Appendix A: Discussion of Model (2)

Sufficient conditions to establish model (2) of the main manuscript are given in the following Proposition 1, and we consider a more general covariance structure of $\gamma(t, s; t', s)$; that is, we assume

$$\gamma(t, s; t', s) = g(s)\gamma_0(t, t'), \quad s \in \mathbb{R}^n$$

where $\gamma_0(t, t') = \gamma(t, s_0; t', s_0)$ is a covariance function and $g(s)$ is a positive function.

**Proposition 1.** Let $\varepsilon(s, t)$ denote a zero-mean square integrable stochastic process with covariance function satisfying assumption (1). If $\int_0^T \gamma_0(t, t)dt < \infty$, then, for any given $s \in \mathbb{R}^n$ and $t \in [0, T]$, the following expansion holds,

$$\varepsilon(s, t) = g(s)^{1/2} \sum_{j=1}^{\infty} \xi_j(s)\varphi_j(t), \quad (2)$$
where \( \{\xi_j(s)\}_{j=1}^\infty \) is a sequence of stochastic spatial processes that have mean zero and are uncorrelated for a given location \( s \), and \( \{\varphi_j(t)\}_{j=1}^\infty \) is a sequence of fixed eigenfunctions of \( \gamma_0(t,t') \). Moreover, if \( \varepsilon(s,t) \) is a Gaussian process, then \( \{\xi_j(s)\}_{j=1}^\infty \) is a sequence of Gaussian processes.

Proof. Let \( \varepsilon^{(s)}(t) = g(s)^{-1/2}\varepsilon(s,t) \). It follows that \( \varepsilon^{(s)}(t) \) are stochastic processes with a common covariance function \( \gamma_0(t,t') \) at any location \( s \). First, by Karhunen-Loève expansion (Ghanem and Spanos, 1991), for given \( s \in \mathcal{R}_{n^*} \) and \( t \in [0,T] \), we have

\[
\varepsilon^{(s)}(t) = \sum_{j=1}^\infty \xi_j(s)\varphi_j(t),
\]

where \( \{\xi_j(s)\}_{j=1}^\infty \) is a sequence of uncorrelated random variables. For given \( s \in \mathcal{R}_{n^*}, \omega \in \Omega, \varepsilon^{(s)}(t) \) is a deterministic function and therefore, \( (\varepsilon^{(s)}(t), \varphi_j(t)) = \int_0^T \varepsilon^{(s)}(t)\varphi_j(t)dt \) is well-defined in \( L^2([0,T]) \). Moreover, for given \( s \in \mathcal{R}_{n^*}, \), we have

\[
\int_\Omega \int_0^T |\varepsilon^{(s)}(t)\varphi_j(t)|dtdP \leq \left\{ \int_\Omega \int_0^T \varepsilon^{(s)}(t)^2dtdP \int_0^T \varphi_j(t)^2dtdP \right\}^{1/2} \leq \left\{ \int_0^T \gamma_0(t,t)dt \int_\Omega 1dP \right\} < \infty.
\]

By Fubini’s Theorem, \( \int_0^T \varepsilon^{(s)}(t)\varphi_j(t)dt \) is measurable in \( (\Omega, \mathcal{F}, P) \), and therefore, we have \( \xi_j(s) = \left( \sum_{j=1}^\infty \xi_j(s)\varphi_j(t), \varphi_j(t) \right) = \int_0^T \varepsilon^{(s)}(t)\varphi_j(t)dt \) is a stochastic process.

Next, we show that if \( \varepsilon(s,t) \) is a Gaussian process, then \( \xi_j(s) \) is a Gaussian process; that is, for any finite integer \( m > 0 \), \( (\xi_j(s_1), \ldots, \xi_j(s_m))^T \) is a multivariate Gaussian random vector. Let

\[
\xi_j(s_k)_n = \sum_{i=1}^n \varepsilon^{(s_k)}(iT/n)\varphi_j(iT/n)(T/n),
\]

for \( k = 1, \ldots, m \), and by definition, \( \xi_j(s_k)_n \to \xi_j(s_k) \) for each \( s_k \) as \( n \to \infty \). Since \( \varepsilon^{(s)}(t) \) is a Gaussian process in \( \mathcal{R}_{n^*} \times [0,T] \), \( (\xi_j(s_1)_n, \ldots, \xi_j(s_m)_n)^T \) is a multivariate Gaussian random vector, for any finite integer \( n > 0 \), with the characteristic function

\[
E\left\{ e^{\sum_{k=1}^m \tau_k \xi_j(s_k)_n} \right\} = e^{-(1/2)\tau^T \Pi_n \tau},
\]
where \( \tau = (\tau_1, \ldots, \tau_m)^T \), and \( \Pi_n \) is an \( m \times m \) matrix whose \((k, k')\)th component is

\[
\left( \frac{T^2}{n^2} \right) (\varphi_j(T/n), \ldots, \varphi_j(T)) \left[ \text{cov}(\varepsilon(s_k)(iT/n), \varepsilon(s_{k'})(iT/n)) \right]_{i,i'=1}^n
\]

\[
(\varphi_j(T/n), \ldots, \varphi_j(T))^T.
\]  

As \( n \to \infty \), the term in (4) \( \to \int_0^T \int_0^T \gamma_0(t, s_k; t', s_{k'}) \varphi_j(t) \varphi_j(t') dt dt' \), where \( \gamma_0(t, s; t', s') = \text{cov}\{\varepsilon(s)(t), \varepsilon(s')(t')\} \) is the covariance function of \( \varepsilon(s)(t) \).

Since \( e^{i \sum_{k=1}^m \tau_k \xi_j(s_k)} \to e^{i \sum_{k=1}^m \tau_k \xi_j(s_k)} \) as \( n \to \infty \), and \( |e^{i \sum_{k=1}^m \tau_k \xi_j(s_k)}| < 2 \), by the dominated convergence theorem, the characteristic function of \((\xi_j(s_1), \ldots, \xi_j(s_m))^T \)

\[
E \left\{ e^{i \sum_{k=1}^m \tau_k \xi_j(s_k)} \right\} = \lim_{n \to \infty} E \left\{ e^{i \sum_{k=1}^m \tau_k \xi_j(s_k)} \right\} = e^{- (1/2) \tau^T \Pi \tau}
\]

where \( \Pi = \left[ \int_0^T \int_0^T \gamma_0(t, s_k; t', s_{k'}) \varphi_j(t) \varphi_j(t') dt dt' \right]_{k,k'=1}^m \). Thus, the result of model 1 holds.

\[ \square \]

Since \( T \) is fixed, the condition \( \int_0^T \gamma_0(t, t) dt < \infty \) is satisfied if the variance of \( \varepsilon(s, t) \) is bounded over \([0, T]\). Without replications of the spatial process, \( g(s) \) cannot be estimated. However, if \( g(s) \) can be estimated using say prior information, then a more general, spatially varying covariance function \( \gamma(t, s; t', s) \) would be possible. In the following proof, we assume \( g(s) \) is known, and let \( g(s) = 1 \) for \( s \in \mathcal{R}_n \).

**Appendix B: Proof of Theorem 1**

Consider the covariance function \( \gamma_0(t, t') \) as defined in Section 2. Its Karhunen-Loève expansion can be written as \( \gamma_0(t, t') = \sum_{j=1}^\infty \lambda_j \varphi_j(t) \varphi_j(t') \), where \( \lambda_j \) and \( \varphi_j(t) \) are the eigenvalues and eigenfunctions of \( \gamma_0(t, t') \). The sample covariance function can be written as \( \tilde{\gamma}_0(t, t') = n^{-1} \sum_{i=1}^n \{ \varepsilon(s_i, t) + v(s_i, t) \} \{ \varepsilon(s_i, t') + v(s_i, t') \} \). Now, let \( \mathcal{T} \) denote the collection of functions defined on \([0, T]\) with square integrable second-order derivatives; that is, \( \mathcal{T} = \{ f(t) : f''(t) \in L^2([0, T]) \} \). In fact, \( \mathcal{T} \) is a linear subspace of \( L^2[0, T] \). For \( f(t) \in \mathcal{T} \), we define two operators \( \Gamma \) and \( \tilde{\Gamma} \),

\[
\Gamma f(t) = \int_0^T \gamma_0(t, t') f(t') dt, \quad \tilde{\Gamma} f(t) = \int_0^T \tilde{\gamma}_0(t, t') f(t') dt.
\]
Moreover, let $P_k f = f - \sum_{j=1}^{k-1} (f, \varphi_j) \varphi_j$ be the projection of $f(t)$ onto the space perpendicular to $\{\varphi_1(t), \ldots, \varphi_{k-1}(t)\}$ by the inner product $(\cdot, \cdot)$, and $\tilde{P}_k f = f - \sum_{j=1}^{k-1} (f, \tilde{\varphi}_j) \tilde{\varphi}_j$ be the projection of $f(t)$ onto the space perpendicular to $\{\tilde{\varphi}_1(t), \ldots, \tilde{\varphi}_{k-1}(t)\}$ by the inner product $(\cdot, \cdot)$. Finally, by definition, we have $\|\varphi_j(t)\| = \|\tilde{\varphi}_j(t)\| = \|\tilde{\varphi}_j(t)\| = 1$.

In geostatistics, asymptotic results depend on the asymptotic framework adopted. For example, Zhang (2004) showed that under the infill asymptotics (i.e., increasingly dense sampling locations in a fixed spatial domain), maximum likelihood does not yield consistent estimates for all parameters; also see Ying (1991), Lahiri (1996), Loh (2005). Here, we focus on the increasing domain asymptotics such that spatial domain increases at the same rate as the number of sampling locations. Under this increasing domain framework, consistency of $\hat{\gamma}_0(t, t')$ can be established.

Now, we give a proof of Theorem 1.

**Proof.** First, we show that, under (A.1)–(A.4), we have

$$\sup_{\|f\|_a \leq 1} \|\hat{\gamma} - \Gamma\| f \rightarrow_P 0. \tag{5}$$

Denote $\hat{\Gamma} f = n^{-1} \sum_{i=1}^{n} \int_0^T \epsilon(s_i, t) f(t) dt$, $\hat{\Gamma}_2 f = n^{-1} \sum_{i=1}^{n} \int_0^T \epsilon(s_i, t) v(s_i, t') f(t) dt$, $\hat{\Gamma}_3 f = n^{-1} \sum_{i=1}^{n} \int_0^T v(s_i, t) \epsilon(s_i, t') f(t) dt$, and $\hat{\Gamma}_4 f = n^{-1} \sum_{i=1}^{n} \int_0^T v(s_i, t) v(s_i, t') f(t) dt$. By the following inequality

$$\sup_{\|f\|_a \leq 1} \|\hat{\Gamma} - \Gamma\| f \leq \sup_{\|f\|_a \leq 1} \|\hat{\Gamma}_1 - \Gamma\| f + \sup_{\|f\|_a \leq 1} \|\hat{\Gamma}_2 f\| + \sup_{\|f\|_a \leq 1} \|\hat{\Gamma}_3 f\| + \sup_{\|f\|_a \leq 1} \|\hat{\Gamma}_4 f\| \equiv (I_1) + (I_2) + (I_3) + (I_4),$$

it suffices to show that $(I_i) \rightarrow_P 0$ for each $i = 1, 2, 3, 4$.

Proof of $(I_1) \rightarrow_P 0$. Note that $L^2([0, T])$ is a separable space (Billingsley, 1995), and an orthonormal basis in $L^2([0, T])$ can be constructed by expanding $\{\varphi_j(t) : j = 1, \ldots, \infty\}$ to $\{\varphi_j(t) : j = 1, \ldots, \infty\} \cup \{\phi_j(t) : j = 1, \ldots, \infty\}$. For any $f(t) \in T \subseteq L^2([0, T])$, we have $f(t) = \sum_{j=1}^{\infty} \alpha_j(f) \varphi_j(t) + \sum_{k=1}^{\infty} \beta_k(f) \phi_k(t)$, where $\alpha_j(f) = (f, \varphi_j)$ and $\beta_k(f) = (f, \phi_k)$.\]
Moreover, for \( k = 1, 2, \ldots \), we have

\[
\Gamma \phi_k(t) = \int_0^T \gamma(t, t') \phi_k(t) dt = \sum_{j=1}^{\infty} \lambda_j \varphi_j(t') \int_0^T \varphi_j(t) \phi_k(t) dt = 0,
\]

\[
\hat{\Gamma}_1 \phi_k(t) = n^{-1} \sum_{i=1}^{n} \varepsilon(s_i, t') \int_0^T \varepsilon(s_i, t) \phi_k(t) dt
\]

\[
= n^{-1} \sum_{i=1}^{n} \varepsilon(s_i, t') \sum_{j=1}^{\infty} \xi_j(s_i) \int_0^T \varphi_j(t) \phi_k(t) dt = 0.
\]

For any \( f \) with \( \|f(t)\|_a \leq 1 \), we have \( \alpha_j(f)^2 \leq \|f(t)\| \leq \|f(t)\|_a \leq 1 \). Thus, we also have

\[
(I_1)^2 = \sup_{\|f\|_a \leq 1} \| \hat{(\Gamma_1 - \Gamma)} f(t) \|^2
\]

\[
= \sum_{j=1}^{\infty} \alpha_j(f)^2 \| \hat{(\Gamma_1 - \Gamma)} \varphi_j(t) \|^2 \leq \sum_{j=1}^{\infty} \| \hat{(\Gamma_1 - \Gamma)} \varphi_j(t) \|^2.
\]

Therefore, it suffices to show \( \sum_{j=1}^{\infty} E \| \hat{(\Gamma_1 - \Gamma)} \varphi_j(t) \|^2 \to 0 \).

Next, we quantify \( E \| \hat{(\Gamma_1 - \Gamma)} \varphi_j(t) \|^2 \). By the fact that \( \Gamma(t, t') \varphi_j(t) = \lambda_j \varphi_j(t') \) and \( \hat{\Gamma}_1(t, t') \varphi_j(t) = n^{-1} \sum_{i=1}^{n} \varepsilon(s_i, t') \xi_j(s_i) \), we have

\[
E \| \hat{(\Gamma_1 - \Gamma)} \varphi_j(t) \|^2 = E \int \left( \sum_{i=1}^{n} \varepsilon(s_i, t') \xi_j(s_i) - \lambda_j \varphi_j(t') \right)^2 dt'
\]

\[
= n^{-2} E \int \left( \sum_{i=1}^{n} \varepsilon(s_i, t') \xi_j(s_i) \right)^2 dt' - 2n^{-1} E \int \lambda_j \varphi_j(t') \left( \sum_{i=1}^{n} \varepsilon(s_i, t') \xi_j(s_i) \right) dt'
\]

\[
+ \lambda_j^2 \int \varphi_j(t'^2) dt' \equiv n^{-2}(II_1) - 2n^{-1}(II_2) + \lambda_j^2.
\]

Straightforward calculation yields simplification of \( (II_1) \) and \( (II_2) \) such that

\[
(II_1) = \sum_{i=1}^{n} \sum_{i'=1}^{n} E \int \varepsilon(s_i, t') \xi_j(s_i) \varepsilon(s_{i'}, t') \xi_j(s_{i'}) dt'
\]

\[
= \sum_{i=1}^{n} \sum_{i'=1}^{n} E \left\{ \xi_j(s_i) \xi_j(s_{i'}) \int \varepsilon(s_i, t') \varepsilon(s_{i'}, t') dt' \right\}
\]

\[
= \sum_{i=1}^{n} \sum_{i'=1}^{n} E \left\{ \xi_j(s_i) \xi_j(s_{i'}) \int \left( \sum_{j'=1}^{\infty} \xi_j(s_i) \varphi_j(t') \right) \left( \sum_{j''=1}^{\infty} \xi_j(s_{i'}) \varphi_{j''}(t') \right) dt' \right\}
\]

\[
= \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j'=1}^{\infty} E \left\{ \xi_j(s_i) \xi_j(s_{i'}) \xi_j(s_i) \xi_j(s_{i'}) \right\}
\]
Combining both equations above, we have

\[
(I1_2) = \lambda_j \sum_{i=1}^{n} E \int \varepsilon(s_i, t') \xi_j(s_i) \varphi_j(t') dt'
\]

\[
= \lambda_j \sum_{i=1}^{n} E \int \left( \sum_{j'=1}^{\infty} \xi_j(s_i) \varphi_j(t') \right) \xi_j(s_i) \varphi_j(t') dt'
\]

\[
= \lambda_j \sum_{i=1}^{n} \sum_{j'=1}^{\infty} E(\xi_j'(s_i)\xi_j(s_i)) \int \varphi_j'(t')\varphi_j(t') dt' = n\lambda_j^2 \int \varphi_j(t')^2 dt' = n\lambda_j^2.
\]

Combining both equations above, we have

\[
E\|\hat{\Gamma}_1 - \Gamma\|\varphi_j(t)\|^2 = n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j'=1}^{\infty} E\{\xi_j(s_i)\xi_{j'}(s_i)\xi_j(s_{i'})\xi_{j'}(s_{i'})\} - \lambda_j^2.
\]

\[
n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j'=1}^{\infty} E\{\xi_j(s_i)\xi_{j'}(s_i)\xi_j(s_{i'})\xi_{j'}(s_{i'})\} + n^{-2} \left[ \sum_{i=1}^{n} \sum_{i'=1}^{n} E\{\xi_j(s_i)^2\xi_{j'}(s_{i'})^2\} - \lambda_j^2 \right].
\]

Further, for a Gaussian process \(\xi_j(s)\), we have

\[
E\{\xi_j(s_i)^2\xi_j(s_{i'})^2\} = 2E\{\xi_j(s_i)\xi_j(s_{i'})\}^2 + \lambda_j^2.
\]

Thus,

\[
E\|\hat{\Gamma}_1 - \Gamma\|\varphi_j(t)\|^2 = n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j'=1}^{\infty} E\{\xi_j(s_i)\xi_j(s_{i'})\} E\{\xi_{j'}(s_i)\xi_{j'}(s_{i'})\}
\]

\[
+ n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} [E\{\xi_j(s_i)\xi_j(s_{i'})\}]^2.
\]

Finally, we have

\[
\sum_{j=1}^{\infty} E\|\hat{\Gamma}_1 - \Gamma\|\varphi_j(t)\|^2 \leq n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{\infty} \sum_{j'=1}^{\infty} E\{\xi_j(s_i)\xi_j(s_{i'})\} E\{\xi_{j'}(s_i)\xi_{j'}(s_{i'})\}
\]

\[
+ n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{\infty} [E\{\xi_j(s_i)\xi_j(s_{i'})\}]^2,
\]

and therefore \((I_1) \xrightarrow{P} 0\).
Proof of \((I_i) \xrightarrow{P} 0\), for \(i = 2, 3\). For a function \(g(t, t') \in L^2([0, T] \times [0, T])\), it is easy to show that

\[
\left\| \int_0^T g(t, t') f(t) dt \right\|^2 = \int_0^T \left\{ \int_0^T g(t, t') f(t) dt \right\}^2 dt' \\
\leq \int_0^T \left\{ \int_0^T g(t, t')^2 dt \right\} \left\{ \int_0^T f(t)^2 dt \right\} dt' = \|f(t)\|^2 \int_0^T \int_0^T g(t, t')^2 dt dt'.
\]

Therefore, we have

\[
\sup_{\|f\| \leq 1} \left\| \int_0^T g(t, t') f(t) dt \right\|^2 \leq \int_0^T \int_0^T g(t, t')^2 dt dt'. \tag{7}
\]

By (7), we have

\[
E(I_2)^2 = E(I_3)^2 \leq n^{-2} \int_0^T \int_0^T \left\{ \sum_{i=1}^n \sum_{i'=1}^n \varepsilon(s_i, t') \varepsilon(s_i, t) \right\}^2 dt dt' \\
= n^{-2} \sum_{i=1}^n \sum_{i'=1}^n \int_0^T \int_0^T \varepsilon(s_i, t') \varepsilon(s_i, t) \varepsilon(s_{i'}, t') \varepsilon(s_{i'}, t) dt dt' \\
= n^{-2} \sum_{i=1}^n \sum_{i'=1}^n \int_0^T \int_0^T \varepsilon(s_i, t') \varepsilon(s_i, t) \varepsilon(s_{i'}, t') \varepsilon(s_{i'}, t) dt dt' \\
= n^{-2} \sum_{i=1}^n \int_0^T \int_0^T \varepsilon(s_i, t') \varepsilon(s_i, t) \varepsilon(s_{i'}, t') \varepsilon(s_{i'}, t) dt dt' \\
= n^{-2} \sum_{i=1}^n \int_0^T \gamma_0(t, t) dt \rightarrow 0.
\]

Thus, \((I_i) \xrightarrow{P} 0\) holds for \(i = 2, 3\).

Proof of \((I_4) \xrightarrow{P} 0\). By (7), we have

\[
E(I_4)^2 \leq n^{-2} \int_0^T \int_0^T \left\{ \sum_{i=1}^n \varepsilon(s_i, t') \varepsilon(s_i, t) \right\}^2 dt dt' \\
= n^{-2} \sum_{i=1}^n \sum_{i'=1}^n \int_0^T \int_0^T \varepsilon(s_i, t') \varepsilon(s_i, t) \varepsilon(s_{i'}, t') \varepsilon(s_{i'}, t) dt dt' \\
= n^{-2} \sum_{i=1}^n \int_0^T \int_0^T \varepsilon(s_i, t') \varepsilon(s_i, t) \varepsilon(s_{i'}, t') \varepsilon(s_{i'}, t) dt dt' \\
+ n^{-2} \sum_{i \neq i'} \int_0^T \gamma_0(t, t) \gamma_0(t', t') \varepsilon(t, t') dt dt' = n^{-1} \left\{ \int_0^T \gamma_0(t, t) dt \right\}^2 \rightarrow 0.
\]
Thus, \((I_4) \xrightarrow{P} 0\) and (5) is established.

Define the statement \(S_j\): as \(n \to \infty\),

\[
\alpha[\varphi_j, \varphi_j] \xrightarrow{P} 0; \quad (8a)
\]

\[
(\varphi_j, \varphi_j)^2 \xrightarrow{P} 1; \quad (8b)
\]

\[
\lambda_j \xrightarrow{P} \lambda_j. \quad (8c)
\]

Since \(\|\varphi_j\|^2 = 1\), (8a) and (8b) are equivalent to (9a) and (9b), respectively,

\[
\|\varphi_j\|^2 \xrightarrow{P} 1; \quad (9a)
\]

\[
\|\varphi_j(t) - \varphi_j(t)\| \xrightarrow{P} 0. \quad (9b)
\]

First, we show that if \(S_j\) holds for all \(j < k\), then as \(n \to \infty\),

\[
\sup_{\|f\| \leq 1} |(P_k^1 f, \Gamma P_k^1 f) - (\tilde{P}_k^1 f, \tilde{\Gamma} P_k^1 f)| \xrightarrow{P} 0. \quad (10)
\]

Note that

\[
\sup_{\|f\| \leq 1} \|(f, \varphi_j) \varphi_j - (f, \varphi_j) \varphi_j\| \leq \sup_{\|f\| \leq 1} \|(f, \varphi_j) \varphi_j - (f, \varphi_j) \varphi_j\|
\]

\[
+ \sup_{\|f\| \leq 1} \|(f, \varphi_j) \varphi_j - (f, \varphi_j) \varphi_j\| + \sup_{\|f\| \leq 1} \|(f, \varphi_j) \varphi_j - (f, \varphi_j) \varphi_j\|
\]

\[
\equiv (III_1) + (III_2) + (III_3).
\]

By (9a) and (9b), we have

\[
(III_1) = \sup_{\|f\| \leq 1} \|(f, \varphi_j) \varphi_j - (f, \varphi_j) \varphi_j\| + \sup_{\|f\| \leq 1} \|(f, \varphi_j) \varphi_j - (f, \varphi_j) \varphi_j\|
\]

\[
\leq \sup_{\|f\| \leq 1} \|f(t)\|^1/2 \|\varphi_j(t)\|^{1/2} \|\varphi_j - \varphi_j\| + \sup_{\|f\| \leq 1} \|f(t)\|^1/2 \|\varphi_j(t)\|^{1/2} \|\varphi_j - \varphi_j\|
\]

\[
= o_p(1).
\]

By \(\varphi_j(t) = \varphi_j(t)/\|\varphi_j(t)\|\) and (9a), we have

\[
(III_2) = \|(f, \varphi_j) \varphi_j\| \|1 - \|\varphi_j\|^2\| \leq \|f\|^{1/2} \|\varphi_j\|^3/2(1 - \|\varphi_j\|^2) \leq (1 - \|\varphi_j\|^2) = o_p(1).
\]

By the definition of \((\cdot, \cdot)_\alpha\), we have

\[
(III_3) = \sup_{\|f\| \leq 1} \alpha[f, \varphi_j] \|\varphi_j\| \leq \sup_{\|f\| \leq 1} (\alpha[f, f])^{1/2}(\alpha[\varphi_j, \varphi_j])^{1/2} \|\varphi_j\| = o_p(1).
\]
Thus, for \( j < k \), we have \( \sup_{\|f\|_{\alpha} \leq 1} \| (f, \varphi_j) \varphi_j - (f, \tilde{\varphi}_j) \alpha \tilde{\varphi}_j \| = o_p(1) \), and
\[
\sup_{\|f\|_{\alpha} \leq 1} \| (\tilde{P}_k \supset P_k) f \| \leq \sum_{j=1}^{k-1} \sup_{\|f\|_{\alpha} \leq 1} \| (f, \varphi_j) \varphi_j - (f, \tilde{\varphi}_j) \alpha \tilde{\varphi}_j \| \xrightarrow{P} 0. \tag{11}
\]
Also by (5), we have \( \sup_{\|f\|_{\alpha} \leq 1} \| (\tilde{\Gamma} - \Gamma) f \| \xrightarrow{P} 0 \), which implies
\[
\sup_{\|f\|_{\alpha} \leq 1} \| (\tilde{\Gamma} - \Gamma) \tilde{P}_k f \| \xrightarrow{P} 0. \tag{12}
\]
Combining (11) and (12), it follows that
\[
\sup_{\|f\|_{\alpha} \leq 1} \| (P_k f, \Gamma P_k f) - (\tilde{P}_k f, \tilde{\Gamma} \tilde{P}_k f) \| \leq \sup_{\|f\|_{\alpha} \leq 1} \| (P_k f, \Gamma P_k f) - (\tilde{P}_k f, \Gamma P_k f) \|
\]
\[
+ \sup_{\|f\|_{\alpha} \leq 1} \| (\tilde{P}_k f, \Gamma P_k f) - (\tilde{P}_k f, \tilde{\Gamma} \tilde{P}_k f) \| + \sup_{\|f\|_{\alpha} \leq 1} \| (\tilde{P}_k f, \Gamma \tilde{P}_k f) - (\tilde{P}_k f, \tilde{\Gamma} \tilde{P}_k f) \| \xrightarrow{P} 0,
\]
and thus, (10) is established. The remainder of the proof follows (Silverman, 1996) and we omit the details. \(\square\)

Remarks: To verify the regularity conditions (A.3) and (A.4), we consider the case in which \( \xi_j(s), j = 1, 2, \ldots \), are stationary (detrended) Gaussian process with its covariance function from the Matérn family. That is,
\[
\rho_j(d) = \frac{1}{\Gamma(\nu_j)} \left( \frac{d}{2r_j} \right)^{\nu_j} 2\mathcal{K}_{\nu_j} \left( \frac{d}{r_j} \right),
\]
where \( r_j > 0 \) is a range parameter controlling the rate of autocorrelation decay with lag distance, \( \nu_j > 0 \) is a shape parameter controlling the smoothness of the Gaussian process, and \( \mathcal{K}_{\nu_j}(\cdot) \) is a modified Bessel function of the second kind of order \( \nu_j \). For \( \mathcal{R}^* = [0, n^{1/2}] \times [0, n^{1/2}] \subset \mathbb{R}^2 \), sufficient conditions for (A.3) and (A.4) are

(i) There exist constants \( r > 0 \) and \( \nu > 0 \), such that \( r_j \geq r \) and \( \nu_j \leq \nu \).

(ii) The distance between any two sampling sites is greater than a constant.
Proof. Since \( K_\nu(x) \propto e^{-x}x^{-1/2}\{1 + O(1/x)\} \) when \(|x| \to \infty\), there exist constants \( c_1, c_2 > 0 \), such that for \(|d| > c_1\), we have

\[
\rho_j(d) \leq c_2 d^{\nu-1/2} \exp \left(-\frac{d}{r}\right).
\]

Recall that there exists a constant \( C \), such that \( \sum_{j=1}^{\infty} \lambda_j = C \). By (i), for any two sampling locations \( s_i \) and \( s_{i'} \), we have

\[
\sum_{j=1}^{\infty} E\{\xi_j(s_i)\xi_j(s_{i'})\} \leq \sum_{j=1}^{\infty} \lambda_j c_2 d_{i'i'}^{\nu-1/2} \exp\{-\frac{d_{i'i'}/r}{}} = C c_2 d_{i'i'}^{\nu-1/2} \exp\{-\frac{d_{i'i'}/r}{}} \}
\]

where \( d_{i'i'} = \|s_i - s_{i'}\| \).

For a given sampling location \( s_i \), we have

\[
\sum_{i'=1}^{n} \sum_{j=1}^{\infty} \sum_{j'=1}^{\infty} E\{\xi_j(s_i)\xi_j(s_{i'})\} E\{\xi_{j'}(s_i)\xi_{j'}(s_{i'})\} \leq C^2 c_2^2 \sum_{i'=1}^{n} d_{i'i'}^{\nu-1} \exp\{-2\frac{d_{i'i'}/r}{}}
\]

\[
\sum_{i'=1}^{n} \sum_{j=1}^{\infty} |E\{\xi_j(s_i)\xi_j(s_{i'})\}|^2 \leq \sum_{i'=1}^{n} \sum_{j=1}^{\infty} \lambda_j^2 c_2^2 d_{i'i'}^{\nu-1} \exp\{-2\frac{d_{i'i'}/r}{}} \leq C^2 c_2^2 \sum_{i'=1}^{n} d_{i'i'}^{\nu-1} \exp\{-2\frac{d_{i'i'}/r}{}}.
\]

By (ii), the sampling density of any subset of \( \mathcal{R}_n \subset \mathbb{R}^2 \) is bounded by a constant, say \( \varrho \). Let \( B_m = \{i' : mh < d_{i'i'} \leq (m + 1)h\} \), where \( h \) is independent of \( n \). Thus, \( |B_m| \leq (2m + 1)\varrho \pi h^2 \) and \( \mathcal{R}_n \subset \bigcup_{m=0}^{\lfloor(2n)^{1/2}/h\rfloor} B_m \), where \( \lfloor \cdot \rfloor \) denotes the floor function, and \( |B_m| \) denotes the cardinality of a discrete set \( B_m \).

\[
\sum_{i'=1}^{n} d_{i'i'}^{\nu-1} \exp\{-2\frac{d_{i'i'}/r}{}} \leq \pi c_1^2 \varrho C + \sum_{m=0}^{\lfloor(2n)^{1/2}/h\rfloor} (2m + 1)\varrho \pi h^2 (mh)^{2\nu-1} \exp\{-2\frac{m(h)/r}{}}
\]

\[
\leq \pi c_1^2 \varrho C + \pi \varrho h^2 + \sum_{m=1}^{\infty} 3m\varrho \pi h^2 (mh)^{2\nu-1} \exp\{-2\frac{m(h)/r}{}}.
\]

As \( h \to 0 \), we have \( \sum_{m=1}^{\infty} mh^2 (mh)^{2\nu-1} \exp\{-2\frac{m(h)/r}{}} \to \int_0^{\infty} u^{2\nu} \exp\{-2\frac{u/r}{\}} du = O(1) \).

Therefore, (A.3) and (A.4) hold. \( \square \)
Appendix C: Proof of Theorem 3

Recall the sample covariance function \( \tilde{\gamma}_0(t, t') = n^{-1} \sum_{i=1}^{n} \tilde{y}(s_i, t)\tilde{y}(s_i, t') \), and define the operator \( \hat{\Gamma} \) as \( \hat{\Gamma}f(t) = \int_{0}^{T} \tilde{\gamma}_0(t, t')f(t')dt \) for \( f(t) \in \mathcal{T} \).

**Proof.** The proof of Theorem 3 is similar to that of Theorem 1, but we replace \( \tilde{y}(s, t), \hat{y}_s, \hat{y} \) and \( \hat{\Gamma} \) with \( y(s, t), \bar{y}_s, \bar{y} \) and \( \bar{\Gamma} \), respectively. Further, we replace (5) with Lemma 1 below.

**Lemma 1.** Under (A.1)–(A.4) and (A.11), we have

\[
sup_{\|f\|_\alpha \leq 1} \| (\hat{\Gamma} - \Gamma) f \| \xrightarrow{P} 0.
\]

**Proof.** Since \( \hat{y}(s, t) = y(s, t) - \mu(s, t) + \mu(s, t) - \bar{\mu}(s, t) \), we have

\[
sup_{\|f\|_\alpha \leq 1} \| (\hat{\Gamma} - \Gamma) f \| \leq sup_{\|f\|_\alpha \leq 1} \| \hat{\Gamma} f - \Gamma f \| + sup_{\|f\|_\alpha \leq 1} \| \hat{\Gamma}_2 f \| + sup_{\|f\|_\alpha \leq 1} \| \hat{\Gamma}_3 f \|
\]

where \( \hat{\Gamma}_2 f = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \{ y(s, t) - \mu(s, t) \} \{ \mu(s, t') - \bar{\mu}(s, t') \} f(t')dt \) and \( \hat{\Gamma}_3 f = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \{ \mu(s, t') - \bar{\mu}(s, t') \} \{ \mu(s, t) - \bar{\mu}(s, t) \} f(t')dt \). To show Lemma 1, it suffices to show that \( (IV_i) \xrightarrow{P} 0 \), respectively, for \( i = 1, \ldots, 4 \). By (5), we have \( (IV_1) \xrightarrow{P} 0 \).

Proof of \( (IV_i) \xrightarrow{P} 0, i = 2, 3 \). By (7), we have

\[
E(IV_2)^2 = E(IV_3)^2 \leq n^{-2}E \int_{0}^{T} \int_{0}^{T} \left[ \sum_{i=1}^{n} \{ y(s, t') - \mu(s, t') \} \{ \mu(s, t) - \bar{\mu}(s, t) \} \right]^2 dt'dt'
\]

\[
= n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \int_{0}^{T} \int_{0}^{T} E[\{ y(s, t') - \mu(s, t') \} \{ \mu(s, t) - \bar{\mu}(s, t) \} \{ y(s', t'') - \mu(s', t'') \} \{ \mu(s', t) - \bar{\mu}(s', t) \}] dt'dt'.
\]

Applying Cauchy-Schwarz inequality twice and (A.11), we have

\[
E[\{ y(s, t') - \mu(s, t') \} \{ \mu(s, t) - \bar{\mu}(s, t) \} \{ y(s', t'') - \mu(s', t'') \} \{ \mu(s', t) - \bar{\mu}(s', t) \}] \leq c_n^{1/2} \left[ E\{ y(s, t') - \mu(s, t') \}^4 \right]^{1/4} \left[ E\{ y(s', t'') - \mu(s', t'') \}^4 \right]^{1/4}.
\]

11
Therefore,
\[
E(IV_2)^2 = E(IV_3)^2 \leq c_n^{1/2} \left\{ \int_0^T \left[ E\{y(s_i, t) - \mu(s_i, t)\}^4 \right]^{1/4} dt \right\}^{1/2}
\]
\[
= c_n^{1/2} (3T^4)^{1/8} \left\{ \int_0^T \gamma_0(t)^{3/8} dt \right\}^{1/2} \to 0.
\] (13)

Thus, \( IV_i \xrightarrow{P} 0 \), for \( i = 2, 3 \).

Proof of \( IV_4 \xrightarrow{P} 0 \). Similarly, it is straightforward to show that
\[
E(IV_4)^2 \leq n^{-2} E \int_0^T \int_0^T \left[ \sum_{i=1}^n \{ \mu(s_i, t') - \bar{\mu}(s_i, t') \} \{ \mu(s_i, t) - \bar{\mu}(s_i, t) \} \right]^2 dt dt'
\]
\[
= c_n T^2 \to 0.
\]

Therefore, \( IV_4 \xrightarrow{P} 0 \). \qed

Appendix D: Proof of Theorem 2 and 4

Theorem 2 is a special case of Theorem 4 without the mean trend and parameter \( \beta \). Thus, it suffices to prove Theorem 4.

Proof. By Theorem 2 of Mardia and Marshall (1984), it suffices to show that

(B.1) As \( n \to \infty \), \( \| \Sigma \|_2 = O(1) \), \( \| D_l \Sigma \|_2 = O(1) \), \( \| D_{l'} \Sigma \|_2 = O(1) \) for all \( l, l' = 1, \ldots, qm \).

(B.2) For some \( \delta > 0 \), there exist positive constants \( C_l^* \) such that \( \| D_l \Sigma \|_F^2 \leq C_l^* n^{-1/2 - \delta} \) for \( l = 1, \ldots, qm \).

(B.3) For any \( l, l' = 1, \ldots, qm \), \( a_{ll'} = \lim_{n \to \infty} \{ t_{ll'}^* (t_{ll'}^*)^{-1/2} \} \) exists and \( A^* = [a_{ll'}]_{l,l'=1}^{qm} \) is nonsingular.

(B.4) There exists a positive constant \( c_0^* \), such that \( \| \Sigma^{-1} \|_2 < c_0^* < \infty \).

By (A.12), we have \( \| \Phi^T \Phi \|_2 = O(1) \). Since \( \Sigma = \Phi^T \Lambda \Phi + \sigma^2 I_n \),
\[
\| \Sigma \|_2 \leq \| \Lambda \|_2 \| \Phi^T \Phi \|_2 + \sigma^2 \leq \max_{j=1,\ldots,l} \| \Lambda_j \|_2 + \sigma^2 \leq O(m) O(1) + \sigma^2 = O(1).
\]
Thus, (B.1) is established.

For $D_t \Sigma = \Phi^T D_t \Lambda \Phi$ with $\vartheta_l = \theta_{j,k}$,

$$\|D_t \Sigma\|_2 = \|\Phi^T D_k \Lambda_j \Phi\|_2 = O(m) \|D_k \Lambda_j\|_2.$$  

For $D_{tt'} \Sigma = \Phi^T D_{tt'} \Lambda \Phi$ with $\vartheta_l = \theta_{j,k}$ and $\vartheta_l = \theta_{j,k}$

$$\|D_{tt'} \Sigma\|_2 = \|\Phi^T D_{k\ell} \Lambda_j \Phi\|_2 = O(m) \|D_{k\ell} \Lambda_j\|_2 \text{ for } j = j',$$  

$$\|D_{tt'} \Sigma\|_2 = 0, \text{ for } j \neq j'.$$

Thus, (B.1) is established.

For (B.2), it can be seen through the following equation,

$$\|D_t \Sigma\|_F^2 = \text{tr}\{\Phi^T (D_t \Lambda) \Phi \Phi^T (D_t \Lambda) \Phi \} \geq \text{tr}\{m I_n (D_t \Lambda) m I_n (D_t \Lambda)\}$$

$$= m^2 \sum_{i=1}^n \{\mu_i(D_t \Lambda)\}^2 = m^2 \sum_{j=1}^j \|D_t \Lambda_j\|_F^2 = (m^2 \sum_{j=1}^j 1/C_l) n^{1/2+\delta},$$

for $\vartheta_l = \theta_{j,k}$.

For (B.3), if $\vartheta_l = \theta_{j,k}$, $\vartheta_{l'} = \theta_{j',k'}$ and $j \neq j'$, $t_{l'} = 0$. Therefore, $A^* = \text{diag}\{A^*_1, \ldots, A^*_n\}$, where $A^*_j = [a^*_{l']_j k', k=1}]$, $\vartheta_l = \theta_{j,k}$ and $\vartheta_{l'} = \theta_{j,k}$. Moreover, we have $t_{l'}* = \text{tr}\{(\Lambda_j + \sigma^2 I_n)^{-1}(D_k \Lambda_j)(A_j \sigma^2 I_n)^{-1}(D_k \Lambda_j)\} = w_{j,kj',k'}$, for $\vartheta_l = \theta_{j,k}$ and $\vartheta_{l'} = \theta_{j,k'}$. Therefore, $A^*_j = A_j$ is nonsingular, and (B.3) is established.

For (B.4), the smallest eigenvalue of $\Sigma$, $\mu_n(\Sigma) \geq \sigma^2$, and therefore, $\|\Sigma^{-1}\|_2 \leq \sigma^{-2}$.

Therefore, by Theorem 2 of Mardia and Marshall (1984), we obtain

$$\left(\begin{array}{c}
    H(\beta_0)^{1/2} \\
    E\{\ell''(\beta_0, \delta_0)\}^T \\
    E\{\ell''(\beta_0, \delta_0)\}
\end{array}\right) \left\{\begin{array}{c}
    \hat{\beta} \\
    \hat{\delta}
\end{array}\right\} - \left(\begin{array}{c}
    \beta_0 \\
    \delta_0
\end{array}\right) \overset{D}{\rightarrow} N(0, I_{p+q}),$$

where $E\{\ell''(\beta_0, \delta_0)\}$ is a $p \times q$ matrix with $(i,j)$th element $-\frac{\partial^2 \ell(\lambda, \mu(t), \theta, \sigma^2)}{\partial \beta_i \partial \theta_j}|_{\beta=\beta_0, \delta=\delta_0}$. Note that $\frac{\partial^2 \ell(\lambda, \mu(t), \theta, \sigma^2)}{\partial \beta_i \partial \theta_j} = X^T \Sigma^{-1}(D_j \Sigma) \Sigma^{-1}(Y - X \beta)$ and $E\{\ell''(\beta_0, \delta_0)\} = 0$. Thus, Theorem 4 is proved.
Appendix F: Remarks on Non-Gaussian Processes

Theorems 1 and 3 hold for non-Gaussian stochastic processes under two additional assumptions:

(C.1) \( \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{\infty} \text{cov}\{\xi_j(s_i)^2, \xi_j(s_{i'})^2\} = o(n^2) \)

(C.2) For detrended geostatistical functional data, \( \mathbb{E}\{y(s_i, t') - \mu(s_i, t')\}^4 < \infty \).

The proofs are similar to those of Theorems 1 and 3 for Gaussian processes. However, in (5), (6) is replaced by \( \mathbb{E}\{\xi_j(s_i)^2 \xi_j(s_{i'})^2\} = \text{cov}\{\xi_j(s_i)^2, \xi_j(s_{i'})^2\} + \lambda_j^2 \), and for (I1), we have

\[
\sum_{j=1}^{\infty} \mathbb{E}\|(\hat{\Gamma}_1 - \Gamma)\varphi_j(t)\|^2 \leq n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{\infty} \sum_{j'=1}^{\infty} \mathbb{E}\{\xi_j(s_i)\xi_j(s_{i'})\} \mathbb{E}\{\xi_{j'}(s_i)\xi_{j'}(s_{i'})\} \\
+n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \sum_{j=1}^{\infty} \text{cov}\{\xi_j(s_i)^2, \xi_j(s_{i'})^2\}.
\]

With (C.1) replacing (A.4), (I1) \( \overset{P}{\longrightarrow} 0 \) holds. In Lemma 1, a fourth moment condition (C.2) is required in (13) to show

\[
E(IV_2)^2 = E(IV_3)^2 \leq c_n^{1/2} \left[ E\{y(s_i, t') - \mu(s_i, t')\}^4 \right]^{1/4} \left[ E\{y(s_{i'}, t') - \mu(s_{i'}, t')\}^4 \right]^{1/4} \rightarrow 0,
\]

and thus, \( (IV_i) \overset{P}{\longrightarrow} 0 \), for \( i = 2, 3 \).

Remarks: For the increasing domain, the average distance \( n^{-2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \|s_i - s_{i'}\|_2 \rightarrow \infty \) as \( n \rightarrow \infty \). Since as distance grows, the spatial correlation becomes weaker, and here Condition (C.1) means that the average covariance goes to zero. Moreover, in Theorem 2 and Theorem 4 as well as Step II of estimation procedure, Gaussian assumption cannot be relaxed.

Appendix G: Further Simulation Study

In Section 5.1 of the main manuscript, our proposed method (GFD) is compared with two alternative approaches (ALT1 and ALT2). Here, we investigate the performance of the
proposed method with different numbers of eigenfunctions through simulation studies. For the proposed method with \( J \) eigenfunctions, denoted as \( \text{GFD}_J \), \( J = 2, 3, 4 \). For the simulated datasets in Section 5.1, our methods with \( J = 2, \ldots, 4 \) are carried out, where datasets are simulated from an underlying geostatistical functional data with 2 eigenfunctions, as specified in Section 5.1 of the paper.

<table>
<thead>
<tr>
<th>( n )</th>
<th>Method</th>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \beta_3 )</th>
<th>( \beta_4 )</th>
<th>( \beta_5 )</th>
<th>( \beta_6 )</th>
<th>( \beta_7 )</th>
<th>MSPE (_1)</th>
<th>MSPE (_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>True Values</td>
<td>4.000</td>
<td>3.000</td>
<td>2.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>GFD(_2)</td>
<td>4.001</td>
<td>2.992</td>
<td>1.998</td>
<td>1.007</td>
<td>0.002</td>
<td>0.000</td>
<td>-0.001</td>
<td>2.610</td>
<td>1.106</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.031</td>
<td>0.030</td>
<td>0.029</td>
<td>0.031</td>
<td>0.033</td>
<td>0.027</td>
<td>0.030</td>
<td>0.715</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>GFD(_3)</td>
<td>4.001</td>
<td>2.992</td>
<td>1.997</td>
<td>1.007</td>
<td>0.003</td>
<td>0.001</td>
<td>-0.001</td>
<td>2.611</td>
<td>1.117</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.031</td>
<td>0.030</td>
<td>0.029</td>
<td>0.031</td>
<td>0.033</td>
<td>0.027</td>
<td>0.030</td>
<td>0.714</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>GFD(_4)</td>
<td>4.002</td>
<td>2.992</td>
<td>1.997</td>
<td>1.007</td>
<td>0.004</td>
<td>0.000</td>
<td>-0.001</td>
<td>2.609</td>
<td>1.126</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.030</td>
<td>0.031</td>
<td>0.029</td>
<td>0.032</td>
<td>0.033</td>
<td>0.028</td>
<td>0.030</td>
<td>0.715</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>ALT(_1)</td>
<td>3.996</td>
<td>2.991</td>
<td>2.001</td>
<td>1.011</td>
<td>0.005</td>
<td>-0.005</td>
<td>0.004</td>
<td>3.720</td>
<td>3.775</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.053</td>
<td>0.055</td>
<td>0.054</td>
<td>0.058</td>
<td>0.056</td>
<td>0.053</td>
<td>0.057</td>
<td>1.130</td>
<td>0.453</td>
</tr>
<tr>
<td></td>
<td>ALT(_2)</td>
<td>4.001</td>
<td>2.992</td>
<td>1.998</td>
<td>1.007</td>
<td>0.002</td>
<td>0.000</td>
<td>-0.001</td>
<td>3.715</td>
<td>1.109</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.031</td>
<td>0.030</td>
<td>0.029</td>
<td>0.032</td>
<td>0.033</td>
<td>0.027</td>
<td>0.030</td>
<td>1.133</td>
<td>0.108</td>
</tr>
</tbody>
</table>

Table 1: The mean, standard deviation (SD) of regression coefficient estimates and mean square prediction errors under GFD\(_2\), GFD\(_3\), GFD\(_4\), ALT\(_1\) and ALT\(_2\), for sample size \( n = 100 \).

In Table 1, the regression coefficient estimates of our methods with different eigenfunctions have similar biases and standard deviations. Compared with the two competing methods, our method with different eigenfunctions has performance similar to the traditional functional data analysis, while outperforms ordinary least squares in terms of smaller standard deviations. For prediction, GFD\(_2\) slightly outperforms GFD\(_3\) and GFD\(_4\), since the datasets are simulated from models with 2 eigenfunctions. All three methods (GFD\(_2\), GFD\(_3\) and GFD\(_4\)) outperform the two competing methods (ALT\(_1\) and ALT\(_2\)).

In
short, if we correctly specify the number of eigen-components for the geostatistical functional data, the result is the best. If we choose a larger number of eigen-components, the prediction results are not as good, but are still much better than the alternative methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \lambda_1 )</th>
<th>( r_1 )</th>
<th>( \lambda_2 )</th>
<th>( r_2 )</th>
<th>( \sigma^2 )</th>
<th>( \lambda_3 )</th>
<th>( r_3 )</th>
<th>( \lambda_4 )</th>
<th>( r_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Values</td>
<td>2.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.30</td>
<td>1.00</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GFD_2</td>
<td>2.312</td>
<td>0.469</td>
<td>0.516</td>
<td>0.319</td>
<td>1.017</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SD</td>
<td>0.451</td>
<td>0.136</td>
<td>0.085</td>
<td>0.107</td>
<td>0.049</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GFD_3</td>
<td>2.312</td>
<td>0.468</td>
<td>0.516</td>
<td>0.317</td>
<td>0.990</td>
<td>0.066</td>
<td>1.266</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SD</td>
<td>0.451</td>
<td>0.136</td>
<td>0.085</td>
<td>0.105</td>
<td>0.049</td>
<td>0.007</td>
<td>2.371</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GFD_4</td>
<td>2.312</td>
<td>0.467</td>
<td>0.516</td>
<td>0.316</td>
<td>0.971</td>
<td>0.066</td>
<td>1.307</td>
<td>0.057</td>
<td>2.810</td>
</tr>
<tr>
<td>SD</td>
<td>0.451</td>
<td>0.136</td>
<td>0.085</td>
<td>0.105</td>
<td>0.050</td>
<td>0.007</td>
<td>2.417</td>
<td>0.006</td>
<td>3.131</td>
</tr>
</tbody>
</table>

Table 2: The mean, standard deviation (SD) of spatial-temporal coefficients estimates under GFD_2, GFD_3 and GFD_4 for sample size \( n = 100 \).

In Table 2, the parameter estimates from our method with different eigenfunctions are obtained. In the case of GFD_2 where the number of eigenfunction is correctly specified, the simulation results are close to the true values, as illustrated in Table 2 of the main manuscript. In the cases of GFD_3 and GFD_4 where we specify more eigenfunctions than the true model, the estimates for the first two eigen-components are similar to those of GFD_2, and are close to the true values. For additional eigen-components, the estimates of the additional eigenvalues, \( \lambda_3 \) and \( \lambda_4 \), are quite small compared with the estimates of \( \lambda_1 \) and \( \lambda_2 \).

For sample size \( N = 1000, 2000, 3000 \), the corresponding computational time of proposed methods is 0.1, 0.5 and 1.5 minutes, respectively, while the method without utilizing Sherman-Morrison-Woodbury formula and Sylvester’s determinant theorem needs 1.5, 11 and 45 minutes for the same size. The computational burden can be further relieved by utilizing some form of approximation, such as blocking or tapering, in the covariance matrix inversion.
References


