Robust spatial functional linear regression

Ufuk Beyaztas **Han Lin Shang** **Abhijit Mandal**

Marmara University Macquarie University University of Texas at El Paso

August 16, 2023

Table of contents

INTRODUCTION

SCALAR-ON-FUNCTION REGRESSION

- Let $\mathcal{H} \rightarrow$ a separable $\mathcal{L}_2(\mathcal{I})$ Hilbert space,
- $\mathcal{I}, f : \mathcal{I} \rightarrow \mathbb{R}$ satisfying $\int_{\mathcal{I}} f^2(t)dt < \infty$.
- Consider a random sample $\{Y_i, \mathcal{X}_i(t) : i = 1, \ldots, n\}$ from $\{Y \in \mathbb{R}, \mathcal{X}(t) \in \mathcal{L}_2(\mathcal{I})$

Then, the scalar-on-function linear regression model is of the form:

$$Y = \alpha + \int_{\mathcal{I}} \mathcal{X}(t)\beta(t)dt + \epsilon,$$

where $\alpha \in \mathbb{R}$, $\beta(t) \in \mathcal{L}_2(\mathcal{I})$, and $\epsilon \sim N(0, \sigma^2)$. For reviews, see Morris, 2015 and Reiss et al., 2017.

Introduction

Variations of the scalar-on-function regression

Linear functional predictor regression:

- Crainiceanu et al., 2009; Fang et al., 2009; Goldsmith et al., 2012; James, 2002; Marx and Eilers, 1999

Nonlinear functional predictor regression:

- Ait-Saidi et al., 2009; James, 2005; Muller et al., 2013; Yao and Muller, 2010

Nonparametric functional regression:

- Ferraty and Vieu, 2006

INTRODUCTION

METHODS TO ANALYZE DATA WITH SPATIAL DEPENDENCE

Discrete data matrix:

- Anselin, 1998; Cressie and Wikle, 2015; Lesage and Pace, 2009; Schabenberger and Gotway, 2017; Wang et al., 2019

Spatially correlated functional data (kriging):

- Aguilera-Morillo et al., 2017; Giraldo et al., 2017; Menafoglio and Secchi, 2017; Nerini and Mante, 2010; Zhang et al., 2011; Zhang et al., 2016

INTRODUCTION
METHODS TO ANALYZE DATA WITH SPATIAL DEPENDENCE

When the predictor is scalar:

- Spatial autoregressive model
- Spatial error model
- Spatial Durbin model

Spatial autoregressive model is representative of Spatial error and Spatial Durbin models.
Spatial autoregressive model in discrite data matrix:

- Case, 1991; Kelejian and Prucha, 2001; Lee, 2004, 2007; Lesage and Pace, 2009; Ord, 1975; Topa, 2001

Introduction

Methods to analyze data with spatial dependence

When the response is scalar and predictor is functional:

- Aw and Cabral, 2021; Hu et al., 2020; Huang et al., 2021; Pineda-Rios et al., 2019

Existing methods use maximum likelihood estimator to estimate the regression parameters.

Model, notations, and nomenclature

Spatial functional linear regression model

- Let $\{Y_s, \mathcal{X}_s(t), s \in \mathcal{D} \subset \mathbb{R}^d, d \geq 1\}$, where $\{Y \in \mathbb{R}, \mathcal{X}(t) \in \mathcal{L}_2(\mathcal{I}), \rightarrow$ a pair whose observations are observed over a discrete fixed subset of $\mathcal{D}$ composed of $s_1, \ldots, s_n$.
- To simplify notations, denote the spatial unit s_i by i.

Then, the spatial functional linear regression model is of the form:

$$Y = \alpha \mathbf{1}_n + \rho \boldsymbol{W} Y + \int_{\mathcal{I}} \mathcal{X}(t) \beta(t) dt + \epsilon,$$

where $\alpha \in \mathbb{R}$, $\mathbf{1}_n$ is the n-dimensional vector of ones, $\rho \in (-1, 1)$ is the unknown spatial autocorrelation parameter, $\boldsymbol{W} = (w_{ii'})_{n \times n}$, $\beta(t) \in \mathcal{L}_2(\mathcal{I})$, and $\epsilon \sim N(0, \sigma^2)$.

Model, notations, and nomenclature

Spatial weights

- $n \times n$-dimensional weight matrix $\boldsymbol{W}$ can be regularly or irregularly located on spatial domain $\mathcal{D}$.
- Let $d_{ii'} \rightarrow$ the distance between i and i' (for the model, for instance, geographic distance, economic distances, or a more general distance) .
- Let $m(\cdot) \rightarrow$ monotonically decreasing function.
- $w_{ii'} = m(d_{ii'})$.
- $w_{ii'}$s are row-normalized, i.e.,

$$w_{ii'} = \frac{m(d_{ii'})}{\sum_{i'=1}^{n} m(d_{ii'})}.$$

Model, notations, and nomenclature

Another look at the model

The spatial functional linear regression model may be re-expressed as follows:

$$Y = (\mathbb{I}_n - \rho \boldsymbol{W})^{-1}\alpha \mathbf{1}_n + (\mathbb{I}_n - \rho \boldsymbol{W})^{-1} \int_{\mathcal{I}} \mathcal{X}(t)\beta(t)dt + (\mathbb{I}_n - \rho \boldsymbol{W})^{-1}\epsilon.$$

- Y is not stationary because $\mathbb{E}[Y|\mathcal{X}(t)] \neq \int_{\mathcal{I}} \mathcal{X}(t)\beta(t)dt$.
- Y_is are spatially correlated because of $(\mathbb{I}_n - \rho \boldsymbol{W})^{-1}\epsilon$.

Model, notations, and nomenclature

Estimation: General basis expansion

- Suppose the functional predictor $\mathcal{X}(t)$ and regression coefficient function $\beta(t)$ are generated by an M-dimensional basis system $\gamma(t) = \{\gamma_m(t) : m = 1, \ldots, M\} \in \mathcal{L}_2(\mathcal{I})$.
- Then,

$$\mathcal{X}(t) = \sum_{m=1}^{M} a_m \gamma_m(t), \qquad \beta(t) = \sum_{m=1}^{M} \beta_m \gamma_m(t),$$

  where $a_m = \int_{\mathcal{I}} \mathcal{X}(t)\gamma_m(t)dt$ and $\beta_m = \int_{\mathcal{I}} \beta(t)\gamma_(t)dt$.
- The spatial functional linear regression model in the finite-dimensional space:

$$Y \approx \alpha \mathbf{1}_n + \rho \boldsymbol{W} Y \sum_{m=1}^{M} a_m \beta_m + \epsilon.$$

MODEL, NOTATIONS, AND NOMENCLATURE

ESTIMATION: FUNCTIONAL PRINCIPAL COMPONENT BASIS

- Let $\mathcal{C}(t_1, t_2) = \text{Cov}[\mathcal{X}(t_1), \mathcal{X}(t_2)]$ denote the covariance function of $\mathcal{X}(t)$ satisfying $\int_0^1 \int_0^1 \text{Cov}[\mathcal{X}(t_1), \mathcal{X}(t_2)] dt1 dt2 < \infty$.
- By Mercer's Theorem,
$$\mathcal{C} = \sum_{m \geq 1} \lambda_m \psi_m(t_1)\psi_m(t_2), \qquad \forall t_1, t_2 \in [0, \mathcal{I}],$$
$\{\psi_m(t) : m = 1, \ldots, M\}$ are orthonormal eigenfunctions in $\mathcal{L}_2(\mathcal{I})$ corresponding to the non-negative eigenvalues $\{\lambda_m : m = 1, \ldots, M\}$ with $\lambda_m \geq \lambda_{m+1}$.
- The empirical version of $\mathcal{C}(t_1, t_2)$ is $\widehat{\mathcal{C}}(t_1, t_2) = \frac{1}{n}\mathcal{X}_i(t_1)\mathcal{X}_i(t_2)$ or $\widehat{\mathcal{C}}(t_1, t_2) = \sum_{j=1}^{n} \widehat{\lambda}_j \widehat{\psi}_j(t_1)\widehat{\psi}_j(t_2)$. Thus, we have $\mathcal{X}(t) = \sum_{m=1}^{M} \widetilde{a}_m \widehat{\psi}_m(t)$, where $\widetilde{a}_m = \int_{\mathcal{I}} \mathcal{X}(t)\widehat{\psi}_m(t)dt$. Similarly, $\beta(t) = \sum_{m=1}^{M} \widetilde{\beta}_m \widehat{\psi}_m(t)$, where $\widetilde{\beta}_m = \int_{\mathcal{I}} \beta(t)\widehat{\psi}(t)dt$.
- The sample version of the model:
$$Y \approx \alpha \mathbf{1}_n + \rho \boldsymbol{W} Y \sum_{m=1}^{M} \widetilde{a}_m \widetilde{\beta}_m + \epsilon.$$

Model, notations, and nomenclature

Estimation: Functional partial-least squares basis

Step 1. Let h denote the iteration number. For $h = 1$, let $\mathcal{X}_h(t) = \mathcal{X}(t)$ and $Y_h = Y$.

Step 2. Obtain the weight function $\kappa_h(t) = \frac{\mathbb{E}[Y_h \mathcal{X}_h(t)]}{\mathbb{E}[Y_h \mathcal{X}_h(t)]\|}$ by maximizing $\text{Cov}_{\|\kappa_h(t)\|=1}[Y_h \int_{\mathcal{I}} \mathcal{X}_h(t) \kappa_h(t) dt]$, where $\|\kappa_h(t)\| = \sqrt{\int_{\mathcal{I}} \kappa_h^2(t) dt}$.

Step 3. Construct two separate regression models:

$$\mathcal{X}_h(t) = p_h(t) a_h + \epsilon_h^x, \qquad Y_h = q_h a_h + \epsilon_h^y,$$

where $_h = \int_{\mathcal{I}} \mathcal{X}_h(t) \kappa_h(t) dt$, $p_h(t) = \frac{\mathbb{E}[\mathcal{X}_h(t) a_h]}{\|a_h\|^2}$, and $q_h = \frac{Y_h a_h}{\|a_h\|^2}$.

Step 4. Go back to Step 2. (stop when $h = m$).

Estimation

Maximum likelihood estimator

Let $\boldsymbol{A} = (a_1, \ldots, a_M)^\top$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_M)^\top$, $\boldsymbol{Z} = [\mathbf{1}_n, \boldsymbol{A})$, and $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$. Then,

$$Y \approx \rho \boldsymbol{W} Y + \boldsymbol{Z}\boldsymbol{\theta} + \epsilon. \tag{1}$$

Let $\boldsymbol{\Theta} = (\boldsymbol{\theta}^\top, \sigma, \rho)^\top$, Then, the log-likelihood function for (1) is:

$$\begin{aligned}\ell(\boldsymbol{\Theta}; Y) = &-\frac{n}{2}\log(2\pi) - n\log(\sigma) + \log|\det(\boldsymbol{I}_n - \rho\boldsymbol{W})| \\ &- \frac{1}{2\sigma^2}[(\boldsymbol{I}_n - \rho\boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}]^\top[(\boldsymbol{I}_n - \rho\boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}]\end{aligned}$$

The maximum likelihood estimator is then, $\widehat{\boldsymbol{\Theta}}_{ML} = \arg\max_{\boldsymbol{\Theta}} \ell(\boldsymbol{\Theta}; Y)$.

Estimation

Maximum likelihood estimator

Let $\eta_{ML}(\boldsymbol{\Theta}; Y) = \partial\ell(\boldsymbol{\Theta}; Y)/\partial\boldsymbol{\Theta} = \mathbf{0}$, then, the estimating equation is (Tho et al., 2023):

$$\eta_{ML}(\boldsymbol{\Theta}; Y) = \begin{bmatrix} \frac{1}{\sigma^2}\boldsymbol{Z}^\top[(\boldsymbol{I}_n - \rho\boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}] \\ \frac{1}{\sigma^3}[(\boldsymbol{I}_n - \rho\boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}]^\top[(\boldsymbol{I}_n - \rho\boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}] - \frac{n}{\sigma} \\ \frac{1}{\sigma^2}(\boldsymbol{W}Y)^\top[(\boldsymbol{I}_n - \rho\boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}] - \text{trace}[\boldsymbol{W}(\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}] \end{bmatrix}$$

ESTIMATION
M-ESTIMATOR

- $\boldsymbol{r} = [(\boldsymbol{I}_n - \rho \boldsymbol{W})Y - \boldsymbol{Z}\boldsymbol{\theta}]/\sigma$
- $\varphi_c : \mathbb{R}^n \to \mathbb{R}^n \to$ Huber function with $c > 0$, i.e., $\varrho_c(r_i) = r_i \min\{1, c/|r_i|\}$.
- $\boldsymbol{G}(\rho) = \boldsymbol{W}(\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1}$.
- $\widetilde{\varrho}(c) = 2c^2[1 - \Phi(c)] - 2c\phi(c) - 1 + 2\Phi(c)$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of standard normal distribution, respectively.
- Let $\widehat{\Theta}$ is defined as the solution to $\eta_R(\boldsymbol{\Theta}; Y) = \mathbf{0}$.

ESTIMATION

M-ESTIMATOR

Let $\eta_{ML}(\boldsymbol{\Theta}; Y) = \partial\ell(\boldsymbol{\Theta}; Y)/\partial\boldsymbol{\Theta} = \mathbf{0}$, then, the estimating equation is (Tho et al., 2023):

$$\eta_R(\boldsymbol{\Theta}; Y) = \begin{bmatrix} \boldsymbol{Z}^\top \varphi_{c_1}(\boldsymbol{r}) \\ \varphi_{c_2}^\top(\boldsymbol{r})\varphi_{c_2}(\boldsymbol{r}) - n\widetilde{\varrho}(c_2) \\ \frac{1}{\sigma}[\boldsymbol{G}(\rho)\boldsymbol{Z}\boldsymbol{\theta}]^\top \varphi_{c_3}(\boldsymbol{r}) + \varphi_{c_3}^\top(\boldsymbol{r})\boldsymbol{G}^\top(\rho)\varphi_{c_3}(\boldsymbol{r}) - \text{trace}[\boldsymbol{G}(\rho)]\widetilde{\varrho}(c3) \end{bmatrix}$$

ESTIMATION

ITERATIVE ESTIMATION ALGORITHM

Step 0. **Intput:** Initial values $\widehat{\boldsymbol{\Theta}}^{[0]} = [\widehat{\boldsymbol{\theta}}^{[0]}, \widehat{\sigma}^{[0]}, \rho^{[0]}$, and the tuning parameters $[c_1, c_2, c_3] = [1.4, 2.4, 1.65$.

Step 1. At the h-th iteration, update the regression coefficient

$$\widehat{\boldsymbol{\theta}}^{[h+1]} = [\boldsymbol{Z}^\top \boldsymbol{D}_w(\widehat{\boldsymbol{\Theta}}^{[h]}; Y)\boldsymbol{Z}]^\top \boldsymbol{Z}^\top \boldsymbol{D}_w(\widehat{\boldsymbol{\Theta}}^{[h]}; Y)(Y - \rho^{[h]} \boldsymbol{W} Y),$$

where $\boldsymbol{D}_w(\widehat{\boldsymbol{\Theta}}^{[h]}; Y) = \text{diag}[\omega_1(\widehat{\boldsymbol{\Theta}}^{[h]}; Y), \ldots, \omega_n(\widehat{\boldsymbol{\Theta}}^{[h]}; Y)]$, $\omega_i(\widehat{\boldsymbol{\Theta}}^{[h]}; Y) = \varrho_{c_1}[r_i(\widehat{\boldsymbol{\Theta}}^{[h]}; Y)]r_i(\widehat{\boldsymbol{\Theta}}^{[h]}; Y)^{-1}$ if $r_i(\widehat{\boldsymbol{\Theta}}^{[h]}; Y) \neq 0$ and $\omega_i(\widehat{\boldsymbol{\Theta}}^{[h]}; Y) = 1$ if $r_i(\widehat{\boldsymbol{\Theta}}^{[h]}; Y) = 0$.

Step 2. Update the scale estimate

$$\widehat{\sigma}^{[h+1]} = \left[\frac{\widehat{\sigma}^{[h]}}{n\widetilde{\varrho}(c_2)} \varphi_{c_2}[\boldsymbol{r}(\widehat{\boldsymbol{\theta}}^{[h]}, \widehat{\sigma}^{[h]}, \rho^{[h]}; Y)]^\top [\boldsymbol{r}(\widehat{\boldsymbol{\theta}}^{[h]}, \widehat{\sigma}^{[h]}, \rho^{[h]}; Y)] \right]^{1/2}.$$

Step 3. Update spatial autocorrelation estimate by solving

$$\rho^{[h+1]} = \arg\min_{\rho} [\eta_R(\rho; \widehat{\boldsymbol{\theta}}^{[h+1]}, \widehat{\sigma}^{[h+1]}, Y)]^2.$$

Repeat Step 1. - Step 3. until convergence, where convergence is achieved when $\|\widehat{\boldsymbol{\Theta}}^{[h+1]} - \widehat{\boldsymbol{\Theta}}^{[h]}\|_2 < \varepsilon$ for some sufficiently small ε

ESTIMATION
PREDICTION

The M-estimator for $\beta(t)$ is $\widehat{\beta}(t) = \sum_{m=1}^{M} \widehat{\beta}_m \gamma_m(t)$.
Asymptotic properties:

- Consistency?
- Rates of convergence?

Prediction:

$$\widehat{Y} = (\mathbb{I}_n - \widehat{\rho}\boldsymbol{W})^{-1}\widehat{\alpha}\mathbf{1}_n + (\mathbb{I}_n - \widehat{\rho}\boldsymbol{W})^{-1}\int_{\mathcal{I}} \mathcal{X}(t)\widehat{\beta}(t)dt.$$

MONTE-CARLO EXPERIMENTS

- Proposed method: RfSAC
- Maximum likelihood based methods:
  - Basis expansion + maximum likelihood estimate: fSAC
  - Functional principal component basis + maximum likelihood estimate: fPCA
  - Functional partial least squares basis + maximum likelihood estimate: fPLS
- Functional linear regression: fLM

MONTE-CARLO EXPERIMENTS

DATA GENERATION PROCESS

- $\mathcal{X}(t) = \sum_{j=1}^{5} \zeta_j \nu_j(t)$, where $\zeta_j \sim N(0, 4j^{-3/2})$ and $\nu_j(t) = \sin(j\pi t) - \cos(j\pi t)$.
- Trajectories are generated at 101 equally spaced point in the interval $[0, 1]$.
- $\beta(t) = \sin(2\pi t)$.
- $\boldsymbol{W} = (w_{ii'})_{n\times n}$, where $w_{ii'} = (1/d_{ii'})/(\sum_{i'=1}^{n} 1/d_{ii'})$ for $i \neq i'$ and zero otherwise, and $d_{ii'} = |i - i'|$.
- $\rho_0 = 0.5$.
- $\epsilon \sim N(0, \sigma^2)$.
- $\alpha = 0$.
- $$Y = (\mathbb{I}_n - \rho\boldsymbol{W})^{-1} \int_{\mathcal{I}} \mathcal{X}(t)\beta(t)dt + (\mathbb{I}_n - \rho\boldsymbol{W})^{-1}\epsilon.$$

Monte-Carlo experiments

Metrics

100 Monte-Carlo runs are performed.

- $\text{ISE} = \int_0^1 [\beta(t) - \widehat{\beta}(t)]^2 dt$.
- $\text{Bias}_\rho = \widehat{\rho} - \rho_0$.
- $\text{Bias}_\sigma = \widehat{\sigma} - \sigma_0$.

In each experiment, $n_{\text{test}} = 100$ independent samples are generated as the test sample.

- $\text{MSPE} = \frac{1}{100} \sum_{i=1}^{100} (\widehat{Y}_i - Y_i)^2$.

Monte-Carlo experiments

Scenario-I

- $\sigma_0 = 0.5$.
- $n = [50, 100, 250, 500]$

Monte-Carlo experiments

Scenario-I

Table. Computed mean ISE, Bias_ρ, Bias_σ, and MSPE values over 100 Monte-Carlo replications for different sample sizes.

n	Metric	Method				
		fLM	fSAC	fPCA	fPLS	RfSAC
$n = 50$	ISE	0.106	0.089	**0.082**	0.104	0.092
	Bias_ρ	–	**0.073**	0.077	**0.073**	**0.073**
	Bias_σ	–	0.049	**0.042**	0.049	0.049
	MSPE	0.490	0.350	**0.346**	0.350	0.362
$n = 100$	ISE	0.063	**0.058**	0.066	0.075	0.060
	Bias_ρ	–	0.056	0.056	0.056	**0.054**
	Bias_σ	–	0.015	**0.010**	0.015	0.017
	MSPE	0.411	0.297	**0.294**	0.297	0.297
$n = 250$	ISE	0.041	**0.040**	0.056	0.053	**0.040**
	Bias_ρ	–	0.022	0.023	0.022	**0.018**
	Bias_σ	–	0.009	**0.007**	0.009	0.010
	MSPE	0.330	**0.266**	**0.266**	**0.266**	**0.266**
$n = 500$	ISE	0.033	**0.032**	0.053	0.046	**0.032**
	Bias_ρ	–	**0.014**	**0.014**	**0.014**	0.015
	Bias_σ	–	0.049	**0.042**	0.049	0.049
	MSPE	0.328	**0.250**	**0.250**	**0.250**	**0.250**

Monte-Carlo experiments

Scenario-II

- $n = 250$.
- $\sigma_0 = [0.25, 0.50, 0.75, 1.00]$

Monte-Carlo experiments

Scenario-II

Table. Computed mean ISE, Bias_ρ, Bias_σ, and MSPE values over 100 Monte-Carlo replications for different σ_0 values.

σ_0	Metric	Method				
		fLM	fSAC	fPCA	fPLS	RfSAC
$\sigma_0 = 0.25$	ISE	0.032	**0.029**	0.051	0.043	**0.029**
	Bias_ρ	–	$<$ 0.000	$<$ 0.000	$<$ 0.000	$<$ 0.000
	Bias_σ	–	0.007	**0.004**	0.007	0.007
	MSPE	0.154	**0.067**	0.069	**0.067**	**0.067**
$\sigma_0 = 0.50$	ISE	0.063	**0.058**	0.066	0.075	0.060
	Bias_ρ	–	0.056	0.056	0.056	**0.0540**
	Bias_σ	–	0.015	**0.010**	0.015	0.017
	MSPE	0.411	0.297	**0.294**	0.297	0.297
$\sigma_0 = 0.75$	ISE	0.056	**0.052**	0.063	0.064	0.053
	Bias_ρ	–	0.032	0.032	0.032	**0.029**
	Bias_σ	–	0.013	**0.011**	0.013	0.014
	MSPE	0.678	0.625	0.625	0.625	**0.620**
$\sigma_0 = 1.00$	ISE	0.092	0.093	**0.083**	0.101	0.099
	Bias_ρ	–	0.096	0.096	0.096	**0.091**
	Bias_σ	–	0.009	**0.007**	0.009	0.009
	MSPE	1.194	1.112	**1.110**	1.112	1.111

MONTE-CARLO EXPERIMENTS

SCENARIO-III, CASE-I

The generated scalar response is contaminated by outliers at $[1\%,\ 5\%,\ 10\%]$ contamination levels (C_L).

- $n = 50$.
- $rho_0 = 0.5$.
- $\sigma_0 = 0.5$.
- $\epsilon \sim (1 - C_L)N(0, \sigma_0^2) + C_L N(5, 4)$.

Monte-Carlo experiments

Scenario-III, Case-I

Table. Computed mean ISE, Bias_ρ, Bias_σ, and MSPE values over 100 Monte-Carlo replications for $n = 50$ and different C_Ls.

σ_0	Metric	Method				
		fLM	fSAC	fPCA	fPLS	RfSAC
$C_L = 0\%$	ISE	0.106	**0.089**	0.082	0.104	0.092
	Bias_ρ	–	**0.073**	0.077	**0.073**	**0.073**
	Bias_σ	–	0.049	**0.042**	0.049	0.049
	MSPE	0.490	0.350	**0.346**	0.350	0.362
$C_L = 1\%$	ISE	0.112	0.087	**0.079**	0.104	0.092
	Bias_ρ	–	0.049	**0.046**	0.049	0.048
	Bias_σ	–	0.038	**0.033**	0.038	0.039
	MSPE	0.438	0.325	**0.316**	0.325	0.336
$C_L = 5\%$	ISE	0.460	0.443	0.216	0.338	**0.109**
	Bias_ρ	–	0.199	0.189	0.199	**0.093**
	Bias_σ	–	-0.541	-0.554	-0.541	**-0.028**
	MSPE	0.683	0.638	0.587	0.638	**0.377**
$C_L = 10\%$	ISE	1.595	1.625	0.624	0.934	**0.291**
	Bias_ρ	–	0.303	0.308	0.303	**0.105**
	Bias_σ	–	-1.086	-1.117	-1.086	**-0.247**
	MSPE	2.039	1.917	1.757	1.917	**0.520**

MONTE-CARLO EXPERIMENTS

SCENARIO-III, CASE-II

- $n = 250$.
- $rho_0 = 0.5$.
- $\sigma_0 = 0.5$.
- $\epsilon \sim (1 - C_L)N(0, \sigma_0^2) + C_L N(5, 4)$.

Monte-Carlo experiments

Scenario-III, Case-II

Table. Computed mean ISE, Bias_ρ, Bias_σ, and MSPE values over 100 Monte-Carlo replications for $n = 250$ and different C_Ls.

σ_0	Metric	Method				
		fLM	fSAC	fPCA	fPLS	RfSAC
$C_L = 0\%$	ISE	0.041	**0.040**	0.056	0.053	**0.040**
	Bias_ρ	–	0.022	0.023	0.022	**0.018**
	Bias_σ	–	0.009	**0.007**	0.009	0.010
	MSPE	0.330	**0.266**	**0.266**	**0.266**	**0.266**
$C_L = 1\%$	ISE	0.056	0.052	0.062	0.065	**0.044**
	Bias_ρ	–	0.036	0.034	0.036	**0.014**
	Bias_σ	–	-0.176	-0.178	-0.176	**-0.001**
	MSPE	0.396	0.303	0.303	0.303	**0.293**
$C_L = 5\%$	ISE	0.125	0.124	0.099	0.124	**0.044**
	Bias_ρ	–	0.064	0.065	0.064	**0.013**
	Bias_σ	–	-0.717	-0.720	-0.717	**-0.072**
	MSPE	0.639	0.578	0.574	0.578	**0.309**
$C_L = 10\%$	ISE	0.240	0.235	0.168	0.218	**0.058**
	Bias_ρ	–	0.021	0.021	0.021	**-0.021**
	Bias_σ	–	-1.165	-1.168	-1.165	**-0.252**
	MSPE	1.575	1.363	1.356	1.363	**0.337**

MONTE-CARLO EXPERIMENTS

SCENARIO-III, CASE-III

- $n = 500$.
- $\rho_0 = 0.5$.
- $\sigma_0 = 0.5$.
- $\epsilon \sim (1 - C_L)N(0, \sigma_0^2) + C_L N(5, 4)$.

MONTE-CARLO EXPERIMENTS

SCENARIO-III, CASE-III

Table. Computed mean ISE, Bias_ρ, Bias_σ, and MSPE values over 100 Monte-Carlo replications for $n = 500$ and different C_Ls.

σ_0	Metric	Method				
		fLM	fSAC	fPCA	fPLS	RfSAC
$C_L = 0\%$	ISE	0.033	**0.032**	0.053	0.046	**0.032**
	Bias_ρ	–	**0.014**	**0.014**	**0.014**	0.015
	Bias_σ	–	0.049	**0.042**	0.049	0.049
	MSPE	0.328	**0.250**	**0.250**	**0.250**	**0.250**
$C_L = 1\%$	ISE	0.067	0.063	0.065	0.074	**0.054**
	Bias_ρ	–	0.033	0.033	0.033	**0.013**
	Bias_σ	–	-0.202	-0.205	-0.202	**0.003**
	MSPE	0.396	0.312	0.310	0.312	**0.286**
$C_L = 5\%$	ISE	0.071	0.069	0.073	0.080	**0.033**
	Bias_ρ	–	0.039	0.040	0.039	**0.008**
	Bias_σ	–	-0.755	-0.757	-0.755	**-0.084**
	MSPE	0.590	0.530	0.523	0.530	**0.272**
$C_L = 10\%$	ISE	0.114	0.109	0.083	0.109	**0.038**
	Bias_ρ	–	0.077	0.077	0.077	**0.018**
	Bias_σ	–	-1.178	-1.180	-1.178	**-0.258**
	MSPE	1.493	1.387	1.385	1.387	**0.360**

Spanish weather data example

- The dataset includes averaged daily log-precipitation (scalar response) and temperature (functional predictor) from 73 weather stations in Spain from 1980 to 2009.
- Observations of the temperature considered the functions of days ($1 \leq t \leq t$.
- **Aim:** explore the functional relationship between the log-precipitation and temperature taking into account the spatial dependence between weather stations.

SPANISH WEATHER DATA EXAMPLE

Figure. The locations of 73 weather stations on map of Spain.

Spanish weather data example

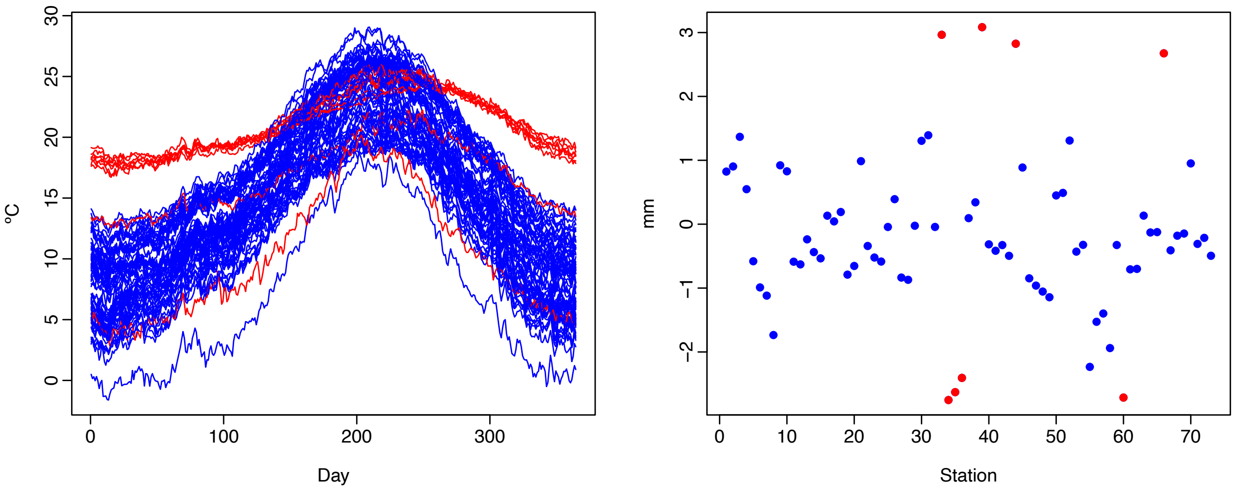

Figure. Graphical displays of the Spanish weather data; temperature (left panel) and log-precipitation (right panel).

SPANISH WEATHER DATA EXAMPLE

- The distances between weather stations are computed using the great circle formula based on the geographic coordinates (latitudes and longitudes) of the weather stations. The weight matrix $\boldsymbol{W}$ is then constructed via the bisquare weight function, where weights are $w_{ii'} = K(d_{ii'}/h)$ for $d_{ii'} < h$ and $w_{ii'} = 0$ for $i = i'$ and $d_{ii'} > h$.
- Divide the data into two parts: a training sample with size $n_{\text{train}} = 50$ and a test sample with size $n_{\text{test}} = 50$.
- Build model on training sample to predict log-precipitation in test sample.
- Repeat the above process 100 times and compute MSPE.

SPANISH WEATHER DATA EXAMPLE

Table. Computed mean *MSPE* values and their standard errors (SE) over 100 Monte-Carlo replications for the Spanish weather data.

Metric	Method				
	fLM	fSAC	fPCA	fPLS	RfSAC
MSPE	0.585	0.585	0.512	0.487	**0.317**
SE	0.178	0.179	0.153	0.1543	**0.122**

Conclusion

- Scalar-on-function linear regression model vs spatial functional linear regression model.
- Robust and efficient.
- Limitation?

References I

Aguilera-Morillo, M. C., Durban, M., & Aguilera, A. M. (2017). Prediction of functional data with spatial dependence: A penalized approach. *Stoch Environ Res Risk Assess*, *31*(1), 7–22.

Ait-Saidi, A., Ferraty, F., Kassa, R., & Vieu, P. (2009). Cross-validated estimations in the single-functional index model. *Statistics*, *42*(6), 475–494.

Anselin, L. (1998). *Spatial econometrics: Methods and models*.

Aw, A., & Cabral, E. N. (2021). Functional SAC model: With application to spatial econometrics. *S Afr Stat J*, *55*(1), 1–13.

Case, A. C. (1991). Spatial patterns in household demand. *Econometrica*, *59*(4), 953–965.

Crainiceanu, C. M., Staicu, A.-M., & Di, C.-Z. (2009). Generalized multilevel functional regression. *J Am Stat Assoc*, *104*(488), 1550–1561.

Cressie, N., & Wikle, C. K. (2015). *Statistics for spatio-temporal data*.

Fang, Y., Yuejiao, F., & Lee, T. C. M. (2009). Functional mixture regression. *Biostatistics*, *12*(2), 341–353.

Ferraty, F., & Vieu, P. (2006). *Nonparametric functional data analysis: Theory and practice*.

Giraldo, R., Delicado, P., & Mateu, J. (2017). Spatial prediction of a scalar variable based on data of a functional random field. *Comunicaciones en Estadistica*, *10*(2), 315–344.

Goldsmith, J., Crainiceanu, C. M., Caffo, B., & Reich, D. (2012). Longitudinal penalized functional regression for cognitive outcomes on neuronal tract measurements. *J R Stat Soc*, *61*(3), 453–469.

References II

Hu, Y., Wu, S., Feng, S., & Jin, J. (2020). Estimation in partial functional linear spatial autoregressive model. *Mathematics*, *8*, 1680.

Huang, T., Saporta, G., Wang, H., & Wang, S. (2021). A robust spatial autoregressive scalar-on-function regression with t-distribution. *Adv Data Anal Classif*, *15*, 57–81.

James, G. M. (2002). Generalized linear models with functional predictors. *J R Stat Soc Ser B (Stat Methodol)*, *64*(3), 411–432.

James, G. M. (2005). Functional adaptive model estimation. *J Am Stat Assoc*, *100*(470), 565–576.

Kelejian, H. H., & Prucha, I. R. (2001). A generalized moments estimator for the autoregressive parameter in a spatial model. *Int Econ Rev*, *40*(2), 509–533.

Lee, L.-F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, *72*(6), 1899–1925.

Lee, L.-F. (2007). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *J Econ*, *137*(2), 489–514.

Lesage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*.

Marx, B. D., & Eilers, P. H. C. (1999). Generalized linear regression on sampled signals and curves: A P-spline approach. *Technometrics*, *41*(1), 1–13.

Menafoglio, A., & Secchi, P. (2017). Statistical analysis of complex and spatially dependent data: A review of object oriented spatial statistics. *Eur J Oper Res*, *258*(2), 401–410.

Morris, J. S. (2015). Functional regression. *Ann Rev Stat Appl*, *2*(1), 321–359.

References III

- Muller, H.-G., Wu, Y., & Yao, F. (2013). Continuously additive models for nonlinear functional regression. *Biometrika*, *100*(3), 607–622.
- Nerini, D., & Mante, P. (2010). Cokriging for spatial functional data. *J Multivar Anal*, *101*(2), 409–418.
- Ord, K. (1975). Estimation methods for models of spatial interaction. *J Am Stat Assoc*, *70*(349), 120–126.
- Pineda-Rios, W., Giraldo, R., & Porcu, E. (2019). Functional sar models: With application to spatial econometrics. *Spat Stat*, *29*, 145–159.
- Reiss, P. T., Goldsmith, J., Shang, H. L., & Ogden, R. T. (2017). Methods for scalar-on-function regression. *Int Stat Rev*, *85*(2), 228–249.
- Schabenberger, O., & Gotway, C. A. (2017). *Statistical methods for spatial data analysis*.
- Tho, Z. Y., Francis, D. D., Hui, K. C., Welsh, A. H., & Zou, T. (2023). On the robust estimation of spatial autoregressive models. *Econom Stat*, in press.
- Topa, G. (2001). Social interactions, local spillovers and unemployment. *Rev Econ Stud*, *68*(2), 261–295.
- Wang, H., Gu, J., Wang, S., & Saporta, G. (2019). Spatial partial least squares autoregression: Algorithm and applications. *Chemom Intell Lab Syst*, *184*, 123–131.
- Yao, F., & Muller, H.-G. (2010). Functional quadratic regression. *Biometrika*, *97*(1), 49–61.
- Zhang, J., Clayton, M. K., & Townsend, P. A. (2011). Functional concurrent linear regression model for spatial images. *J Agric Biol Environ Stat*, *16*(1), 105–130.

References IV

Zhang, L., Baladandayuthapani, V., Zhu, H., Baggerly, K. A., Majewski, T., Czerniak, B. A., & Morris, J. S. (2016). Functional car models for large spatially correlated functional datasets. *J Am Stat Assoc*, *111*(514), 772–786.