A COVARIATE-CONDITIONED ORNSTEIN-UHLENBECK BRIDGE MODEL FOR RECONSTRUCTING PROBABILITY FIELD IN SPARSE DOMAIN

Anonymous authors
Paper under double-blind review

ABSTRACT

A natural strategy for inferring the latent state at an unobserved location in space, especially when target observations are sparse, is to use information from neighboring target observations and its auxiliary information. We consider this problem in a setting where target observations are sparse, while auxiliary covariate information is available throughout the space domain. In order to solve the problem, we propose a covariate-conditioned Ornstein-Uhlenbeck (COU) bridge model for reconstructing probability field at unobserved location. The proposed model decomposes the probability field into a deterministic component induced by global covariates and a stochastic residual modeled by an OU process. The local decorrelation structure of the residual is modulated by the covariates, allowing the model to reflect spatial heterogeneity and structural discontinuities. Given observations at two flanking locations, the resulting COU bridge provides probabilistic inference at interior locations. We first perform experiment on spatial data, and then for more generalization we evaluate the model on multiple spatio-temporal datasets. The results show that the proposed model is effective when sparse observations are complemented by informative covariates and the underlying dependence structure is consistent with the bridge formulation.

1 INTRODUCTION

Sparse prediction reflecting spatial dependency becomes challenging when observations of the target quantity are available only at a few locations. When the surrounding domain contains heterogeneous structural information, the relevance of a neighboring observation is determined by not only its geometric distance from the target but also the spatial context between them. This suggests that spatial prediction should account for few observable covariates modify both the expected field and the strength of spatial dependence. Motivated by this setting, we propose a covariate-conditioned bridge framework that combines globally available spatial information with sparse local observations and adapts the spatial dependence structure according to observed covariates.

Interpolation from sparse observations by incorporating spatial information has been extensively studied (Paciorek & Schervish, 2006; Blasi & Furrer, 2025; Tsao et al., 2026; Cao et al., 2026). Classical approaches include kriging with nonstationary covariance structures that account for spatially varying dependence (Paciorek & Schervish, 2006), as well as extensions in which the covariance structure is conditioned on observed covariates (Blasi & Furrer, 2025). More recently, diffusion-based generative models have achieved strong performance in modeling complex latent distributions (Ho et al., 2020; Song et al., 2021), and diffusion-bridge formulations provide a direct way to construct conditional generative processes between endpoint distributions (Liu et al., 2023; Zhou et al., 2024; Yue et al., 2024; Albergo et al., 2025). Motivated by these developments, diffusion-based methods have also been explored for spatial interpolation and reconstruction (Tsao et al., 2026; Cao et al., 2026). However, in some applications, a reliable governing equation or forward model may not be available. This limits the use of diffusion-based approaches that rely on dense reference fields or simulation-generated training targets. Under the these conditions, we seek to infer unobserved targets directly from sparse measurements while using globally available covariates to characterize both the baseline field and the way spatial information propagates across the domain.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

The Ornstein-Uhlenbeck (OU) bridge provides a natural stochastic construction for this purpose. Its Markov structure enables inference at an interior location by conditioning on neighboring endpoint states, while its Gaussian representation yields tractable transition and bridge distributions. More importantly, by allowing the OU decorrelation rate and residual amplitude to vary with spatial covariates, the model can adapt the effective correlation scale along the domain. Regions associated with structural changes can therefore reduce the influence of distant observations, whereas strongly correlated regions allow information to propagate over longer distances. This covariate-dependent dependence structure distinguishes the covariate-conditioned OU (COU) bridge from a stationary bridge in which spatial dependence is determined by a single fixed correlation scale.

Since real-world datasets that match the problem setting and isolate covariate-dependent spatial dependence are difficult to identify, we first use controlled synthetic experiments in which the spatial dependence structure is explicitly considered and its probability field is randomly generated. These experiments allow us to examine whether the COU bridge can recover covariate-induced variation in the dependence structure under different observation and correlation regimes. We further evaluate the proposed model using network-structured spatio-temporal datasets which are commonly used traffic benchmarks (Jagadish et al., 2014; Li et al., 2018). The datasets are typically studied for spatio-temporal forecasting, where graph structures are used to represent spatial dependencies between sensors and their temporal evolution is predicted (Wu et al., 2021; Mei et al., 2023). Our objective is different from conventional time-series forecasting. The proposed model can be extended to spatio-temporal forecasting. Although the target measurements vary over time, temporal information is used only as input covariates, while the prediction at each time point is performed from neighboring endpoint observations and spatial information. Thus, each temporal snapshot is treated as a spatial reconstruction problem rather than as a forecasting task. Through these experiments, we investigate the conditions under which the COU bridge provides an effective mechanism for propagating information between sparsely observed locations.

2 Reconstruction of probability field in sparse domain

2.1 Related works

There are many studies to reconstruct a complete spatial fields from sparse observations, which is a representative inverse problem that appears in various fields such as geoscience, meteorology, fluid mechanics, and scientific computing. Zhuang et al. (2025) propose a score-based diffusion model to estimate spatio-temporal fields from sparse sensor data for various PDE. Zhai et al. (2025) suggest a hybrid transformer-based machine learning framework to reconstruct complex and nonlinear dynamics of chaotic systems under sparse observations and without training data. Luo et al. (2024) introduce an implicit neural representation (INR)-based reconstruction method to construct a continuous representation from sparsely sampled irregular surface temperature data.

2.2 Problem formulation

We consider a class of sparse spatial inference problems in which the quantity of interest is observed only at a limited number of locations, while auxiliary covariate information is available throughout the domain. Our formulation allows the strength of dependence between observations to vary according to the covariate structure along the domain. In particular, globally available covariates may contain continuous variation as well as localized structural changes that affect how information should propagate between sparsely observed locations. We formalize this setting as a probability field reconstruction problem and use it as the basis for the proposed COU bridge.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an underlying probability space and let $\mathcal{S}$ denote a spatial domain. We define a response field as

$$Y : \mathcal{S} \times \Omega \to \mathcal{D},$$

such that, for each fixed location $x \in \mathcal{S}$, $Y(x) : \Omega \to \mathcal{D}$ is a random variable. Let $\gamma : [0, 1] \to \mathcal{S}$ be a path connecting two observed locations A and B, such that $A = \gamma(0)$, $B = \gamma(1)$. For an interior point $u \in (0, 1)$, we denote $M = \gamma(u)$ and define $Y_A = Y(A), Y_B = Y(B), Y_M = Y(M)$.

We here propose a latent field along the path as

$$Z(u) = m(u) + R(u),$$

where $m(u)$ represents the systematic component determined by the observed covariates $x(u)$, which are available over the entire path and determine the baseline probability field, and $R(u)$ represents a spatial residual modeled as a stochastic process. Also, along the path, the underlying probability field is defined through

$$p(u) = \sigma(Z(u)) = \frac{1}{1 + \exp(-Z(u))}.$$

Consider D_A and D_B as the observations available at the two endpoints. We do not initially restrict these observations to a particular sampling distribution. Instead, their relationship with the endpoint states $R(0) = R_A, R(1) = R_B$ is represented through likelihood functions

$$L_A(r) = p(D_A \mid R_A = r), \quad L_B(r) = p(D_B \mid R_B = r).$$

The specific form of these likelihoods depends on the observation mechanism. For instance, the form can be determined by Bernoulli or Binomial observations. Let $q_u(r \mid D_A, D_B)$ denote the conditional density of $R(u)$ given the endpoint observations. Since $R(u)$ is latent, the predictive probability at the interior location M is obtained by marginalizing over its posterior distribution:

$$\begin{aligned} p_M &:= \mathbb{E}[p(u) \mid D_A, D_B] \\ &= \int_{\mathbb{R}} \sigma(m(u) + r) q_u(r \mid D_A, D_B)\, dr \\ &= \mathbb{E}\left[\sigma(m(u) + R(u)) \mid D_A, D_B\right]. \end{aligned} \tag{1}$$

Thus the problem is equivalent to characterizing the conditional distribution of the residual process $R(u)$ given the information observed at the two endpoints. In the following subsection, we specify the covariate-induced mean component $m(u)$ and then introduce the residual process $R(u)$ represented by a nonstationary Ornstein-Uhlenbeck process.

3 Derivation of covariate-conditioned Ornstein-Uhlenbeck bridge

The formulation equation 1 suggests separating two sources of information in the latent field. The globally available covariates determine a systematic baseline structure, whereas the remaining uncertainty and spatial dependence must be inferred from the sparse endpoint observations. We therefore decompose the latent field into a deterministic covariate-induced component and a stochastic residual component, and model the latter using a nonstationary OU process whose local dependence structure is conditioned on the covariates. For a detailed derivation of the bridge marginal distribution, we refer the reader to Øksendal (2003); Chen & Georgiou (2015) and the references therein.

3.1 Covariate-induced mean structure

Denote a vector of covariates at the location $\gamma(u)$ as $x(u) = (x_1(u), \ldots, x_n(u))^\top$. We define the covariate-induced mean function as

$$m(u) = f(x(u); \beta)$$

where f is a deterministic function parameterized by β. Applying the inverse logistic transformation gives

$$\operatorname{logit} p(u) = m(u) + R(u).$$

Hence, the residual process can equivalently be written as

$$R(u) = \operatorname{logit} p(u) - m(u),$$

which provides a direct interpretation of the residual field, and $R(u)$ measures the deviation of the latent log-odds from the baseline structure explained by the covariates. We assume that the mean function is sufficiently specified so that the residual process is centered,

$$\mathbb{E}[R(u)] = 0,$$

and model the dependence structure of $R(u)$ separately. In particular, the following subsection models $R(u)$ as a covariate-conditioned nonstationary Ornstein–Uhlenbeck process.

3.2 ORNSTEIN-UHLENBECK RESIDUAL MODEL

Suppose $R(u)$ satisfies the inhomogeneous stochastic differential equation

$$dR(u) = -\kappa(u)R(u)du + \eta(u)dW(u), \tag{2}$$

where $\kappa(u)$ and $\eta(u)$ are sufficiently regular drift and diffusion coefficient functions, respectively, and $W(u)$ is a standard Wiener process. We assume $R(0) \sim \mathcal{N}(0, V(0))$. If we put $V(u) = \mathrm{Var}(R(u))$ and

$$\nu(u) = \sqrt{V(u)}, \quad \theta(u) = \frac{\eta(u)^2}{2V(u)}, \quad \tau(u) = \int_0^u \theta(s)ds$$

then the process in equation 2 admits the equivalent representation

$$R(u) = \nu(u)\widetilde{W}(\tau(u))$$

where $\widetilde{W}$ is a stationary Ornstein-Uhlenbeck process with unit mean-reversion rate and unit marginal variance. The functions θ and ν correspond to the positive decorrelation rate and residual amplitude respectively, and the covariance is given by

$$\mathrm{Cov}(R(u), R(v)) = \nu(u)\nu(v)e^{-|\tau(u)-\tau(v)|}, \quad u, v \in [0, 1]. \tag{3}$$

The representation of equation 3 separates two distinct forms of spatial variation. The decorrelation rate $\theta(u)$ determines how rapidly information is lost along the path, whereas $\nu(u)$ determines the marginal magnitude of the residual variation.

For $0 \leq u < v \leq 1$, define

$$\Delta\tau_{uv} = \tau(v) - \tau(u) = \int_u^v \theta(s)ds.$$

Since $R(u)$ is a Gaussian Markov process, its transition distribution is

$$R(v) \mid R(u) = r \sim \mathcal{N}\left(\frac{\nu(v)}{\nu(u)}e^{-\Delta\tau_{uv}}r,\ \nu^2(v)\left(1 - e^{-2\Delta\tau_{uv}}\right)\right).$$

Thus, the integrated quantity $\Delta\tau_{uv}$, rather than the Euclidean distance between the two locations alone, determines the amount of decorrelation between them.

We parameterize the two functions by

$$\log\theta(u) = \alpha_0 + \boldsymbol{\alpha}^\top x_\theta(u), \quad \log\nu^2(u) = \gamma_0 + \boldsymbol{\gamma}^\top x_\nu(u)$$

where $x_\theta(u)$ and $x_\nu(u)$ denote covariates associated with the local dependence rate and the residual amplitude, respectively. This parameterization guarantees $\theta(u) > 0$ and $\nu(u) > 0$, and therefore ensures that $\tau(u)$ is monotonically increasing and that the covariance function is positive semidefinite.

3.3 BRIDGE DISTRIBUTION

Let $R(u) = R_M$ and $\nu(u) = \nu_M$. The covariance matrix of the two endpoint states is denoted by K_{AB} and k_M indicates the covariance vector between R_M and the endpoint states. Since (R_A, R_M, R_B) is jointly Gaussian, conditioning on the latent endpoint states induces

$$R_M \mid R_A = r_A, R_B = r_B \sim \mathcal{N}(\mu_M(r_A, r_B), v_M) \tag{4}$$

where

$$\mu_M(r_A, r_B) = k_M^\top K_{AB}^{-1}\boldsymbol{r}_{AB}, \quad v_M = \nu_M^2 - k_M^\top K_{AB}^{-1}k_M, \quad \boldsymbol{r}_{AB} = (r_A, r_B)^\top.$$

Using the likelihood functions introduced in section 2.2 their posterior density is

$$q_{AB}(r_A, r_B \mid D_A, D_B) \propto L_A(r_A)L_B(r_B)\varphi\left(\boldsymbol{r}_{AB}; 0, K_{AB}\right),$$

where $\varphi(\,\cdot\,; 0, K_{AB})$ denotes the bivariate Gaussian density with zero mean and covariance K_{AB}. Therefore, the conditional density of the residual at the target location is

$$q_u(r_M \mid D_A, D_B) = \iint p(r_M \mid r_A, r_B)q_{AB}(r_A, r_B \mid D_A, D_B)dr_A dr_B.$$

Substituting this expression into the predictive probability in equation 1 yields

$$p_M = \mathbb{E}_{R_A,R_B|D_A,D_B}\left[\mathbb{E}_{R_M|R_A,R_B}\left[\sigma(m(u)+R_M)\right]\right].$$

For a general non-Gaussian observation likelihood, the posterior distribution of the endpoint states is not available in closed form. Therefore we apply a Laplace approximation. Define

$$\hat{\boldsymbol{r}}_{AB} \equiv (\hat{r}_A, \hat{r}_B)^\top = \arg\max_{\boldsymbol{r}_{AB}} \left\{ \log L_A(r_A) + \log L_B(r_B) - \frac{1}{2}\boldsymbol{r}_{AB}^\top K_{AB}^{-1} \boldsymbol{r}_{AB} \right\}$$

and

$$W = \operatorname{diag}\left(-\frac{\partial^2}{\partial r_A^2} \log L_A(\hat{r}_A),\ -\frac{\partial^2}{\partial r_B^2} \log L_B(\hat{r}_B) \right).$$

The posterior distribution of the endpoint states is then approximated by

$$q_{A,B}(\boldsymbol{r}_{AB} \mid D_A, D_B) \approx \mathcal{N}\left(\boldsymbol{r}_{AB}; \hat{\boldsymbol{r}}_{AB}, \Sigma_{AB}\right), \tag{5}$$

where $\Sigma_{AB} = \left(K_{AB}^{-1} + W\right)^{-1}$. If we put $\mathbf{c} = K_{AB}^{-1} k_M$, the Gaussian bridge distribution derived in 4 is

$$R_M \mid R_A = r_A, R_B = r_B \sim \mathcal{N}(\mathbf{c}^\top \boldsymbol{r}_{AB}, v_M).$$

By the Laplace approximation in 5, we have

$$q_u(r_M \mid D_A, D_B) \approx \mathcal{N}(r_M; \mu_q, v_q), \quad \mu_q = \mathbf{c}^\top \hat{\boldsymbol{r}}_{AB}, \quad v_q = v_M + \mathbf{c}^\top \Sigma_{AB} \mathbf{c}.$$

Then the predictive probability given as in equation 1 is evaluated by using Gauss-Hermite quadrature

$$p_M \approx \frac{1}{\sqrt{\pi}} \sum_{n=1}^{K} w_n \sigma(m(u) + \mu_q + \sqrt{2v_q} t_n)$$

where t_n and w_n denote the Gauss-Hermite nodes and weights, respectively.

4 Experimental results

4.1 Experiments on synthetic dataset

We first conducted experiments to evaluate whether the proposed COU bridge can recover spatial dependence that varies systematically with observed covariates on synthetic dataset. In particular, the experiments were designed to distinguish covariate effects on the mean structure from those on the local dependence and residual variability. Following the problem formulation in Section 2.2, we generate a probability field over a one-dimensional spatial domain $s \in [0, L]$ as

$$z(s) = m(s) + R(s), \quad p(s) = \sigma(z(s))$$

where $m(s) = \beta^\top x(s)$ denotes the covariate-induced mean and $R(s)$ is the OU residual process. The baseline hyperparameters used to generate the probability field are summarized in Table 4 of Appendix A.2. The residual trajectories were generated using the exact OU transition in the transformed coodrinate τ. Sparse observations were then generated at selected locations s_i from $K_i \sim \text{Binomial}(n_i, p(s_i))$.

Table 1: Global covariates used in the synthetic experiments. The covariates include two smoothly varying features, a piecewise-constant regime feature, and a narrow structural-band indicator. Here, L_c and L_e denote the mean regime length and mean spacing between structural bands, respectively.

Covariate	Definition	Parameters
$x_1(s)$	Gaussian white noise smoothed with kernel s.d. ℓ_u	$\ell_u = 12$
$x_2(s)$	$\sin(2\pi s/P) + 0.5\, s/L$	$P = 60$
$x_3(s)$	i.i.d. $N(0,1)$ levels on segments with segments with mean length L_c	$L_c = 25$
$x_4(s)$	$\mathbf{1}_{\{\mathrm{s}-\mathrm{s_e}\leq \mathrm{w}/2\}}$, where the interface locations s_e have mean spacing L_e	$L_e = 8$, $w = 1.25$

The covariate vector $x(s) = (x_1(s), x_2(s), x_3(s), x_4(s))^\top$ contains both smooth varying and piecewise-constant spatial features, as summarized in Table 1. The first two covariates represent

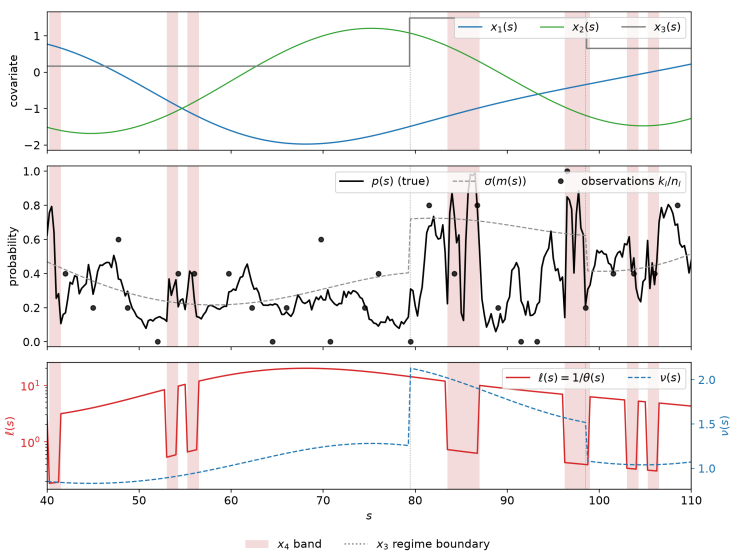

Figure 1: **Synthetic data generating setup.** Top: spatial covariates used to construct the mean and nonstationary dependence structure; shaded regions denote the structural-band covariate, and dotted vertical lines indicate regime boundaries. Middle: true probability $p(s)$, the covariate-induced baseline probability $\sigma(m(s))$, and observed proportions k_i/n_i. Bottom: local correlation length $\ell(s)$ and residual amplitude $\nu(s)$.

smoothly varying spatial effects. The regime variable $x_3(s)$ changes piecewise across the domain and, under the baseline parameterization, affects both the mean structure and residual amplitude. The band indicator $x_4(s)$ introduces narrow regions of increased decorrelation rate through $\theta(s)$.

Figure 1 illustrates one realization of the synthetic data-generating process. The upper panel shows the three spatial covariates $x_1(s), x_2(s)$ and $x_3(s)$, while the shaded regions indicate the structural-band covariate $x_4(s)$. The middle panel presents the true probability $p(s)$, the probability induced by the covariate mean alone, and the observed empirical proportions k_i/n_i. The lower panel shows the resulting local correlation length $\ell(s) = 1/\theta(s)$ and residual amplitude $\nu(s)$. In particular, the structural bands produce sharp local reduction in $\ell(s)$, whereas changes in the regime feature modify the residual amplitude through $\nu(s)$.

We compared the proposed COU bridge, in which both $\theta(s)$ and $\nu(s)$ vary with covariates, against a covariate mean-only model, linear interpolation, and a stationary Gaussian process (GP) (Bolin et al., 2024). The stationary GP and stationary COU bridge models share the same exponential covariance structure but differ in their treatment of the observations. The GP baseline is fitted to empirical logits using a Gaussian observation approximation, whereas the stationary COU bridge retains the binomial observation likelihood, with a Laplace approximation used for latent-Gaussian posterior computations. The remaining gap to the COU bridge with true parameters indicates that part of the reconstruction error is attribute to parameter estimation.

Prediction was evaluated at masked interior locations. For probability field reconstruction, the posterior predictive probability was compared with the true probability $p(s)$ using RMSE and CRPS. Predictive performance for held-out observations was additionally evaluated using negative log-likelihood (NLL), Brier score, and AUROC. Table 2 shows the results under the baseline synthetic setting.

Under this baseline setting, the stationary residual models improved upon the covariate-mean and linear-interpolation baselines, while the COU bridge model further reduced RMSE and CRPS relative to the stationary OU model. The remaining gap to the oracle result indicates the extent to which performance may still be limited by parameter estimation.

Table 2: Predictive performance under the baseline synthetic setting. Lower values indicate better performance for RMSE, CRPS, NLL, and Brier score, whereas higher values indicate better performance for AUROC.

Model	RMSE ↓	CRPS ↓	NLL ↓	Brier ↓	AUROC ↑
Covariate mean	0.200	0.119	0.605	0.209	0.732
Linear interpolation	0.176	0.097	0.587	0.200	0.755
Stationary GP (Kriging)	0.175	0.106	0.583	0.199	0.759
COU bridge (stationary)	0.160	0.088	0.572	0.194	0.769
COU bridge (nonstationary)	0.154	0.084	0.568	0.193	0.772
COU bridge (true parameter)	0.143	0.075	0.570	0.189	0.781

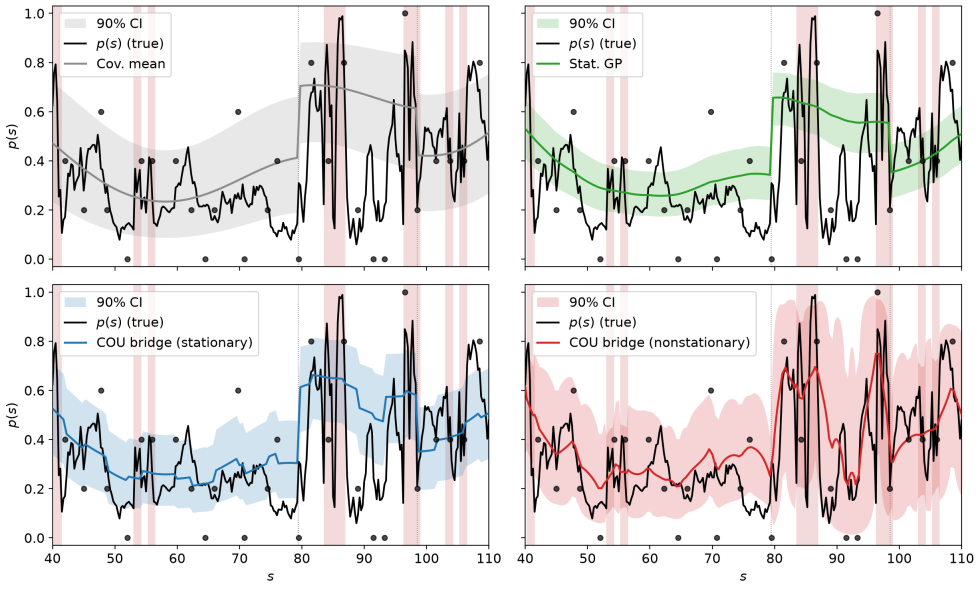

Figure 2: **Reconstruction of the synthetic probability field.** Posterior mean probability fields and 90% posterior intervals are shown for the covariate-mean model (top left), stationary GP (top right), stationary COU bridge (bottom left), and nonstationary COU bridge (bottom right). The black curve denotes the true probability field, and the points are the observed proportions k_i/n_i. Shaded vertical regions indicate structural bands, and dotted lines indicate regime boundaries.

How the four models reconstruct the same probability field is illustrated in Figure 2. Compared with the stationary models, the COU bridge allows both the posterior mean and uncertainty to change with the local dependence structure, particularly near regions with shortened correlation length.

To examine how the spatial correlation scale affects the performance of the COU bridge, we varied the mean observation spacing relative to the baseline correlation length ℓ_0. Figure 3 shows that the COU bridge achieves substantially lower RMSE than the covariate mean model when the observation spacing is smaller then ℓ_0. As the spacing exceeds the correlation length, this advantage gradually decreases and the COU bridge approaches the covariate mean model. This behavior is expected because neighboring observations contain less information about an interior location once their separation becomes large relative to the underlying correlation scale.

4.2 Experimental results for spatio-temporal datasets

We evaluate the proposed COU bridge model on spatio-temporal traffic datasets: METR-LA (Jagadish et al., 2014), and PEMS-BAY (Li et al., 2018), for which repeated measurements are available over a fixed spatial network. We remark that the experiments are formulated as spatial reconstruc-

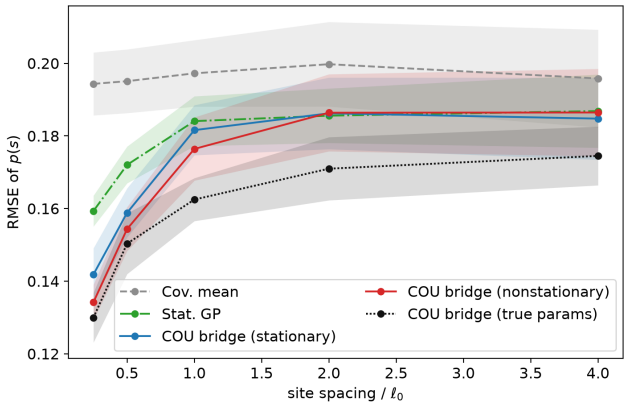

Figure 3: **Effect of observation spacing relative to the correlation length.** RMSE of the reconstructed probability field is shown as a function of the mean observation spacing normalized by the baseline correlation length ℓ_0. Shaded regions denote 90% CI across simulation replicates.

tion rather than temporal forecasting, and true probability field for spatial domain of each dataset is unknown.

We treat each timestamp t as a separate spatial reconstruction problem, while temporal and location-specific covariates are incorporated into the mean function. Let $s_1, \ldots, s_m$ be the ordered sensor locations along a corridor, and let $Y_t(s_i)$ be the observation at location s_i and time t. We model $Y_t(s_i) = m_t(s_i) + R_t(s_i)$, where $m_t(s_i)$ is the component explained by the observed covariates and $R_t(s_i)$ represents the remaining spatially correlated variation. For a masked sensor s_q, conditioning on the neighboring observed sensors gives

$$Y_t(s_q) \mid D_t \sim \mathcal{N}(\mu_{tq}, v_{tq}), \qquad \mu_{tq} = m_t(s_q) + \mathbb{E}[R_t(s_q) \mid D_t],$$

and μ_{tq} is used as the prediction at location s_q and time t. The covariate vectors were constructed from network derived features including normalized corridor position, branch density, and interchange indicates while the adjacency weights were transformed into relative edge distance to define the spatial dependence.

Table 3: Held-out sensor reconstruction on METR-LA and PEMS-BAY datasets. Lower values indicate better performance for all metrics. Best performance within each dataset is shown in bold.

Dataset	Metric	Covariate mean	Stat. GP	COU (stat.)	COU (nonstat.)
METR-LA	RMSE	11.966	11.623	10.917	**10.265**
	CRPS	7.072	6.862	**6.424**	6.634
	NLL	4.086	3.969	**3.900**	4.420
PEMS-BAY	RMSE	8.306	8.380	8.358	**7.028**
	CRPS	4.242	4.342	4.334	**3.706**
	NLL	3.628	3.615	3.612	**3.417**

As shown in Table 3, all methods yield higher RMSE and CRPS values on METR-LA than on PEMS-BAY. This may be attributed to the greater temporal variability of METR-LA, which has been associated with more heterogeneous traffic dynamics and stronger spatial dependencies in previous studies, including Li et al. (2018). For METR-LA, the nonstationary COU bridge achieves the lowest RMSE, suggesting that allowing spatial dependence to vary along the network is beneficial for conditional mean reconstruction under highly variable traffic conditions. However, the stationary COU bridge achieves lower CRPS and NLL values. These results suggest that, while the COU bridge structure is effective in capturing the overall spatial trend, the current nonstationary formulation may not yet fully account for the pronounced temporal fluctuations present in METR-LA.

In contrast, on PEMS-BAY, the nonstationary COU bridge consistently achieves lower RMSE, CRPS, and NLL values than both the covariate mean model and the stationary GP. This result indi-

cates that incorporating residual spatial dependence through the COU bridge improves not only the reconstruction of the predictive mean but also the characterization of predictive uncertainty.

5 Conclusion

In this study, we formulated a spatial inference problem in which globally available spatial covariates are combined with a limited number of sparse observations of the quantity of interest to predict targets at unobserved locations. To address this problem, we proposed a COU bridge model for probabilistic inference at interior locations.

The experimental results showed that the proposed COU bridge can reflect spatial variations well in the dependence structure, including regime-dependent changes in local decorrelation. When the observations are sufficiently dense relative to the underlying correlation length, the proposed model achieved reconstruction performance close to that of the reference model evaluated using the true data-generating parameters. The proposed framework can therefore be useful for reconstructing states in unobserved regions of sensor networks and, more generally, for estimating interior states in spatial domains where dense measurements are difficult to obtain, such as geological applications.

The benefit of the COU bridge, however, relatively decreases when observations are too sparse or when the available spatial covariates contain limited information about variations in the dependence structure. In such cases, separating the field into deterministic and stochastic components provides less advantage over simpler models. In addition, the present study used a relatively simple linear model for the covariate-induced mean function, and the spatio-temporal experiments were not designed to fully model temporal dependence. Future work will investigate more flexible mean functions based on learned feature representations and extend the framework to jointly model spatial and temporal residual dependence.

AI use statement

During the preparation of this work, authors used generative AI tools in order to generate synthetic data, assist with experimental code, implementation, and polish the language of the manuscript. We carefully reviewed and edited the content as needed and take full responsibility for the final content of this work.

References

Michael Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.

Federico Blasi and Reinhard Furrer. A class of modular and flexible covariate-based covariance functions for nonstationary spatial modeling. *Environmetrics*, 36(8):e70038, 2025.

David Bolin, Alexandre B Simas, and Jonas Wallin. Gaussian whittle–matérn fields on metric graphs. *Bernoulli*, 30(2):1611–1639, 2024.

Yuanhao Cao, Haiqi Wang, Shang Wu, Zhenyu Zhang, and Feng Yan. A spatial interpolation method based on denoising diffusion probabilistic model. *International Journal of Geographical Information Science*, 40(1):171–193, 2026.

Yongxin Chen and Tryphon Georgiou. Stochastic bridges of linear systems. *IEEE Transactions on Automatic Control*, 61(2):526–531, 2015.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Hosagrahar V Jagadish, Johannes Gehrke, Alexandros Labrinidis, Yannis Papakonstantinou, Jignesh M Patel, Raghu Ramakrishnan, and Cyrus Shahabi. Big data and its technical challenges. *Communications of the ACM*, 57(7):86–94, 2014.

Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=SJiHXGWAZ.

Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos A Theodorou, Weili Nie, and Anima Anandkumar. I^2sb: Image-to-image schrödinger bridge. *arXiv preprint arXiv:2302.05872*, 2023.

Xihaier Luo, Wei Xu, Balasubramanya T Nadiga, Yihui Ren, and Shinjae Yoo. Continuous field reconstruction from sparse observations with implicit neural networks. In *International Conference on Learning Representations*, volume 2024, pp. 3281–3305, 2024.

Hao Mei, Junxian Li, Zhiming Liang, Guanjie Zheng, Bin Shi, and Hua Wei. Uncertainty-aware traffic prediction under missing data. In *2023 IEEE International Conference on Data Mining (ICDM)*, pp. 1223–1228. IEEE, 2023.

Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations: an introduction with applications*. Springer, 2003.

Christopher J Paciorek and Mark J Schervish. Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics: The official journal of the International Environmetrics Society*, 17(5):483–506, 2006.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

Valerie Tsao, Nathaniel W Chaney, and Manolis Veveakis. Probabilistic spatial interpolation of sparse data using diffusion models. *Artificial Intelligence for the Earth Systems*, 5(1):250049, 2026.

Yuankai Wu, Dingyi Zhuang, Aurelie Labbe, and Lijun Sun. Inductive graph neural networks for spatiotemporal kriging. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 4478–4485, 2021.

Conghan Yue, Zhengwei Peng, Junlong Ma, Shiyan Du, Pengxu Wei, and Dongyu Zhang. Image restoration through generalized ornstein-uhlenbeck bridge. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=oDUJmNCV8D.

Zheng-Meng Zhai, Benjamin D Stern, and Ying-Cheng Lai. Bridging known and unknown dynamics by transformer-based machine-learning inference from sparse observations. *Nature communications*, 16(1):8053, 2025.

Linqi Zhou, Aaron Lou, Samar Khanna, and Stefano Ermon. Denoising diffusion bridge models. In *International Conference on Learning Representations*, volume 2024, pp. 8160–8171, 2024.

Yilin Zhuang, Sibo Cheng, and Karthik Duraisamy. Spatially-aware diffusion models with cross-attention for global field reconstruction with sparse observations. *Computer Methods in Applied Mechanics and Engineering*, 435:117623, 2025.

A Appendix

A.1 Methodologies

Algorithm 1 Covariate-conditioned OU bridge prediction at an interior location $u \in (0,1)$ of the path γ connecting $A = \gamma(0)$ and $B = \gamma(1)$.

Require: covariates $c(s)$ along the path, with dependence covariates $x_\theta(s)$ and $x_\nu(s)$; mean parameters β; dependence parameters $(\alpha_0, \boldsymbol{\alpha}, \gamma_0, \boldsymbol{\gamma})$; endpoint observations D_A, D_B with likelihoods $L_A(r) = p(D_A \mid R_A = r)$, $L_B(r) = p(D_B \mid R_B = r)$; number of Gauss–Hermite nodes K

Ensure: predictive probability p_M and approximate posterior $q_u(r_M \mid D_A, D_B)$

$m(u) \leftarrow f(c(u); \beta)$ ▷ covariate-induced mean, Sec. 3.2

$\theta(s) \leftarrow \exp\{\alpha_0 + \boldsymbol{\alpha}^\top x_\theta(s)\}, \quad \nu(s) \leftarrow \exp\{\tfrac{1}{2}(\gamma_0 + \boldsymbol{\gamma}^\top x_\nu(s))\}, \quad s \in [0,1]; \quad \nu_M \leftarrow \nu(u)$

$\Delta\tau_{0u} \leftarrow \displaystyle\int_0^u \theta(s)\, ds, \quad \Delta\tau_{u1} \leftarrow \int_u^1 \theta(s)\, ds$ ▷ numerical quadrature on the covariate grid

$K_{AB} \leftarrow \begin{pmatrix} \nu(0)^2 & \nu(0)\nu(1)\, e^{-(\Delta\tau_{0u}+\Delta\tau_{u1})} \\ \nu(0)\nu(1)\, e^{-(\Delta\tau_{0u}+\Delta\tau_{u1})} & \nu(1)^2 \end{pmatrix}$ ▷ $\mathrm{Cov}(R(u), R(v)) = \nu(u)\nu(v)e^{-|\tau(u)-\tau(v)|}$

$k_M \leftarrow \big(\nu(0)\,\nu_M\, e^{-\Delta\tau_{0u}},\ \nu_M\,\nu(1)\, e^{-\Delta\tau_{u1}}\big)^\top$

$\hat{\boldsymbol{r}}_{AB} \leftarrow \arg\max_{\boldsymbol{r}_{AB}} \Big\{ \log L_A(r_A) + \log L_B(r_B) - \tfrac{1}{2}\, \boldsymbol{r}_{AB}^\top K_{AB}^{-1} \boldsymbol{r}_{AB} \Big\}$ ▷ two-dim. concave problem; Newton iterations

$W \leftarrow \mathrm{diag}\Big(-\frac{\partial^2}{\partial r_A^2} \log L_A(\hat{r}_A),\ -\frac{\partial^2}{\partial r_B^2} \log L_B(\hat{r}_B) \Big), \quad \Sigma_{AB} \leftarrow \big(K_{AB}^{-1} + W\big)^{-1}$ ▷ Laplace posterior, Eq. (4)

$\boldsymbol{c} \leftarrow K_{AB}^{-1} k_M, \quad v_M \leftarrow \nu_M^2 - k_M^\top K_{AB}^{-1} k_M$ ▷ bridge conditional $R_M \mid \boldsymbol{r}_{AB} \sim \mathcal{N}(\boldsymbol{c}^\top \boldsymbol{r}_{AB},\, v_M)$, Eq. (3)

$\mu_q \leftarrow \boldsymbol{c}^\top \hat{\boldsymbol{r}}_{AB}, \quad v_q \leftarrow v_M + \boldsymbol{c}^\top \Sigma_{AB}\, \boldsymbol{c}$ ▷ $q_u(r_M \mid D_A, D_B) \approx \mathcal{N}(\mu_q, v_q)$

$p_M \leftarrow \dfrac{1}{\sqrt{\pi}} \displaystyle\sum_{n=1}^{K} w_n\, \sigma\big(m(u) + \mu_q + \sqrt{2 v_q}\, t_n\big)$ ▷ Gauss–Hermite quadrature for Eq. (1)

return p_M and (μ_q, v_q)

A.2 EXPERIMENTAL RESULTS

Table 4: Baseline parameter settings for the synthetic data-generating process. The coefficient vector β was rescaled to obtain an equal contribution of the mean and residual components to the total latent-field variance. The quantity $B = \theta_{\text{band}} w$ represents the equivalent decorrelation strength across a structural band.

Specification	**Setting**
Spatial domain and grid	$L = 200, \quad \Delta s = 0.25$
Mean function	$m(s) = \beta^\top x(s)$, with β proportional to $(0.8,\ 0.6,\ 1.0,\ 0)^\top$ and rescaled such that $\dfrac{\mathrm{Var}\{m(s)\}}{\mathrm{Var}\{m(s)\} + \mathrm{Var}\{R(s)\}} = 0.5$
Decorrelation rate	$\log \theta(s) = \alpha_0 + \alpha^\top x(s)$, with $\alpha_0 = -\log \ell_0 = -1.609, \qquad \ell_0 = 5,$ $\alpha = (0.7,\ 0,\ 0,\ 2.773)^\top$
Residual amplitude	$\log \nu^2(s) = \gamma_0 + \gamma^\top x(s)$, with $\gamma_0 = 0, \qquad \nu_0 = 1, \qquad \gamma = (0,\ 0.3,\ 0.8,\ 0)^\top$
Structural-band effect	$\theta_0 = \ell_0^{-1} = 0.2, \qquad \theta_{\text{band}} = 3.2, \qquad w = 1.25,$ $B = \theta_{\text{band}} w = 4, \qquad e^{-B} = 0.018$