USDA ERS Broadband: *Comparison of boundary analysis methods for spatial models.*

Abstract

This a weekly updated memo documenting progress in statistical modeling for the USDA ERS Broadband Project. Currently we aim to perform exploratory analysis, synthetic experiments comparing methods that detect geographical boundaries separating response behavior within a fixed spatial reference domain. We aim to also decide on a modeling framework suitable for the CoreLogic data.

Meeting: 10/21/2020

We aim to perform synthetic experiments showcasing the performance of spatial regression discontinuity design (see Lee and Lemieux (2010), Hidano et al. (2015), Keele and Titiunik (2015)) under a variety of spatial patterns. All the experiments showcased compare regression discontinuity designs under (a) ordinary least squares, (b) two stage least squares with spatially lagged errors (Hidano et al. (2015), Kelejian and Prucha (1998)) (c) *parametric* fully Bayesian hierarchical spatial regression (Finley et al. (2007)) modeling frameworks. The reason behind the choice of a parametric spatial regression is motivated from a desired setup for comparing regression discontinuity designs with methods for wombling (Womble (1951)). We consider the synthetic data generating process (outlined in Hidano et al. (2015))

$$Y = \beta_0 + \alpha X + \delta\left(sin(\overline{lat}) + cos(\overline{long})\right) + \epsilon,$$

where $\beta_0 = 60$, $\alpha = -1.5$, $\delta = 2.5$, $\epsilon \sim N(0,1)$, $\overline{lat}$ and $\overline{long}$ are the standardized latitude and longitude respectively. This introduces location specific effects to within the response. X is the covariate determining the spatial location based discontinuity, i.e. it is 0 or 1 based on the location of the response. We select $\mathbf{s} = (s_x, s_y)^\top \in (0,20)\times(0,20) \subset \mathbb{R}^2$ as the spatial reference domain. The synthetic patterns considered are shown in figure 1. For the patterns shown the X variable is selected as follows,

1. Sharp Regression Discontinuity:

   (a) $X = 0 \cdot I(s_x < 10) + 1 \cdot I(s_x \geq 10)$, (Open Curve)

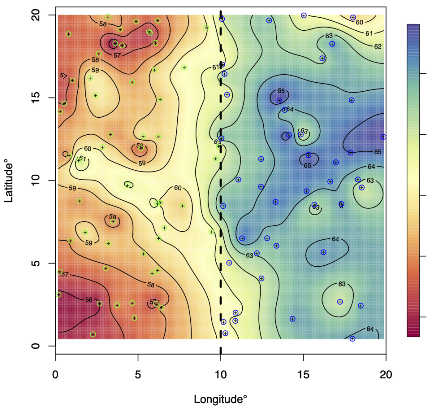
   (b) $X = 1 \cdot I(s_y < s_x) + 0 \cdot I(s_y \geq s_x)$, (Open Curve)

   (d) $X = 1 \cdot I(s_y < s_x^3/400) + 0 \cdot I(s_y \geq s_x^3/400)$, (Open Curve)

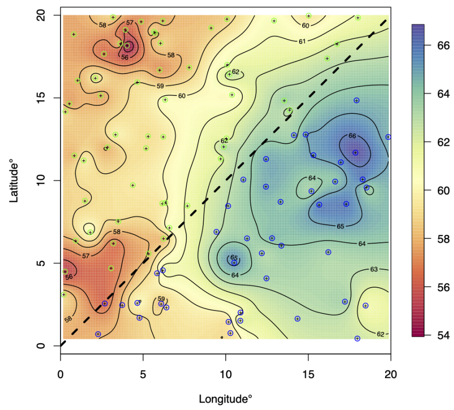
   (e) $X = 1 \cdot I\left((s_y - 10)^2 + (s_x - 10)^2 < 36\right) + 0 \cdot I\left((s_y - 10)^2 + (s_x - 10)^2 \geq 36\right)$ (Closed Curve)

2. Fuzzy Regression Discontinuity:

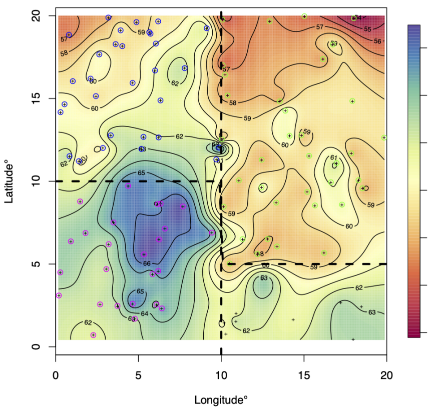
- $X = -5 \cdot I(s_y > 10, s_x \geq 5) + 0.5 \cdot I(s_y < 10, s_x \geq 10) + 5 \cdot I(s_y < 10, s_x < 10) + 0 \cdot I(s_y > 10, s_x < 5)$, (Open Curve)



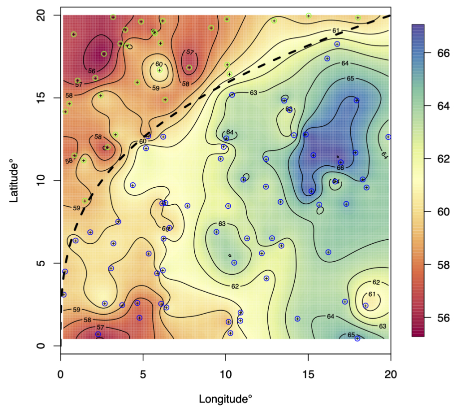
(a) $x = 10$



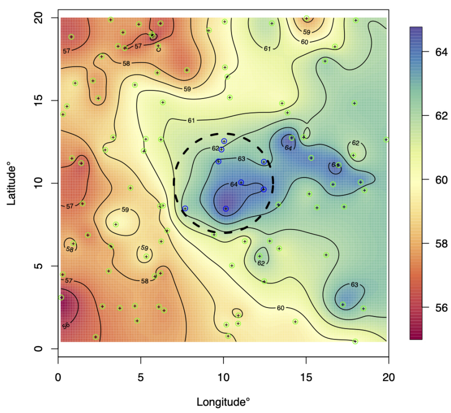
(b) $y = x$



(c) $x = 10, y = 10I(0 < y < 10) + 5I(10 < y < 20)$



(d) $y = x^3/400$



(e) $(y - 10)^2 + (x - 10)^2 = 6^2$

Figure 1: Synthetic spatial patterns.

The models being compared are:

1. OLS Model: $Y = \beta_0 + \alpha X + \epsilon$, $\epsilon \sim N(0, \sigma^2)$

2. OLS Model: $Y = \beta_0 + \alpha X + \lambda WY + u$, $u = \rho Mu + \epsilon$, $\epsilon \sim N(0, \sigma^2)$, $|\lambda| < 1$,$|\rho| < 1$, where M and W are spatial loading matrices (Kelejian and Prucha (1998)) eq. 1.

3. Spatial Model: $Y(\mathbf{s}) = \beta_0 + \alpha X + \mathbf{Z}(\mathbf{s}) + \epsilon$, $\mathbf{Z}(\mathbf{s}) \sim N(0, \Sigma)$, $\epsilon \sim N(0, \tau^2)$, $\Sigma = \sigma^2 \exp(-\phi_s ||\mathbf{s} - \mathbf{s}'||)$, where $\mathbf{s}, \mathbf{s}'$ are locations within the random field. Σ corresponds to the exponential covariance kernel.

It is important to note that the accuracy in estimating the regression discontinuity (if present) is highly dependent on the predictive power of the model. To that end, we use the root mean square error (RMSE) to measure model performance under the various synthetic patterns. Under each pattern we performed 10 replications and the number of locations ranged between $\{100, 200, 500\}$. Tables 1 and 2 shows the superiority in performance for the parametric spatial model based on RMSE.

Table 1: Table showing RMSE and associated standard deviation (in brackets) for estimated parameters under the three different modeling frameworks for synthetic patterns pertaining to sharp regression discontinuity.

Pattern (Sharp Reg. Discontinuity)	N	OLS				2SLS				Spatial			
		β_0	α	Fitted	Cont. Prob.	β_0	α	Fitted	Cont. Prob.	β_0	α	Fitted	Cont. Prob.
$x = 10$	100	3.050 (0.262)	3.326 (0.222)	1.563 (0.074)	0.000	6.975 (0.877)	3.182 (0.403)	1.420 (0.066)	0.000	1.032 (0.341)	1.036 (0.424)	0.800 (0.071)	0.800
	200	3.136 (0.112)	3.390 (0.168)	1.548 (0.058)	0.000	7.156 (0.765)	3.322 (0.215)	1.395 (0.055)	0.000	0.459 (0.202)	0.693 (0.337)	0.798 (0.031)	0.700
	500	3.107 (0.077)	3.399 (0.131)	1.543 (0.036)	0.000	6.998 (0.186)	3.269 (0.109)	1.398 (0.034)	0.000	0.333 (0.166)	0.514 (0.296)	0.882 (0.025)	0.600
$y = x$	100	2.579 (0.248)	2.249 (0.361)	1.941 (0.149)	0.000	6.254 (1.292)	2.198 (0.406)	1.836 (0.151)	0.000	0.670 (0.418)	0.530 (0.349)	0.734 (0.092)	1.000
	200	2.511 (0.125)	2.111 (0.168)	2.047 (0.091)	0.000	6.908 (0.626)	2.108 (0.223)	1.919 (0.097)	0.000	0.272 (0.251)	0.332 (0.236)	0.823 (0.051)	1.000
	500	2.493 (0.096)	2.115 (0.170)	2.019 (0.041)	0.000	6.883 (0.472)	2.002 (0.138)	1.877 (0.047)	0.000	0.236 (0.151)	0.150 (0.103)	0.861 (0.029)	1.000
$y = x^3/400$	100	2.034 (0.134)	2.571 (0.255)	1.960 (0.090)	0.000	4.480 (1.696)	2.263 (0.400)	1.905 (0.099)	0.000	0.745 (0.249)	1.092 (0.377)	0.759 (0.055)	0.700
	200	1.968 (0.126)	2.281 (0.302)	2.077 (0.065)	0.000	5.390 (1.194)	1.800 (0.229)	1.995 (0.082)	0.000	0.244 (0.211)	0.277 (0.213)	0.803 (0.055)	1.000
	500	2.022 (0.042)	2.444 (0.114)	2.059 (0.061)	0.000	4.844 (0.718)	2.125 (0.09)	2.001 (0.067)	0.000	0.275 (0.2)	0.2700 (0.216)	0.886 (0.025)	0.900
$(y - 10)^2 + (x - 10)^2 = 3^2$	100	2.159 (0.224)	1.040 (0.383)	2.193 (0.221)	0.600	5.697 (3.582)	1.259 (0.754)	2.045 (0.145)	0.300	0.574 (0.510)	0.400 (0.338)	0.768 (0.065)	0.900
	200	2.199 (0.263)	1.092 (0.315)	2.214 (0.077)	0.700	5.824 (2.228)	0.604 (0.620)	2.131 (0.102)	0.100	0.387 (0.361)	0.299 (0.271)	0.817 (0.052)	0.900
	500	2.209 (0.119)	1.101 (0.176)	2.232 (0.065)	0.700	5.928 (1.175)	0.317 (0.287)	2.164 (0.058)	0.000	0.265 (0.140)	0.209 (0.145)	0.873 (0.028)	0.900

Table 2: Table showing RMSE and associated standard deviation (in brackets) for estimated parameters under the three different modeling frameworks for a synthetic pattern pertaining to fuzzy regression discontinuity.

Pattern (Fuzzy Reg. Discontinuity)	N	OLS						2SLS						Spatial					
		β_0	α_1	α_2	α_3	Fitted	Cont. Prob.	β_0	α_1	α_2	α_3	Fitted	Cont. Prob.	β_0	α_1	α_2	α_3	Fitted	Cont. Prob.
$x = 10, y = 10I(0 < y < 10) + 5I(10 < y < 20)$	100	2.139 (0.326)	1.219 (0.469)	2.437 (0.403)	2.339 (0.330)	1.489 (0.115)	0.200	6.339 (1.558)	0.822 (0.556)	2.716 (0.409)	2.705 (0.384)	1.355 (0.108)	0.100	0.660 (0.429)	0.840 (0.557)	0.763 (0.480)	0.715 (0.422)	0.738 (0.072)	0.966
	200	2.445 (0.291)	0.891 (0.363)	2.661 (0.508)	2.630 (0.193)	1.517 (0.050)	0.233	6.724 (0.788)	0.523 (0.270)	2.955 (0.442)	3.023 (0.193)	1.375 (0.070)	0.066	0.421 (0.346)	0.343 (0.205)	0.618 (0.465)	0.660 (0.365)	0.822 (0.034)	0.966
	500	2.218 (0.226)	1.130 (0.213)	2.386 (0.222)	2.508 (0.310)	1.505 (0.031)	0.033	5.970 (0.686)	0.726 (0.223)	2.744 (0.252)	2.870 (0.339)	1.403 (0.040)	0.000	0.352 (0.199)	0.298 (0.237)	0.257 (0.260)	0.277 (0.192)	0.852 (0.033)	1.000

For the Bayesian hierarchical spatial model, we have the following parameters, $\boldsymbol{\theta} = \{\boldsymbol{\beta}^\top, \phi_s, \sigma^2, \tau^2\}$, where $\boldsymbol{\beta} = (\beta_0, \alpha)^\top$ with the posterior likelihood being specified by

$$
\begin{aligned}
p(\boldsymbol{\theta}|y) \propto U(\phi_s|a_\phi, b_\phi) \times IG(\sigma^2|a_\sigma, b_\sigma) \times IG(\tau^2|a_\tau, b_\tau) \times N(\boldsymbol{\beta}|\mu_\beta, \Sigma_\beta) \times N(\mathbf{Z}|0, \Sigma(\sigma^2, \phi_s)) \\
\times \prod N(y|\beta_0 + \alpha X, \tau^2)
\end{aligned}
\tag{1}
$$

We use the following hyperparameter specifications: $a_\phi = 0$, $b_\phi = 30$, $a_\sigma = 2, b_\sigma = 1$, $a_\tau = 2, b_\tau = 0.1$, $\mu_\beta = (0, 0)^\top$, $\Sigma_\beta = diag\{1 \times 10^{-3}, 1 \times 10^{-3}\}$. The posterior for ϕ_s is sampled using an *adaptive Metropolis within Gibbs sampler*. It was observed that for each simulation the true value $\tau^2 = 1$ was contained within the highest posterior density (HPD) intervals for the Bayesian hierarchical spatial model.

Scaleable Spatial Processes

We use two approximate spatial processes to provide scaleability to the spatial models viz. Nearest Neighbor Gaussian Processes (NNGP) and Integrated Nested Laplace Approximation (INLA). A synthetic comparison for a sharp regression discontinuity design $x = 10$ is shown below.

Table 3: Table showing RMSE for estimates from NNGP and INLA for a shrap regression discontinuity pattern, $x = 10$.

N	INLA		NNGP		Prediction	
	β	α	β	α	INLA	NNGP
1000	0.353	0.111	0.209	0.089	2.258	2.258
	(0.070)	(0.075)	(0.075)	(0.071)	0.029	0.029
5000	0.247	0.038	0.139	0.036	2.248	2.250
	(0.033)	(0.036)	(0.078)	(0.076)	0.013	0.014
10000	0.113	0.022	0.061	0.016	**4.498**	2.251
	(0.044)	(0.050)	(0.108)	(0.068)	**7.114**	0.005

Wombling

For the purpose of synthetic illustrations for wombling we require differentiability of the underlying random field (as opposed to just continuity in the RD scenario). This is established by using differentiable kernels while constructing Σ for the spatial models. In terms of kernel choices we have,

$$
\Sigma = \begin{cases} \tilde{K}(||\Delta||) = \sigma^2 \exp(-\phi_s ||\Delta||^\nu), & \nu \in [0, 2], \text{ (Power exponential Class)} \\ \tilde{K}(||\Delta||) = \sigma^2 (\phi_s ||\Delta||)^\nu K_\nu(\phi_s ||\Delta||). & \text{(Matérn Class).} \end{cases}
$$

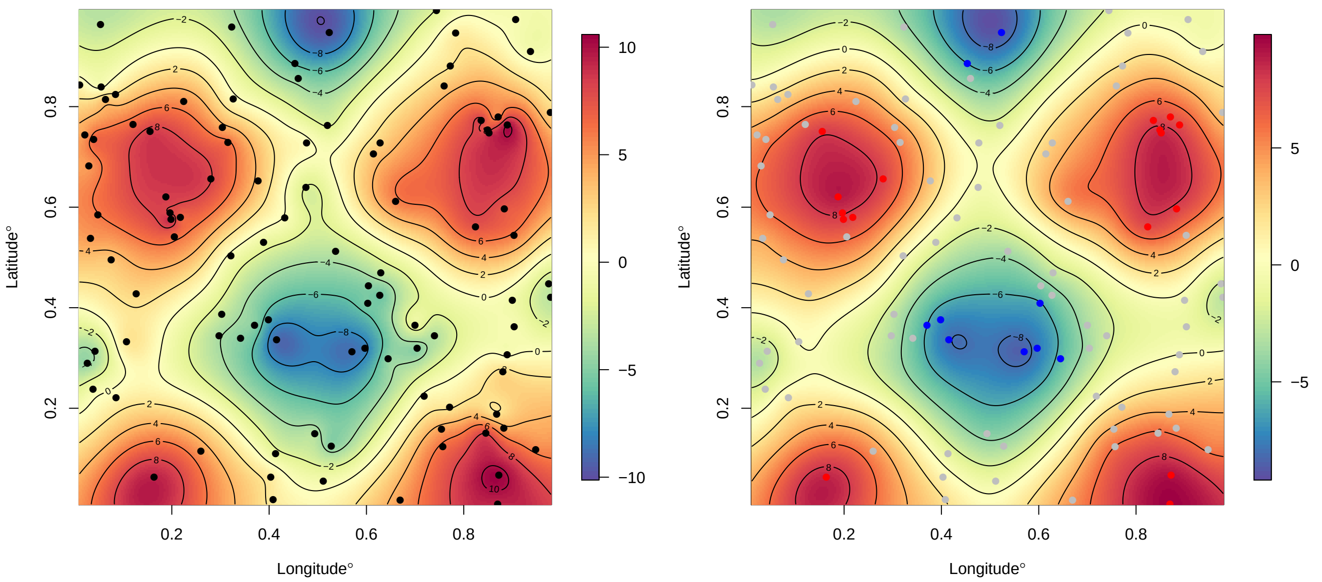
ν controls the smoothness for process realizations. For the power exponential class $\nu = 1$ is the exponential kernel which is just continuous admitting no derivatives, $\nu = 2$ is the squared exponential or gaussian covariance, which is infinitely differentiable. Within the Matérn class we have $\nu = 3/2$ admitting gradients and $\nu = 5/2$ admitting both first and second order gradients. If $\nu \to \infty$ we obtain the Gaussian covariance.

Furthermore, to check the validity of our estimation of gradients it is beneficial to have an analytic function that has a closed form analytic expression of the gradient. We select the following form within the spatial reference domain $\mathbf{s} = (s_x, s_y)^\top \in (0,1) \times (0,1)$ in $\mathbb{R}^2$.

$$Y \sim N(5(\sin(3\pi s_x) + \cos(3\pi s_y)), \tau^2), \qquad\qquad \tau^2 = 1$$

Fig. (2) (a) shows the synthetic spatial pattern. Before we demonstrate gradient estimation, a spatial model with a Matérn kernel with $\nu = 5/2$ is fit to the data. The advantage of fitting such a kernel is that both first and second order gradients exist for process realizations. Given the choice of our synthetic data generating process we should have,

$$\nabla_x Y = 15\pi \cos(3\pi s_x), \qquad \nabla_y Y = -15\pi \sin(3\pi s_y)$$
$$\nabla_{xx} Y = -45\pi^2 \sin(3\pi s_x), \quad \nabla_{yy} Y = -45\pi^2 \cos(3\pi s_y), \quad \nabla_{xy} Y = 0.$$

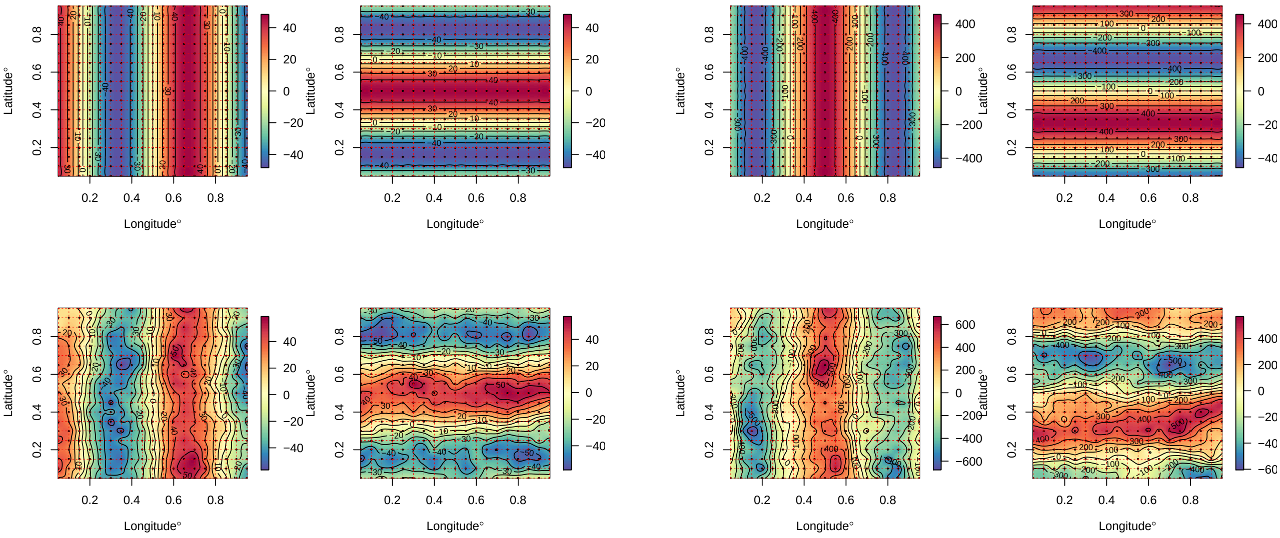


(a) (b)

Figure 2: Interpolated spatial plots showing (a) the true pattern (b) the estimated fit from a spatial model with a Matérn($\nu = 5/2$) kernel. Significant random effects in (b) are color coded accordingly based on their algebraic sign

From the fitted spatial process, we extract posterior samples of process parameters $\{\sigma^2, \phi_s, \tau^2, \mathbf{Z}(\mathbf{s})\}$ to estimate the gradients at regularly spaced points on a grid post-MCMC.

In a realistic setup wombling amounts to locating significant gradients and curvature along points on a curve, once such a curve is available we could replace the points on the grid with points on the curve to



(a) (b)

Figure 3: Interpolated spatial plots showing (a) the true and estimated first order gradients, ∇_x and ∇_y (b) the true and estimated pure second order gradients, ∇_{xx} and ∇_{yy} over 360 points on a regular grid (marked using a dashed red line).

compute gradients/curvature along its normal direction.

Validity of the Continuity assumption: a model-based approach

Along with the conventional distance to boundary approach, using model based information criteria to inform about continuity/differentiability of the random field. For instance, for the first sharp discontinuity pattern we have

(a) Exponential Kernel (Continuous only) **AIC=499.32, BIC = 772.87, DIC= 289.32**

(b) Matérn($\nu = 3/2$) (Derivative exists) AIC=547.94, BIC = 821.48, DIC= 337.94

(c) Matérn($\nu = 5/2$) (Curvature exists) AIC=566.64, BIC = 840.18, DIC= 356.64

(d) Gaussian (Infinitely differentiable) AIC=576.89, BIC =890.44, DIC=376.89

The lower information criteria indicates a better model. The chosen exponential kernel points to only continuity and not differentiability of the random field thereby making it a better fit for RD analysis. As opposed to smooth synthetic patterns like the one considered for wombling presented with Materń($\nu = 5/2$) showing the least information criteria indicating differentiability of the random field.

Table 4: Project Details for the BIPs compared.

RUSID	Project ID	Name	Fisc.Yr.	Obligation Date	Original Obligation	Total Net Obligation (CY 2020)	Adv. Amt.	Resc. Amt.	Year5_Subsc.	Net Grant	Net Loan
VA 1108	VA1108-A39	LUMOS Telephone Inc.	2010	08/10/10	$8,062,088	$7,692,552	$7,692,552	$369,536	Yes	$7,692,552	$-
WV 1103	WV1103-A40	Hardy Telecommunications, Inc.	2010	08/06/10	$31,648,274	$31,622,782	$31,622,782	$25,492	Yes	$22,135,944	$9,486,838

CoreLogic Data: Spatial modeling, regression discontinuity design and wombling

BIP: VA1108

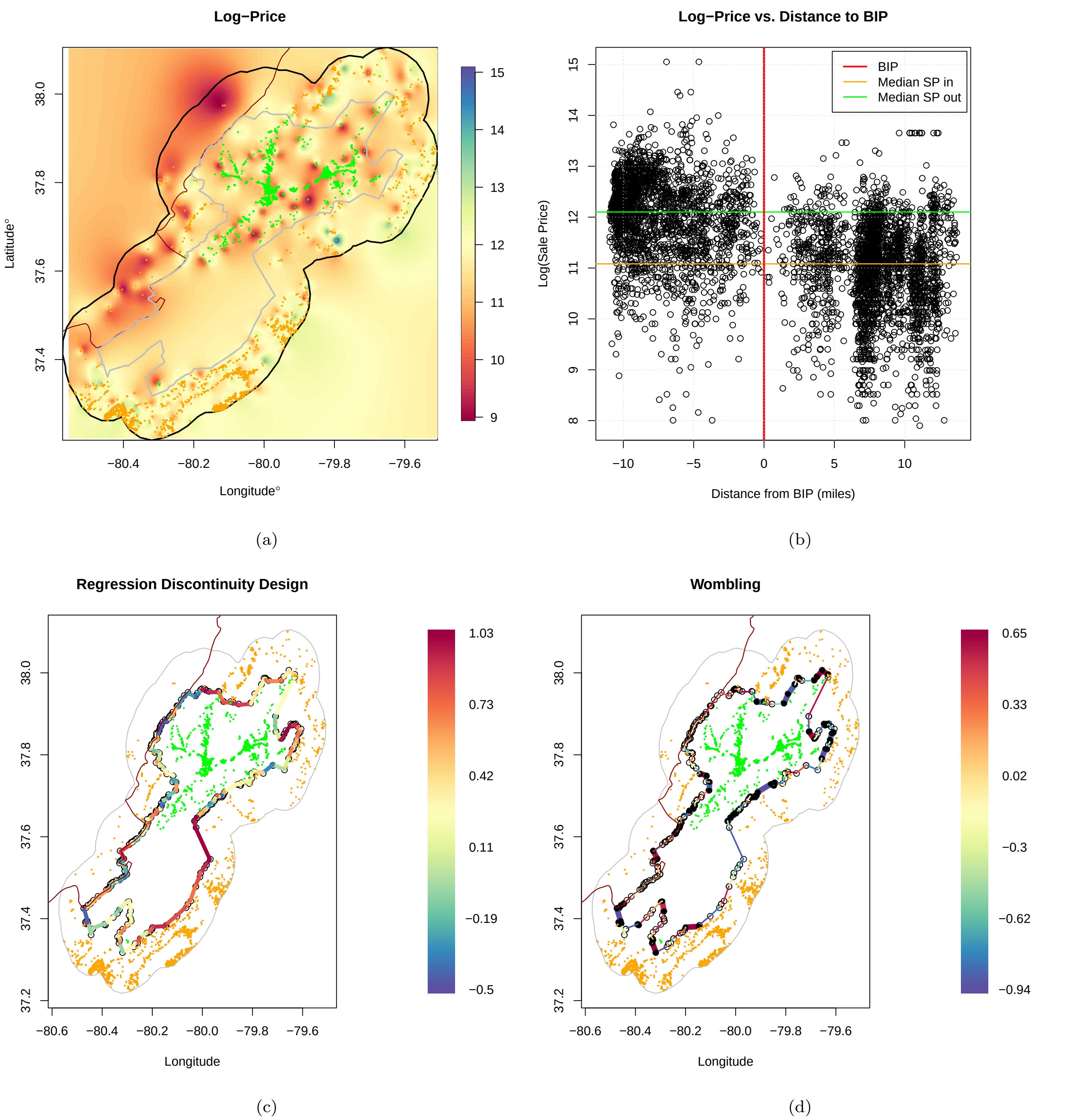


Figure 4: (a) Spatially interpolated surface of log-house prices in and around the BIP (15 mile radius) (b) log-house prices plotted against distance to BIP (c) RDD applied to fitted model, the difference of predictions of log-house prices, $\lim_{d\downarrow 0}\log(\text{Price})_{BIP}(\mathbf{s}+d) - -\lim_{d\uparrow 0}\log(\text{Price})_{BIP}(\mathbf{s}-d)$ (scale in log(US-Dollars))(b) Gradient analysis (Wombling) performed for the BIP (scale in log(US-Dollars)). *Note: Raw estimated gradients are first scaled by length of line segment. Direction (of the associated normal to the line segment) of gradients are inside $\rightarrow$ outside of BIP, i.e. positive gradients $\Rightarrow$ inside $>$ outside log-price and vice-versa.* Significant gradients are marked in bold. Gradient analysis took 22.17 mins.

BIP: WV1103

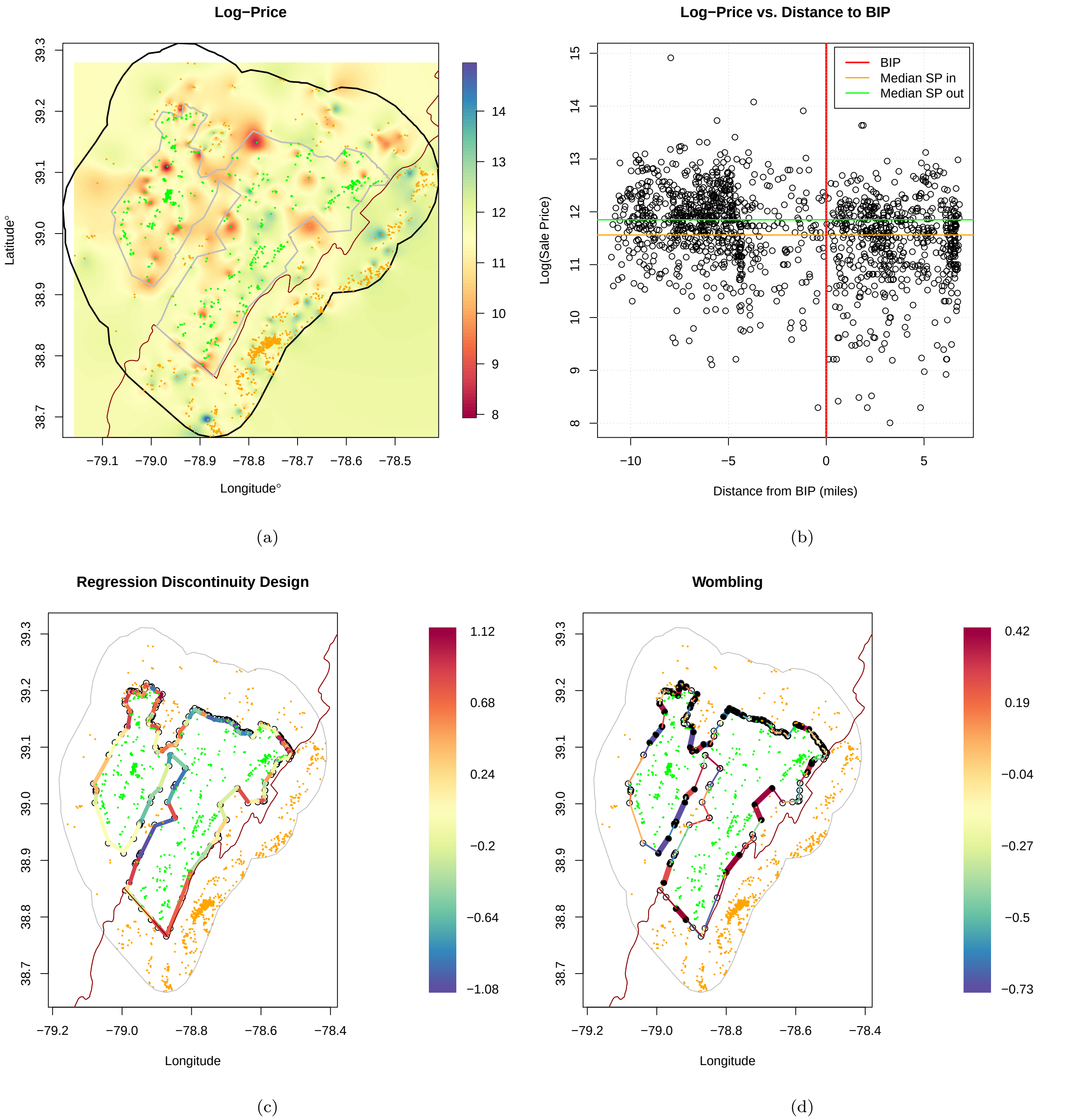


Figure 5: (a) Spatially interpolated surface of log-house prices in and around the BIP (15 mile radius) (b) log-house prices plotted against distance to BIP (c) RDD applied to fitted model, the difference of predictions of log-house prices, $\lim_{d\downarrow 0}\log(\text{Price})_{BIP}(\mathbf{s}+d) - -\lim_{d\uparrow 0}\log(\text{Price})_{BIP}(\mathbf{s}-d)$ (scale in log(US-Dollars))(b) Gradient analysis (Wombling) performed for the BIP (scale in log(US-Dollars)). *Note: Raw estimated gradients are first scaled by length of line segment. Direction (of the associated normal to the line segment) of gradients are inside → outside of BIP, i.e. positive gradients ⇒ inside > outside log-price and vice-versa.* Significant gradients are marked in bold. Gradient analysis took 4.18 mins.

Appendix: Tables

Table 5: Fitted model coefficients from the spatial model for BIP VA1108.

Effect	mean	sd	0.025quant	0.5quant	0.975quant	mode
(Intercept)	**12.6026**	1.4702	9.7164	12.6024	15.4871	12.6022
BIP	**0.4531**	0.0869	0.2822	0.4530	0.6246	0.4526
age	-0.0007	0.0004	-0.0015	-0.0007	0.0002	-0.0007
nbaths	**0.3394**	0.0161	0.3078	0.3394	0.3711	0.3394
ft.^{2}ratio	**0.2094**	0.0305	0.1496	0.2094	0.2692	0.2094
acres	**0.3103**	0.1378	0.0397	0.3103	0.5807	0.3102
land_square_footage	-0.2463	0.1369	-0.5151	-0.2463	0.0223	-0.2462
bedrooms	**0.0133**	0.0067	0.0002	0.0133	0.0265	0.0132
transaction_type:3	**0.2000**	0.0773	0.0482	0.2000	0.3515	0.2000
transaction_type:6	-0.0848	0.2176	-0.5120	-0.0848	0.3421	-0.0849
transaction_type:7	-0.0915	0.1507	-0.3874	-0.0915	0.2041	-0.0915
bldg_code:A0G	0.5533	0.3429	-0.1186	0.5529	1.2272	0.5520
bldg_code:C00	0.2213	0.3072	-0.3809	0.2209	0.8249	0.2202
bldg_code:MA0	-0.2001	0.6688	-1.5129	-0.2002	1.1122	-0.2003
bldg_code:R00	-0.0319	0.1133	-0.2514	-0.0330	0.1940	-0.0352
bldg_code:RM1	-0.4179	0.2616	-0.9302	-0.4184	0.0967	-0.4193
bldg_code:RM2	0.0010	0.1159	-0.2188	-0.0018	0.2368	-0.0075
bldg_code:RS0	**0.3202**	0.0975	0.1391	0.3163	0.5229	0.3082
pri_cat_code:B	-0.2271	0.1906	-0.6014	-0.2271	0.1468	-0.2271
pri_cat_code:D	0.1429	0.3896	-0.6220	0.1429	0.9071	0.1429
zoning:1	-0.0219	0.2897	-0.5911	-0.0218	0.5465	-0.0217
zoning:2	0.2276	0.3507	-0.4605	0.2274	0.9159	0.2271
zoning:3	**-0.5596**	0.2224	-0.9961	-0.5598	-0.1229	-0.5600
zoning:A-1	-0.1804	0.3691	-0.9084	-0.1793	0.5409	-0.1770
zoning:A1	**0.3126**	0.0778	0.1588	0.3129	0.4647	0.3134
zoning:A2	**0.4535**	0.1449	0.1660	0.4545	0.7352	0.4565
zoning:AG1	**0.3082**	0.1244	0.0641	0.3082	0.5526	0.3080
zoning:AG1S	1.2294	0.6904	-0.1262	1.2294	2.5839	1.2294
zoning:AG3	**0.3557**	0.1301	0.1026	0.3548	0.6134	0.3530
zoning:AL	0.5105	0.5055	-0.4822	0.5105	1.5021	0.5106
zoning:AR	0.0948	0.0770	-0.0567	0.0949	0.2459	0.0950
zoning:ARS	-0.1053	0.6763	-1.4331	-0.1053	1.2214	-0.1053
zoning:AV	0.2583	0.3290	-0.3876	0.2583	0.9038	0.2582
zoning:B-1	-0.0796	0.5244	-1.1096	-0.0794	0.9487	-0.0791
zoning:B-2	-0.0832	0.5186	-1.1022	-0.0829	0.9335	-0.0823
zoning:B1	-0.2883	0.2623	-0.8031	-0.2883	0.2264	-0.2884
zoning:C-1	-0.0077	0.5371	-1.0624	-0.0078	1.0463	-0.0078
zoning:C1	-0.3569	0.2547	-0.8569	-0.3569	0.1428	-0.3569
zoning:C2	0.0839	0.2037	-0.3160	0.0838	0.4834	0.0838
zoning:C3	1.1998	0.9429	-0.6516	1.1998	3.0493	1.1999
zoning:CFB2	0.2294	0.7063	-1.1568	0.2292	1.6155	0.2289
zoning:CFBD	0.2497	0.4362	-0.6067	0.2497	1.1054	0.2497
zoning:CFBG	-0.2820	0.2385	-0.7500	-0.2822	0.1865	-0.2825
zoning:CFR1	0.1042	0.1977	-0.2838	0.1040	0.4927	0.1037
zoning:CFR2	-0.2474	0.2017	-0.6430	-0.2476	0.1489	-0.2481
zoning:CFR3	-0.0056	0.2695	-0.5341	-0.0058	0.5236	-0.0062
zoning:CFUK	0.4314	0.3126	-0.1822	0.4314	1.0447	0.4314
zoning:CN	0.6482	0.6677	-0.6631	0.6482	1.9579	0.6484
zoning:FC	**0.7521**	0.2286	0.3045	0.7516	1.2017	0.7508
zoning:IG	-0.0948	0.2144	-0.5149	-0.0952	0.3271	-0.0959
zoning:M1	-0.9551	0.2606	-1.4667	-0.9552	-0.4439	-0.9552

	mean	sd	0.025quant	0.5quant	0.975quant	mode
zoning:PR	**0.6090**	0.2525	0.1140	0.6086	1.1059	0.6078
zoning:R-1	-0.1200	0.5460	-1.1931	-0.1197	0.9503	-0.1189
zoning:R-2	-0.4290	0.6148	-1.6366	-0.4289	0.7767	-0.4286
zoning:R-4	**0.6529**	0.1859	0.2886	0.6526	1.0186	0.6521
zoning:R1	**0.5014**	0.0771	0.3499	0.5015	0.6524	0.5016
zoning:R1S	1.1551	0.6688	-0.1581	1.1551	2.4672	1.1551
zoning:R2	**0.1651**	0.0827	0.0027	0.1651	0.3274	0.1650
zoning:R3	-0.0243	0.1034	-0.2272	-0.0244	0.1787	-0.0245
zoning:R4	0.2247	0.6738	-1.0982	0.2247	1.5467	0.2247
zoning:RR	-0.1352	0.2611	-0.6456	-0.1361	0.3794	-0.1378
zoning:RR-1	0.3831	0.3129	-0.2296	0.3824	0.9992	0.3809
zoning:RRA1	-0.2584	0.3559	-0.9581	-0.2582	0.4393	-0.2577
zoning:RSF	**1.2297**	0.1666	0.9024	1.2297	1.5565	1.2298
after-2010	**-0.0509**	0.0226	-0.0953	-0.0509	-0.0065	-0.0509

Table 6: Fitted model coefficients from the spatial model for BIP WV1103.

Effects	mean	sd	0.025quant	0.5quant	0.975quant	mode
(Intercept)	3.3679	6.2737	-8.9708	3.3742	15.6605	3.3872
bip	0.0781	0.1012	-0.1171	0.0768	0.2804	0.0742
age	0.0001	0.0007	-0.0012	0.0001	0.0015	0.0001
nbaths	0.2203	0.0299	0.1616	0.2203	0.2790	0.2203
sqft_ratio	0.1250	0.1506	-0.1706	0.1250	0.4204	0.1250
acres	-0.5024	0.5859	-1.6547	-0.5018	0.6455	-0.5005
land_square_footage	0.6418	0.5842	-0.5034	0.6411	1.7897	0.6398
bedrooms	0.0418	0.0266	-0.0103	0.0418	0.0939	0.0418
transaction_type:1	0.3591	0.6453	-0.9077	0.3589	1.6255	0.3586
transaction_type:3	0.4646	0.6736	-0.8578	0.4644	1.7867	0.4641
transaction_type:6	0.0298	0.6750	-1.2951	0.0296	1.3547	0.0292
transaction_type:7	0.1683	0.6625	-1.1321	0.1681	1.4687	0.1678
bldg_code:C00	0.7236	0.6691	-0.5899	0.7234	2.0366	0.7232
bldg_code:R00	**0.4091**	0.1419	0.1249	0.4106	0.6846	0.4137
bldg_code:RCA	0.4336	0.5386	-0.6249	0.4339	1.4896	0.4345
bldg_code:RM0	0.2612	0.5095	-0.7397	0.2613	1.2603	0.2617
bldg_code:RM1	0.2101	0.4957	-0.7633	0.2101	1.1827	0.2102
bldg_code:RM2	-0.0583	0.3330	-0.7130	-0.0581	0.5945	-0.0576
bldg_code:RS0	0.2423	22.3610	-43.6599	0.2417	44.1079	0.2423
bldg_code:RT0	0.2397	0.2188	-0.1901	0.2396	0.6693	0.2395
bldg_code:X0M	-0.1946	0.3062	-0.7968	-0.1944	0.4054	-0.1938
pri_cat_code:B	**-0.8413**	0.3812	-1.5898	-0.8413	-0.0933	-0.8414
pri_cat_code:C	-0.8568	0.6461	-2.1251	-0.8570	0.4113	-0.8573
pri_cat_code:D	0.1433	0.3370	-0.5183	0.1433	0.8046	0.1433
zoning:0000	-0.0418	0.2412	-0.5155	-0.0417	0.4313	-0.0416
zoning:A2	0.0490	0.1942	-0.3267	0.0466	0.4385	0.0418
zoning:R2	-0.0646	0.4547	-0.9565	-0.0650	0.8286	-0.0658
zoning:RA	0.2423	22.3610	-43.6599	0.2417	44.1079	0.2423
zoning:RR1	-0.0606	0.2436	-0.5377	-0.0613	0.4202	-0.0627
after-2010	-0.0630	0.0443	-0.1500	-0.0630	0.0240	-0.0629

References

Finley, A., Banerjee, S., Hjelle, Ø., Bivand, R., 2010. MBA: Multilevel B-spline Approximation. R package version 0.0-7. R Foundation for Statistical Computing: Vienna, Austria .

Finley, A.O., Banerjee, S., Carlin, B.P., 2007. spbayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. Journal of statistical software 19, 1.

Hidano, N., Hoshino, T., Sugiura, A., 2015. The effect of seismic hazard risk information on property prices: Evidence from a spatial regression discontinuity design. Regional Science and Urban Economics 53, 113–122.

Keele, L.J., Titiunik, R., 2015. Geographic boundaries as regression discontinuities. Political Analysis 23, 127–155.

Kelejian, H.H., Prucha, I.R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. The Journal of Real Estate Finance and Economics 17, 99–121.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. Journal of economic literature 48, 281–355.

Womble, W.H., 1951. Differential systematics. Science 114, 315–322.