Citation
Baddeley, A. and Davies, T.M. and Rakshit, S. and Nair, G. and McSwiggan, G. 2022. Diffusion Smoothing for Spatial Point Patterns. Statistical Science. 37 (1): pp. 123-142.
http://doi.org/10.1214/21-STS825

SUPPLEMENTARY MATERIAL
for the article
Diffusion Smoothing for Spatial Point Patterns

A. Baddeley T.M. Davies S. Rakshit G.M. Nair G. McSwiggan

Contents

1 New Zealand pā: Additional plots

To emphasise the distinction between the Gaussian and diffusion estimates of intensity for the pā data, we consider a one-dimensional transect across the entrance to the Kaipara Harbour. The transect location is shown in Figure SUP-1 plotted as a line superimposed on the Gaussian and diffusion estimates from Figure 6 of the main article.

Figure SUP-2 plots the values of the two intensity estimates along this transect. It clearly highlights the difference in behaviour between them. The Gaussian estimate resembles a smooth curve interrupted by a gap in the middle of the graph. The diffusion estimate shows markedly different values on either side of the gap. The diffusion estimate is much more plausible for this dataset: Figure 5 of the main article suggests that there are many more pā locations on the southern side than on the northern side of the harbour entrance.

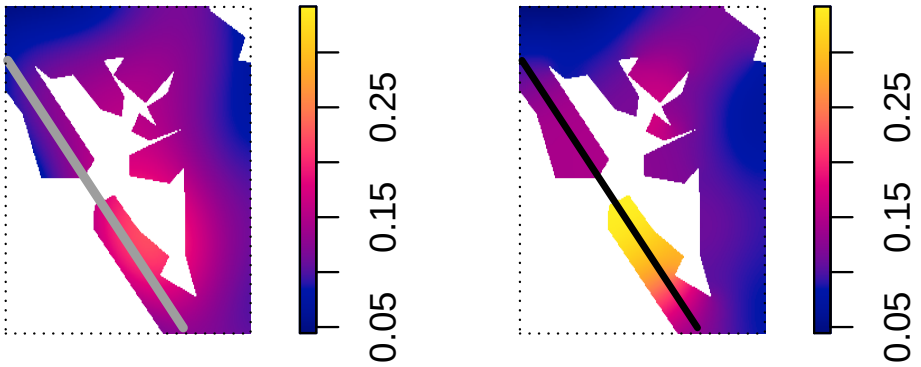

Figure SUP-1: Estimates of the inhomogeneous spatial intensity of pā sites using the fixed-bandwidth Gaussian kernel estimator (left) and the fixed-bandwidth diffusion estimator (right). Superimposed upon both images is a coastal transect running from North-East to South-West along the outer heads of Kaipara Harbour.

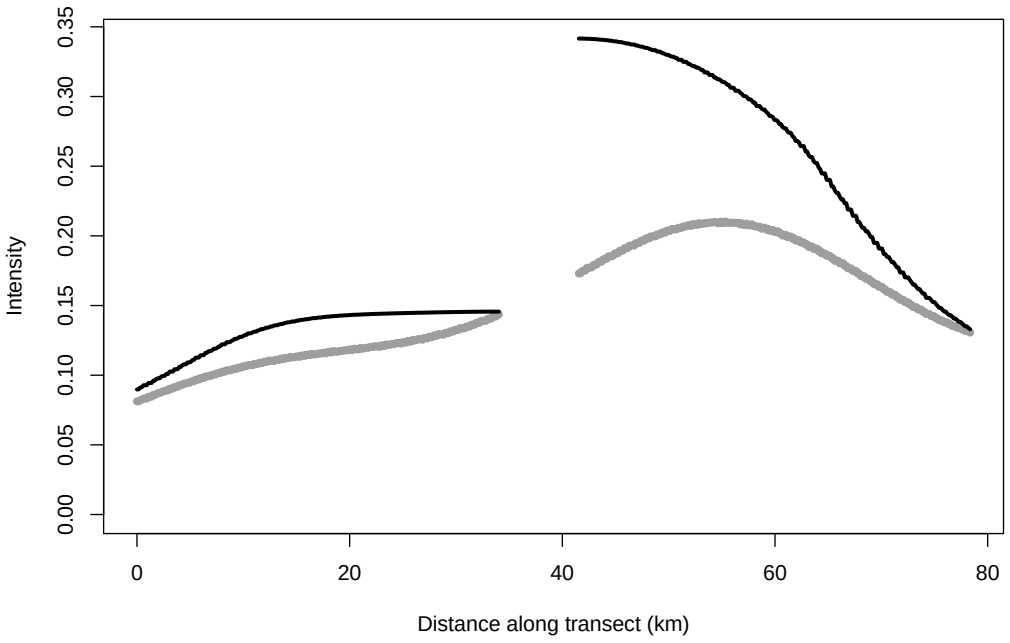

Figure SUP-2: Estimates of the inhomogeneous spatial intensity of pā sites along a coastal transect running from North-East to South-West along the outer heads of Kaipara Harbour. *Thick grey line:* fixed-bandwidth Gaussian kernel estimator. *Thin black line:* fixed-bandwidth diffusion smoother.

2 United Kingdom PBC: Additional commentary

In the analysis of the PBC data presented in Section 5 of the main article, we noted that the VR adaptive estimate closely follows the appearance of the pilot density. This tendency is clear in Figure SUP-3 below, which shows the clipped pilot density, the bandwidth surface, and the VR estimate side-by-side. The final VR estimate exhibits much the same features as the pilot.

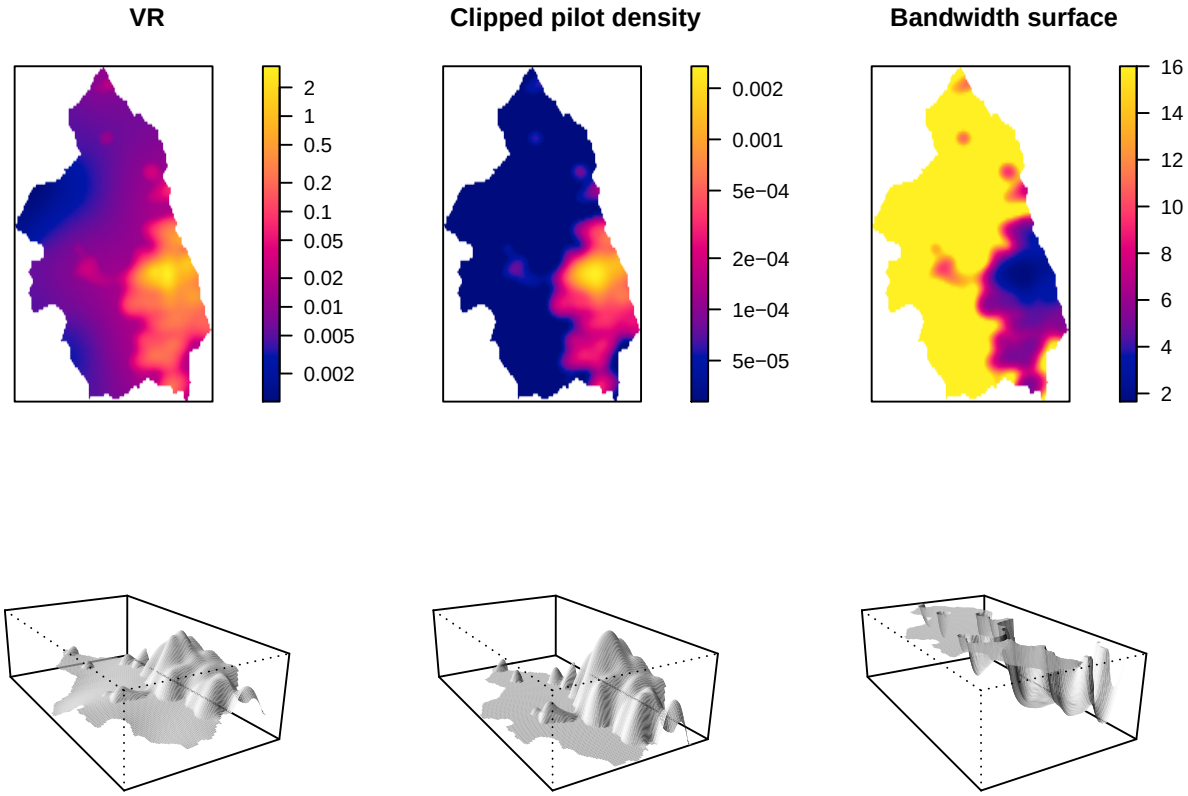

Figure SUP-3: Image (top row) and perspective (bottom row) plots of the log-scaled VR diffusion intensity estimate of PBC cases (left); the log-scaled fixed-bandwidth clipped pilot density estimate (middle); and the bandwidth surface (right).

While this behaviour is expected when the global bandwidth is large, according to the underlying theory, it is more surprising that the phenomenon also occurs at the automatically-selected global bandwidth for this dataset. This suggests that variable-rate diffusion could be problematic if our pilot estimate is biased, incorrect, or in any other way a poor reflection of the truth.

3 Simulation study scenario definitions

Here we provide further details on the simulation study design. In all scenarios we define the study region W (with area $|W|$) as the 'zoomed-in' Kaipara Harbour region of the pā data analysis (see the rightmost panel of Figure 5 in Section 5 of the main article). All scenarios are based on intensity functions that integrate to 500 over W.

3.1 Scenario S1

The intensity for scenario S1, $\lambda_{\text{S1}}(x)$, is a Gaussian mixture with 20 contributing components and a small uniform constant:

$$\lambda_{\text{S1}}(x) = 500 \times \left\{ \frac{0.1}{|W|} + p \sum_{k=1}^{20} \phi\left(x; \mu_k, \Sigma_k\right) \right\}; \qquad x \in W,$$

where $p = 9/200$; $\phi(\cdot; \mu, \Sigma)$ denotes the 2-dimensional normal density function with mean μ and covariance matrix Σ;

$$[\mu_1, \ldots, \mu_{20}]^\top = \begin{bmatrix} 241.9305 & 614.7764 \\ 260.3712 & 592.1339 \\ 280.4925 & 550.0444 \\ 279.9034 & 570.8891 \\ 282.6805 & 541.0712 \\ 263.7468 & 609.5753 \\ 238.3585 & 578.5664 \\ 236.5934 & 587.9653 \\ 267.2214 & 579.4833 \\ 272.1905 & 606.1899 \\ 269.0571 & 603.5392 \\ 285.5144 & 548.6355 \\ 248.8021 & 597.8969 \\ 272.6467 & 572.9020 \\ 275.2568 & 613.0301 \\ 272.9760 & 576.7253 \\ 273.3614 & 560.6413 \\ 254.6338 & 601.3049 \\ 269.9388 & 546.7398 \\ 267.5639 & 610.0257 \end{bmatrix}; \quad \text{and} \quad [\Sigma_1, \ldots, \Sigma_{20}]^\top = \begin{bmatrix} 12.0558490 & -9.5947700 \\ -9.5947700 & 37.1317627 \\ 116.2502686 & -216.2645770 \\ -216.2645770 & 467.1328566 \\ 9.6596418 & 2.9591341 \\ 2.9591341 & 17.3091211 \\ 8.0059818 & -19.0354331 \\ -19.0354331 & 61.2516597 \\ 12.0013587 & 17.7538105 \\ 17.7538105 & 33.3420977 \\ 35.5599334 & 0.1610234 \\ 0.1610234 & 4.5879939 \\ 6.4367478 & 5.4276568 \\ 5.4276568 & 21.6127810 \\ 30.4884863 & -233.3580245 \\ -233.3580245 & 2336.6517745 \\ 8.1630976 & -2.3664032 \\ -2.3664032 & 21.8025636 \\ 48.3347170 & -50.7429596 \\ -50.7429596 & 78.2370783 \\ 7.4581361 & -0.7628775 \\ -0.7628775 & 25.0896456 \\ 16.0805413 & -3.6962223 \\ -3.6962223 & 8.3944104 \\ 10.3860409 & -2.0920954 \\ -2.0920954 & 13.2106916 \\ 5.2310067 & -6.0356562 \\ -6.0356562 & 30.9093578 \\ 11.1050409 & -22.9876265 \\ -22.9876265 & 70.0167070 \\ 5.2843805 & -1.2994042 \\ -1.2994042 & 6.9091569 \\ 50.8747432 & -16.1224983 \\ -16.1224983 & 17.8903858 \\ 10.0996101 & 17.7627155 \\ 17.7627155 & 59.4067690 \\ 20.6817283 & 25.5827807 \\ 25.5827807 & 49.1060179 \\ 10.5920432 & -17.3139131 \\ -17.3139131 & 39.5428916 \end{bmatrix}.$$

3.2 Scenario S2

Let $b_W(x)$ denote the shortest Euclidean distance from the coordinate x to the boundary of the window W. Furthermore, let $m_W = \min\{-b_W(x)\}$ with respect to all locations $x \in W$. Then the intensity $\lambda_{\text{S2}}(x)$ corresponding to scenario S2 is given by

$$\lambda_{\text{S2}}(x) = 500 \frac{\nu(x)}{\int_W \nu(y)\, \mathrm{d}y}; \qquad x \in W,$$

where

$$\nu(x) = \left(-b_W(x) - m_W\right)^2.$$

3.3 Scenario S3

The intensity $\lambda_{\text{S3}}(x)$ representing Scenario S3 is a rescaled, fixed-bandwidth, Gaussian kernel intensity estimate (with uniform edge correction—equation (2) in the main article) of the $n = 369$ observed pā sites $X = \{x_1, \ldots, x_{369}\}$ that fall in the 'zoomed-in' region W, denoted here by $\widehat{\lambda}_\sigma^{(U)}(x|X)$. The bandwidth is set to $\sigma = 7$ km such that

$$\lambda_{\text{S3}}(x) = 500 n^{-1} \widehat{\lambda}_7^{(U)}(x|X); \qquad x \in W.$$

3.4 Scenario S4

Scenario S4 very similar to S3, though now the intensity λ_{S4} is based on a fixed-bandwidth *diffusion* estimate of the same 369 pā observations in X (see equation (2) in the main article), denoted here by $\widehat{\lambda}_t(x|X)$. The diffusion 'time' t is set to provide an equivalent level of smoothing with that of S3, namely $t = \sigma^2 = 49$, such that

$$\lambda_{\text{S4}}(x) = 500 n^{-1} \widehat{\lambda}_{49}(x|X); \qquad x \in W.$$

3.5 Scenario S5

Finally, the intensity λ_{S5} is obtained by rescaling a single realisation of a stationary, isotropic log-Gaussian Cox process (see Møller et al. 1998) on the region W. We set

$$z(x) \leftarrow \exp\left\{Y(x|\mu, \alpha^2, \beta)\right\}; \qquad x \in W,$$

where '$\leftarrow$' should be read as 'generated realisation of' and $Y(x|\mu, \alpha^2, \beta)$ is taken to be a stationary, isotropic Gaussian random field defined by mean μ. The covariance function of Y, $\rho(\cdot)$, is based on an exponential correlation structure given by

$$\rho(r) = \alpha^2 \exp(-r/\beta); \qquad r \geq 0,$$

where r represents Euclidean distance and α^2 and β are the variance and scale parameters respectively. In our implementation we use $\mu = 1$; $\alpha^2 = 2$; and $\beta = 5$ when generating $z(x)$, and simply find

$$\lambda_{\text{S5}}(x) = 500 \frac{z(x)}{\int_W z(y)\,\mathrm{d}y}.$$

Once established, the realised intensity λ_{S5} is held fixed for the duration of the simulations i.e. the function $z(x)$ is *not* generated anew at each iteration.

4 Accuracy of discrete approximation

4.1 Total variation

Tables SUP-1 and SUP-2 below show the counterparts of Tables 1 and 2 from Section 4.4 of the main article, when discretisation is measured using the *total variation* distance. The total variation between two probability distributions P and Q is the maximum value of $|P(A) - Q(A)|$ for any event A. For probability densities, the total variation is equal to one-half the integrated absolute difference between densities.

The conclusion from these tables corresponds to the conclusion for Tables 1 and 2 of the main article: the total variation error in the Euler scheme converges to zero at rate $O(\Delta x)$, while Richardson extrapolation accelerates the convergence to $O((\Delta x)^2)$.

Table SUP-1: Maximum **total variation** error of discrete approximation to heat kernel. Window is the unit square. Single source point at $(0.5, 0.5)$. Bandwidth $\sigma = 0.1$.

	Grid size				
	32	64	128	256	512
4-connected	0.088	0.044	0.022	0.011	0.006
8-connected	0.088	0.044	0.022	0.011	0.006

Table SUP-2: Maximum **total variation** error of discrete approximation to heat kernel **using Richardson extrapolation**. Window is the unit square. Single source point at $(0.5, 0.5)$. Bandwidth $\sigma = 0.1$.

	Finest grid size				
	32	64	128	256	512
4-connected	0.0388	0.0129	0.0033	0.0008	0.0002
8-connected	0.0419	0.0110	0.0028	0.0007	0.0002

4.2 Supremum distance between bivariate CDF's

Additionally we have calculated the supremum distance (maximum absolute difference) between the bivariate cumulative distribution functions. Tables SUP-3 and SUP-4 below provide the results. An upper estimate for this quantity was proposed in Appendix D.

Table SUP-3: Maximum error (maximum absolute discrepancy in bivariate CDF) of discrete approximation to heat kernel. Window is the unit square. Single source point at $(0.5, 0.5)$. Bandwidth $\sigma = 0.1$.

	Grid size				
	32	64	128	256	512
4-connected	0.077	0.038	0.019	0.010	0.005
8-connected	0.075	0.038	0.019	0.009	0.005

Table SUP-4: Maximum error (maximum absolute discrepancy in bivariate CDF) of discrete approximation to heat kernel **using Richardson extrapolation**. Window is the unit square. Single source point at $(0.5, 0.5)$. Bandwidth $\sigma = 0.1$.

	Finest grid size				
	32	64	128	256	512
4-connected	0.0158	0.0063	0.0016	0.0004	0.0001
8-connected	0.0181	0.0054	0.0014	0.0004	0.0001

References

Møller, J., Syversveen, A.-R. & Waagepetersen, R. (1998), 'Log Gaussian Cox processes', *Scandinavian Journal of Statistics* **25**, 451–482.