regions, n	**49 regions**	
Noise level, σ^2	0.2	2
	MSDE	
$\widehat{\beta}_0^{(1)}(t)$	0.0047	0.0576
$\widehat{\beta}_0^{(2)}(t)$	0.0059	0.0651
$\widehat{\beta}_0^{(3)}(t)$	0.0097	0.0582
$\widehat{\beta}_1^{(1)}(t)$	0.0012	0.0097
$\widehat{\beta}_1^{(2)}(t)$	0.0011	0.0090
$\widehat{\beta}_1^{(3)}(t)$	0.0011	0.0114
$\widehat{\beta}_2^{(1)}(t)$	0.0009	0.0092
$\widehat{\beta}_2^{(2)}(t)$	0.0008	0.0093
$\widehat{\beta}_2^{(3)}(t)$	0.0010	0.0098
$\widehat{\psi}_1^{(1)}(t)$	0.0600	0.0752
$\widehat{\psi}_1^{(2)}(t)$	0.0612	0.0689
$\widehat{\psi}_1^{(3)}(t)$	0.0616	0.0948
$\widehat{\psi}_2^{(1)}(t)$	0.0969	0.1401
$\widehat{\psi}_2^{(2)}(t)$	0.0929	0.1335
$\widehat{\psi}_2^{(3)}(t)$	0.1006	0.2084
$\widehat{\psi}_3^{(1)}(t)$	0.0649	0.2304
$\widehat{\psi}_3^{(2)}(t)$	0.0599	0.2700
$\widehat{\psi}_3^{(3)}(t)$	0.0559	0.1585
$\widehat{U}_i^{(1)}(t)$	0.0309	0.2417
$\widehat{U}_i^{(2)}(t)$	0.0294	0.2581
$\widehat{U}_i^{(3)}(t)$	0.0327	0.2605
$\widehat{Y}_i^{(1)}(t)$	0.0292	0.2232
$\widehat{Y}_i^{(2)}(t)$	0.0247	0.1923
$\widehat{Y}_i^{(3)}(t)$	0.0349	0.2146
	MSE	
$\widehat{\alpha}_1$	0.1061	0.1080
$\widehat{\alpha}_2$	0.0741	0.0811
$\widehat{\alpha}_3$	0.0060	0.0801
$\widehat{\nu}$	0.0054	0.0054
$\widehat{\sigma}^{2(1)}$	0.0021	0.0024
$\widehat{\sigma}^{2(2)}$	0.0023	0.0025
$\widehat{\sigma}^{2(3)}$	0.0028	0.0028

Table S3: The mean MSDE of VCFs, deviations $\widehat{U}_i^{(j)}(t)$ and predictions $\widehat{Y}_i^{(j)}(t)$, along with coverage probabilities (CPs, %) and length of the 95% credible bands for simulation cases with aligned and misaligned time points and varying measurement error variance from MV-VCM, MV-VCTM, and MV-VCSTM, respectively. Results are based on two-dimensional multivariate response with $n = 49$ regions and 200 Monte Carlo runs.

Time points	**Aligned**						**Misaligned**					
Noise level	0.2			2			0.2			2		
Model	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM
						MSDE						
$\widehat{\beta}_0^{(1)}(t)$	0.6173	0.3292	0.0033	0.6258	0.3584	0.0178	0.6159	0.3759	0.0130	0.6226	0.4843	0.0590
$\widehat{\beta}_0^{(2)}(t)$	0.6466	0.3260	0.0031	0.6502	0.3487	0.0236	0.6770	0.3700	0.0107	0.6772	0.4912	0.0551
$\widehat{\beta}_1^{(1)}(t)$	0.0074	0.0086	0.0017	0.0139	0.0184	0.0096	0.0211	0.0052	0.0038	0.0149	0.0150	0.0095
$\widehat{\beta}_1^{(2)}(t)$	0.0027	0.0079	0.0013	0.0109	0.0199	0.0093	0.0184	0.0050	0.0036	0.0131	0.0145	0.0096
$\widehat{\beta}_2^{(1)}(t)$	0.0024	0.0040	0.0011	0.0181	0.0397	0.0126	0.0050	0.0072	0.0038	0.0156	0.0172	0.0089
$\widehat{\beta}_2^{(2)}(t)$	0.0020	0.0035	0.0009	0.0150	0.0367	0.0093	0.0041	0.0063	0.0034	0.0143	0.0150	0.0079
$\widehat{U}_i^{(1)}(t)$	/	0.5889	0.0465	/	0.6087	0.3447	/	0.6260	0.2234	/	0.8880	0.3773
$\widehat{U}_i^{(2)}(t)$	/	0.5906	0.0470	/	0.6088	0.3516	/	0.6242	0.2207	/	0.8996	0.3623
$\widehat{Y}_i^{(1)}(t)$	0.7536	0.0289	0.0251	0.7654	0.3537	0.3499	0.8235	0.0429	0.0398	0.8317	0.3104	0.2693
$\widehat{Y}_i^{(2)}(t)$	0.6079	0.0229	0.0205	0.6222	0.2207	0.2152	0.7064	0.0311	0.0278	0.7194	0.2389	0.2241
						CP (%)						
$\widehat{\beta}_0^{(1)}(t)$	21.54	100	100	29.79	98	98.50	20.13	100	100	31.02	97.50	98
$\widehat{\beta}_0^{(2)}(t)$	18.04	99.50	100	25.52	100	100	16.33	100	100	25.09	98.50	99.50
$\widehat{\beta}_1^{(1)}(t)$	99.96	100	100	99.92	98	99	99.96	100	100	99.92	99	99.50
$\widehat{\beta}_1^{(2)}(t)$	100	100	100	100	100	100	100	100	100	99.92	99.50	99.50
$\widehat{\beta}_2^{(1)}(t)$	100	100	100	99.94	100	100	100	100	100	99.94	100	100
$\widehat{\beta}_2^{(2)}(t)$	100	100	100	100	100	100	100	100	100	100	100	100
						Length						
$\widehat{\beta}_0^{(1)}(t)$	0.554	3.446	1.008	0.896	2.167	1.067	0.608	3.520	1.223	1.002	4.205	1.767
$\widehat{\beta}_0^{(2)}(t)$	0.579	4.177	0.900	0.945	2.946	1.335	0.613	4.354	1.744	1.023	4.613	1.967
$\widehat{\beta}_1^{(1)}(t)$	0.430	0.850	0.366	0.764	1.393	0.793	0.314	0.914	0.624	0.571	1.408	0.890
$\widehat{\beta}_1^{(2)}(t)$	0.501	0.900	0.448	0.796	1.354	0.813	0.329	0.963	0.756	0.572	1.525	0.948
$\widehat{\beta}_2^{(1)}(t)$	0.402	0.756	0.397	0.712	1.770	0.994	0.430	0.902	0.670	0.622	1.489	0.943
$\widehat{\beta}_2^{(2)}(t)$	0.405	0.678	0.386	0.714	1.704	0.876	0.346	0.844	0.633	0.560	1.448	0.877

Table S4: The mean MSDE of VCFs, deviations $\widehat{U}_i^{(j)}(t)$ and predictions $\widehat{Y}_i^{(j)}(t)$, along with coverage probabilities (CPs, %) of the 95% credible bands and the corresponding length when $n = 49$ regions from MV-VCM, MV-VCTM, and MV-VCSTM, respectively. Results are based on three-dimensional multivariate response with aligned time points and 200 Monte Carlo runs.

Number of regions, n	**49 regions**					
Noise level, σ^2	0.2			2		
Model	MV-VCM	MV-VCTM	MV-VCSTM	MV-VCM	MV-VCTM	MV-VCSTM
			MSDE			
$\widehat{\beta}_0^{(1)}(t)$	0.4664	0.0155	0.0047	0.6806	0.1079	0.0576
$\widehat{\beta}_0^{(2)}(t)$	0.5008	0.0304	0.0059	0.6808	0.1423	0.0651
$\widehat{\beta}_0^{(3)}(t)$	0.4566	0.0189	0.0097	0.6676	0.1582	0.0582
$\widehat{\beta}_1^{(1)}(t)$	0.0142	0.0021	0.0012	0.0175	0.0125	0.0097
$\widehat{\beta}_1^{(2)}(t)$	0.0200	0.0021	0.0011	0.0202	0.0114	0.0090
$\widehat{\beta}_1^{(3)}(t)$	0.0139	0.0020	0.0011	0.0392	0.0187	0.0114
$\widehat{\beta}_2^{(1)}(t)$	0.0048	0.0014	0.0009	0.0142	0.0128	0.0092
$\widehat{\beta}_2^{(2)}(t)$	0.0089	0.0015	0.0008	0.0157	0.0144	0.0093
$\widehat{\beta}_2^{(3)}(t)$	0.0119	0.0017	0.0010	0.0145	0.0131	0.0098
$\widehat{U}_i^{(1)}(t)$	/	0.0661	0.0309	/	0.3901	0.2417
$\widehat{U}_i^{(2)}(t)$	/	0.0651	0.0294	/	0.3899	0.2581
$\widehat{U}_i^{(3)}(t)$	/	0.0711	0.0327	/	0.4155	0.2605
$\widehat{Y}_i^{(1)}(t)$	0.2691	0.0386	0.0292	0.3365	0.2854	0.2232
$\widehat{Y}_i^{(2)}(t)$	0.2788	0.0256	0.0247	0.3101	0.2328	0.1923
$\widehat{Y}_i^{(3)}(t)$	0.2792	0.0456	0.0349	0.3431	0.2471	0.2146
			CP (%)			
$\widehat{\beta}_0^{(1)}(t)$	16.73	99.50	99.50	33.83	100	100
$\widehat{\beta}_0^{(2)}(t)$	11.98	99.50	99.50	26.10	100	100
$\widehat{\beta}_0^{(3)}(t)$	27.00	99.50	100	33.42	100	100
$\widehat{\beta}_1^{(1)}(t)$	98.31	100	100	99.94	100	100
$\widehat{\beta}_1^{(2)}(t)$	96.50	100	100	99.98	100	100
$\widehat{\beta}_1^{(3)}(t)$	93.48	100	100	99.46	100	100
$\widehat{\beta}_2^{(1)}(t)$	99.83	100	100	99.98	100	100
$\widehat{\beta}_2^{(2)}(t)$	99.83	100	100	99.98	100	100
$\widehat{\beta}_2^{(3)}(t)$	98.48	100	100	99.98	100	100
			Length			
$\widehat{\beta}_0^{(1)}(t)$	0.526	1.062	0.668	1.028	2.974	1.971
$\widehat{\beta}_0^{(2)}(t)$	0.482	1.094	0.636	0.845	2.913	1.996
$\widehat{\beta}_0^{(3)}(t)$	0.509	0.976	0.739	0.884	2.459	1.893
$\widehat{\beta}_1^{(1)}(t)$	0.416	0.492	0.320	0.890	1.205	0.894
$\widehat{\beta}_1^{(2)}(t)$	0.490	0.499	0.315	0.934	1.118	0.892
$\widehat{\beta}_1^{(3)}(t)$	0.409	0.438	0.318	0.920	1.310	0.890
$\widehat{\beta}_2^{(1)}(t)$	0.342	0.382	0.282	0.839	1.188	0.908
$\widehat{\beta}_2^{(2)}(t)$	0.449	0.399	0.297	0.888	1.220	0.915
$\widehat{\beta}_2^{(3)}(t)$	0.341	0.420	0.301	0.848	1.173	0.919

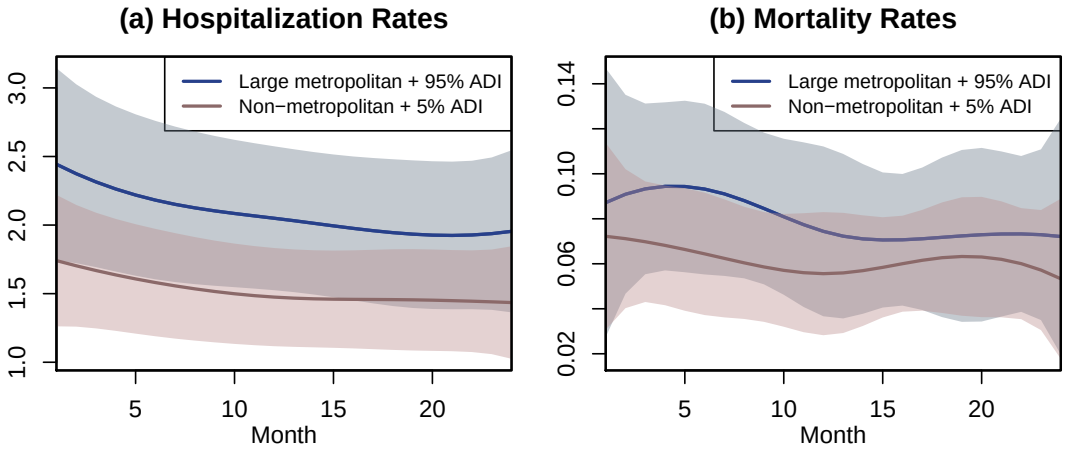

Figure S1: Predicted hospitalization (a) and mortality (b) trajectories for two hypothetical regions, one of large metropolitan with 95% percentile ADI (blue) and the other one of non-metropolitan with 5% percentile ADI (red) while other covariates are kept at mean or reference levels, along with 95% simultaneous credible bands (shaded).

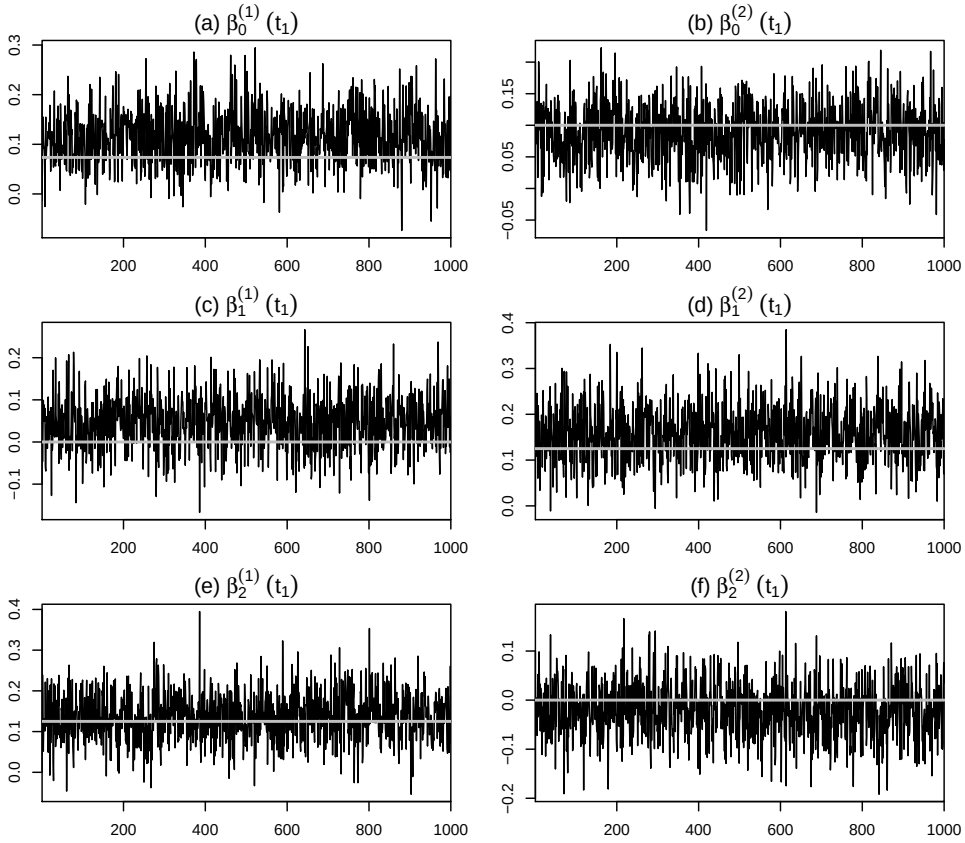

Figure S2: MCMC trace plots for VCFs from Step 5 of the estimation algorithm, where 15000 iterations (5000 burn-in) are used and one out of every 10 samples is returned. The true values of VCFs are represented by the gray horizontal lines. The results are based on two-dimensional multivariate response with aligned time points, $n = 49$ regions and measurement error variance $\sigma^2 = 0.2$.

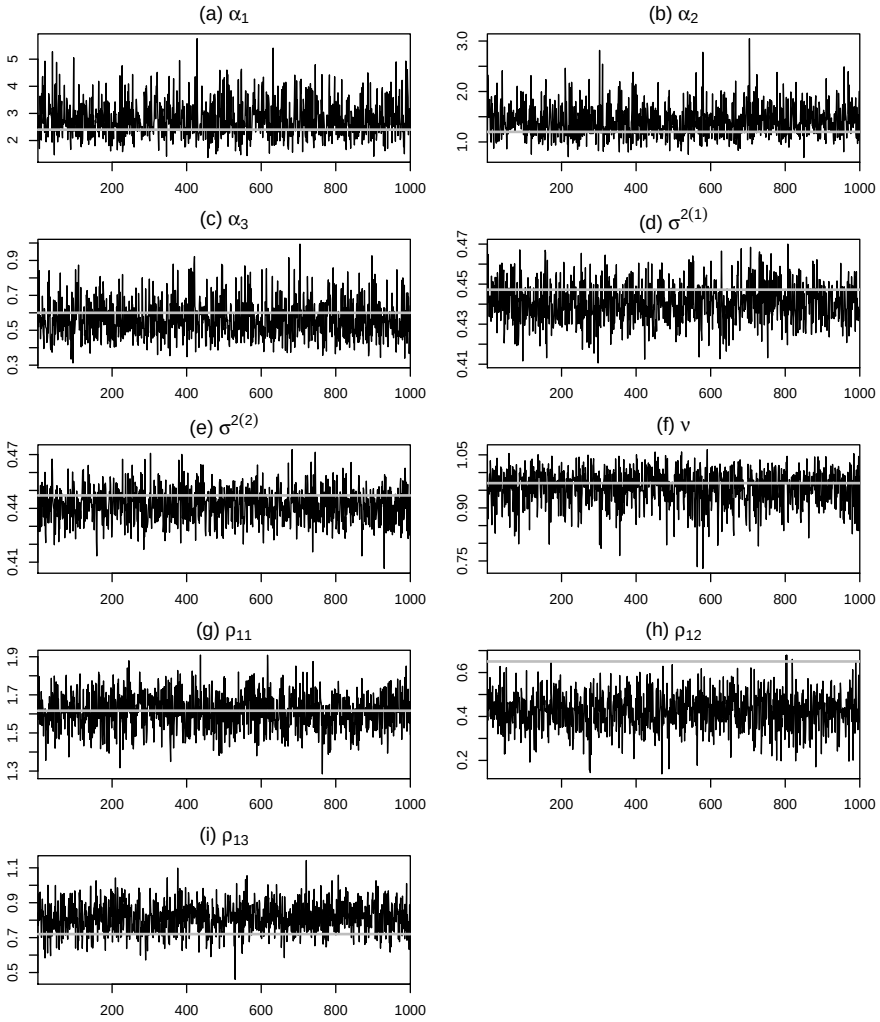

Figure S3: MCMC trace plots for model parameters from Step 5 of the estimation algorithm, where 15000 iterations (5000 burn-in) are used and one out of every 10 samples is returned. The true values of model parameters are represented by the gray horizontal lines. The results are based on two-dimensional multivariate response with aligned time points, $n = 49$ regions and measurement error variance $\sigma^2 = 0.2$.

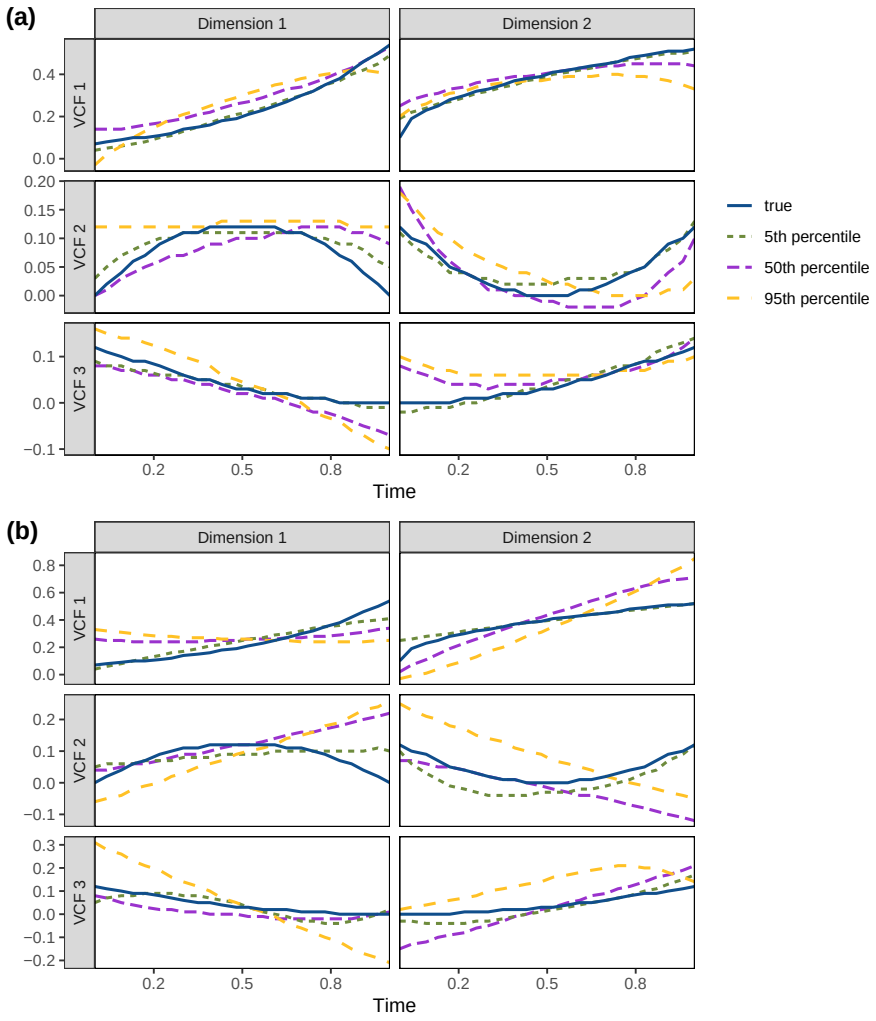

Figure S4: Estimated VCFs from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the simulation set-up with two-dimensional multivariate response with aligned time points, $n = 49$ regions and measurement error variances $\sigma^2 = 0.2$ (a) and $\sigma^2 = 2$ (b). The true VCFs are given in solid blue.

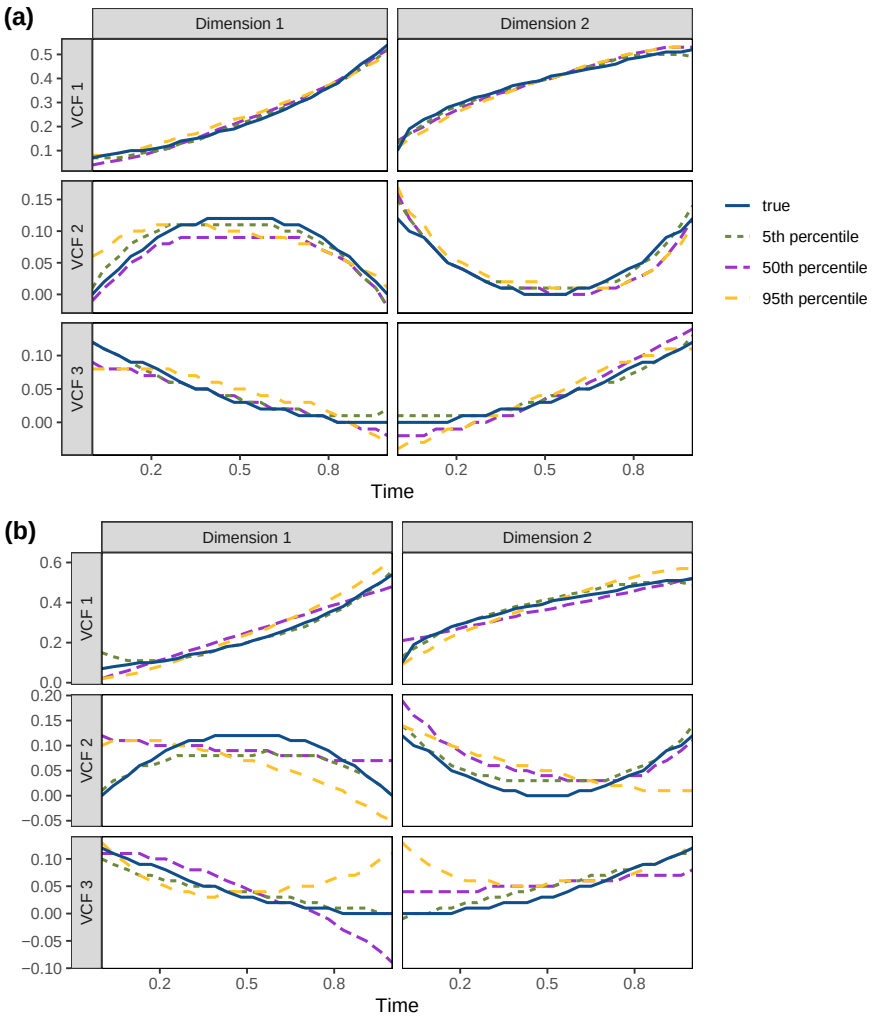

Figure S5: Estimated VCFs from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the simulation set-up with two-dimensional multivariate response with aligned time points, $n = 367$ regions, and measurement error variances $\sigma_j^2 = 0.2$ (a) and $\sigma_j^2 = 2$ (b). The true VCFs are given in solid blue.

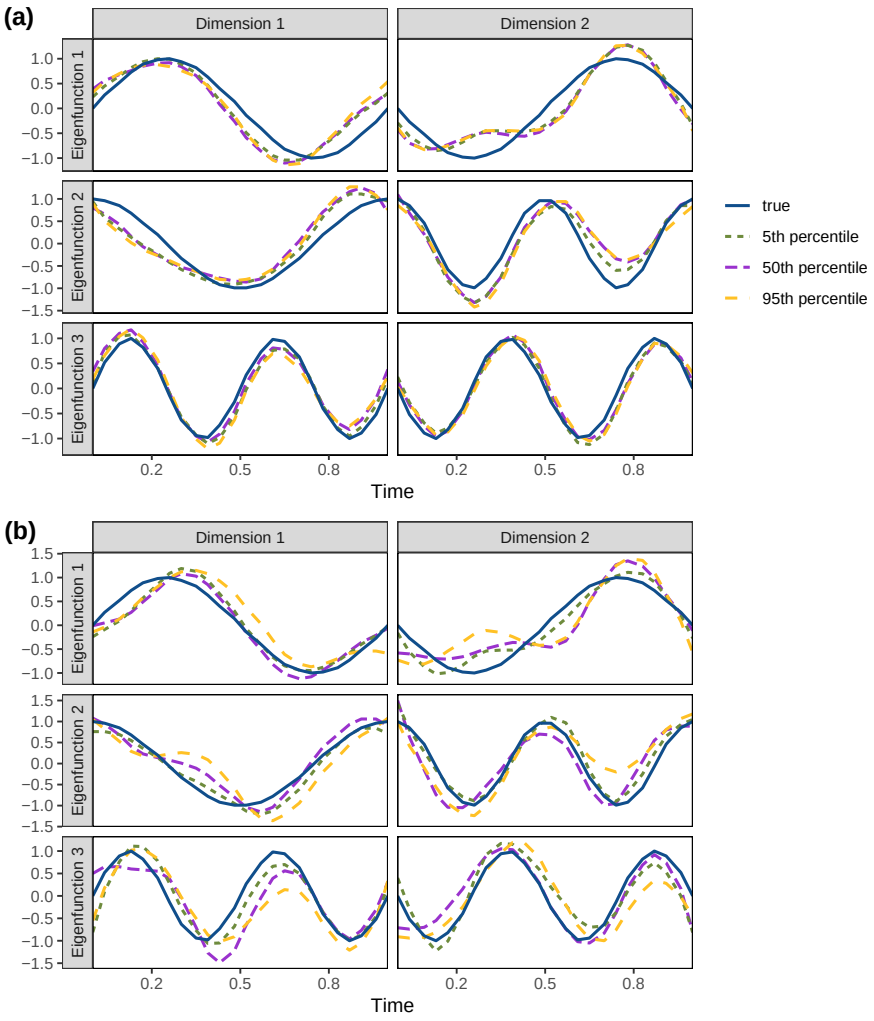

Figure S6: Estimated multivariate eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the set-up with two-dimensional multivariate response with aligned time points, $n = 49$ regions, and $\sigma_j^2 = 0.2$ (a) and $\sigma_j^2 = 2$ (b). The true multivariate eigenfunctions are given in solid blue.

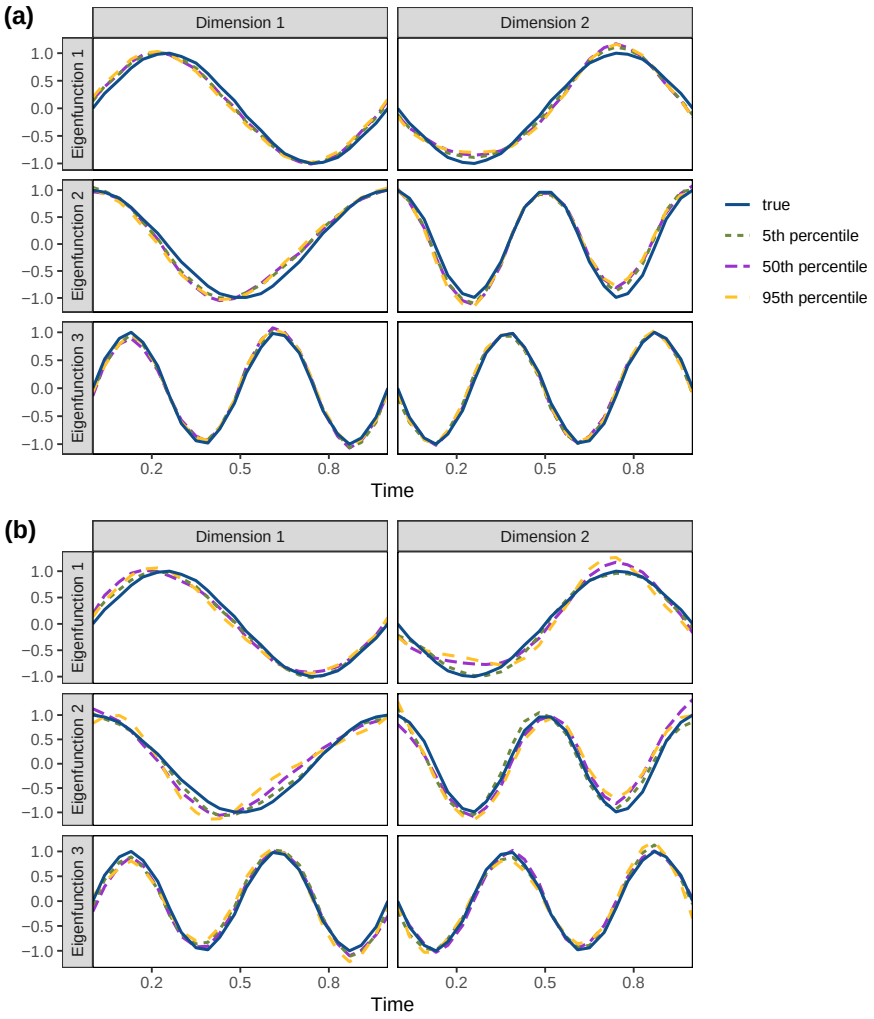

Figure S7: Estimated multivariate eigenfunctions from runs with the 5th, 50th and 95th percentile MSDEs using MV-VCSTM, from the set-up with two-dimensional multivariate response with aligned time points, $n = 367$ regions, and $\sigma_j^2 = 0.2$ (a) and $\sigma_j^2 = 0.2$ (b). The true multivariate eigenfunctions are given in solid blue.