Statistics 222, Spatial Statistics.

Outline for the day:

1. irregular boundaries, fithawkes.r .
2. Exponential density in the plane.
3. Deviance residuals and Voronoi residuals.

Modifying F,G,J,K,L functions to deal with irregular boundaries is in the file custom_obs_window_jkl thanks to Michael Tzen. It is on the course site.

2. Exponential density in the plane.

I originally was going to have this.

Fitting a Pseudo-Likelihood model.
I'm using the model lambda_p (z | z_1, ..., z_k) =
mu + alpha x + beta y + gamma SUM_{i = 1 to k} a1 exp{-a1 D(z_i,z)}
where z = (x,y), and where D means distance.
So, if gamma is positive, then there is clustering; otherwise inhibition.

But $g(r) = a_1 \exp(-a_1 r)$ is actually not a density.
$g(t) = a_1 \exp(-a_1 t)$ is a density, because $\int_0^\infty a_1 \exp(-a_1 t)\, dt = 1$, for $a_1 > 0$,
but not $\int\int a_1 \exp(-a_1 r)\, dx\, dy$.

$a_1 \exp(-a_1 r) / (2\pi r)$ is a spatial density, because
$\iint a_1 \exp(-a_1 r) / (2\pi r)\, dx\, dy = \int_0^{2\pi} \int_0^\infty a_1 \exp(-a_1 r) / (2\pi r)\, r\, dr\, d\phi$
$= \int_0^\infty a_1 \exp(-a_1 r)\, dr$
$= 1.$

So I should fit **lambda_p (z | z_1, ..., z_k) =**
mu + alpha x + beta y + gamma SUM_{i = 1 to k} a1/2π exp{-a1 D(z_i,z)}/D(z_i,z).
This is in day07.r.

Recent methods for point process models for occurrences.

1. Deviance residuals

2. Voronoi residuals

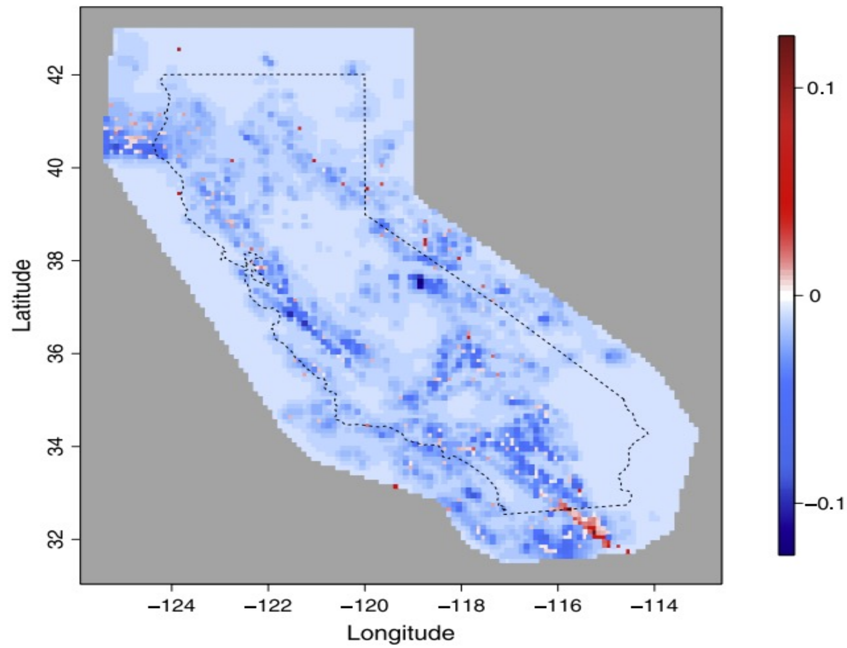
3. Superthinned residuals.

-- Given two competing models, can consider the difference between residuals, number of observed points – number expected, over each pixel. Divide by the estimated SE to get *Pearson residuals* (Baddeley et al. 2005).

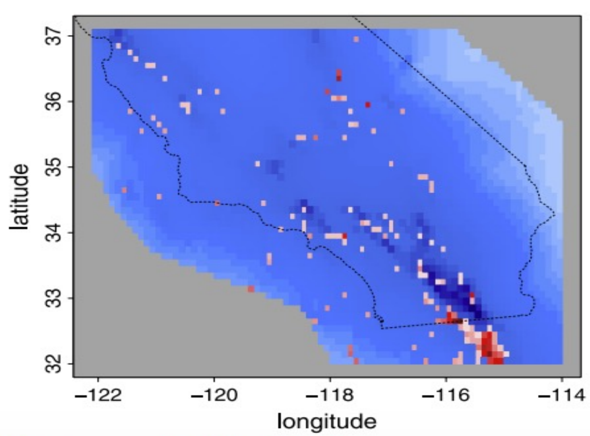
Problem: Hard to interpret. If difference = 3, is this because model A overestimated by 3? Or because model B underestimated by 3? Or because model A overestimated by 1 and model B underestimated by 2?

-- Also, the results are rarely visually appealing or useful.

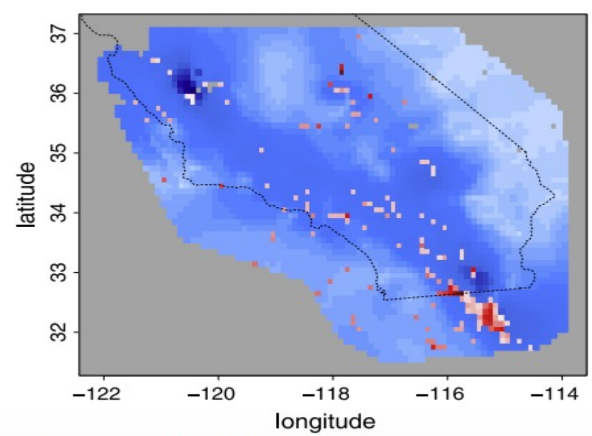
Pearson residuals tend to look just like a map of the points, unless pixels are very large.



(a) Pearson residuals for Model A.



(b) Pearson residuals for Model B.



(c) Pearson residuals for Model C.

With two competing models, it is better to consider the difference between *log-likelihoods*, in each pixel. The result may be called ***deviance residuals*** (Clements et al. 2011), ~ resids from gen. linear models.

$$R_{\mathrm{D}}(B_i) = \sum_{i:(t_i,x_i,y_i)\in B_i} \log(\hat{\lambda}_1(t_i,x_i,y_i)) - \int_{B_i} \hat{\lambda}_1(t,x,y)\,\mathrm{d}t\mathrm{d}x\mathrm{d}y - \left(\sum_{i:(t_i,x_i,y_i)\in B_i} \log(\hat{\lambda}_2(t_i,x_i,y_i)) - \int_{B_i} \hat{\lambda}_2(t,x,y)\,\mathrm{d}t\mathrm{d}x\mathrm{d}y \right).$$

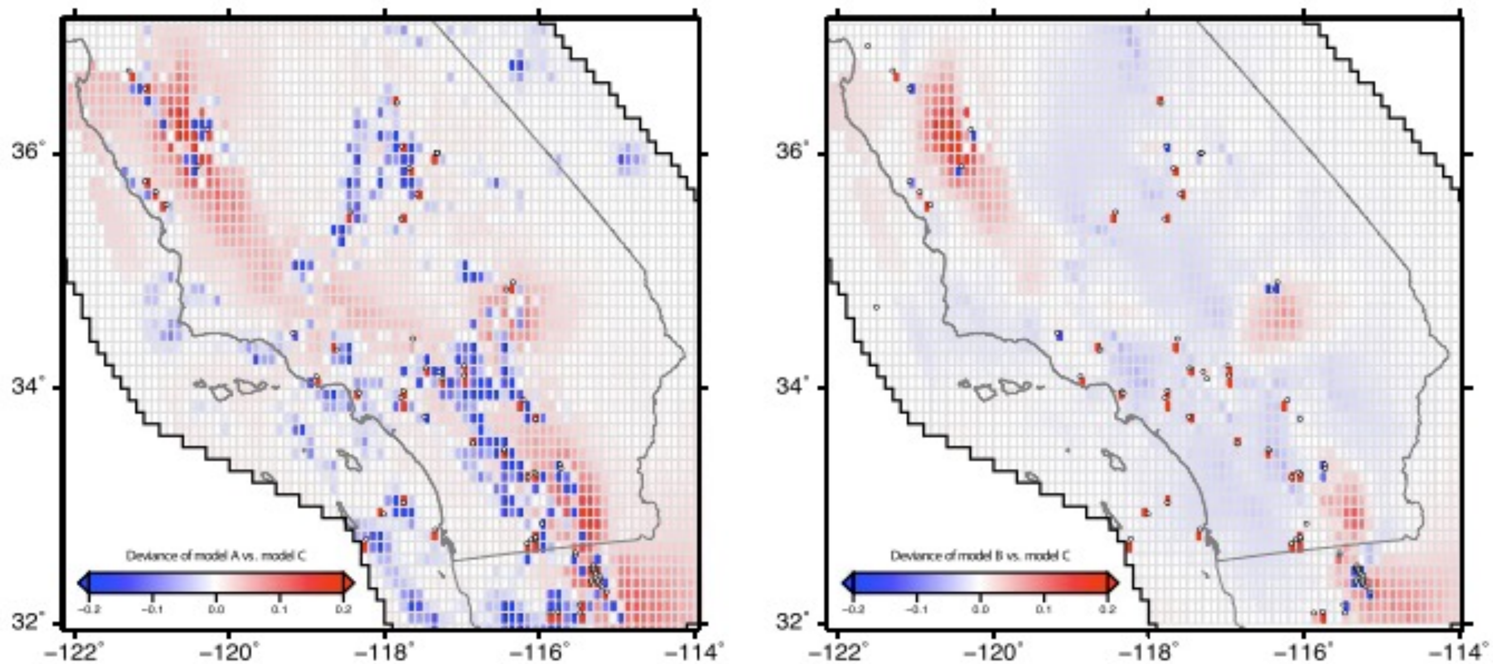


FIG. 4. *Left panel* (a): *deviance residuals for model A versus C. Sum of deviance residuals is* 86.427. *Right panel* (b): *deviance residuals for model B versus C. Sum of deviance residuals is* −7.468.

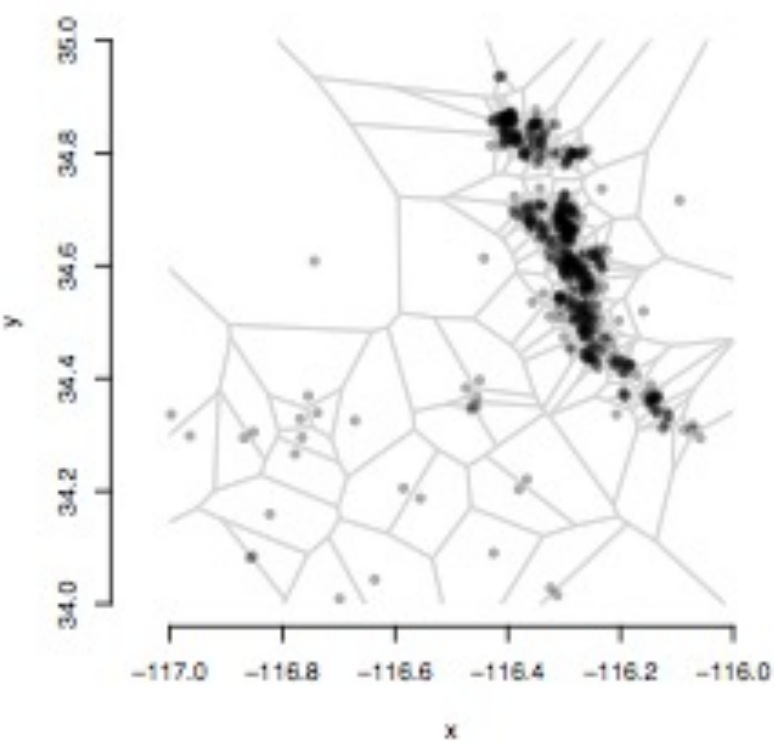
Voronoi residuals (Bray et al. 2013)

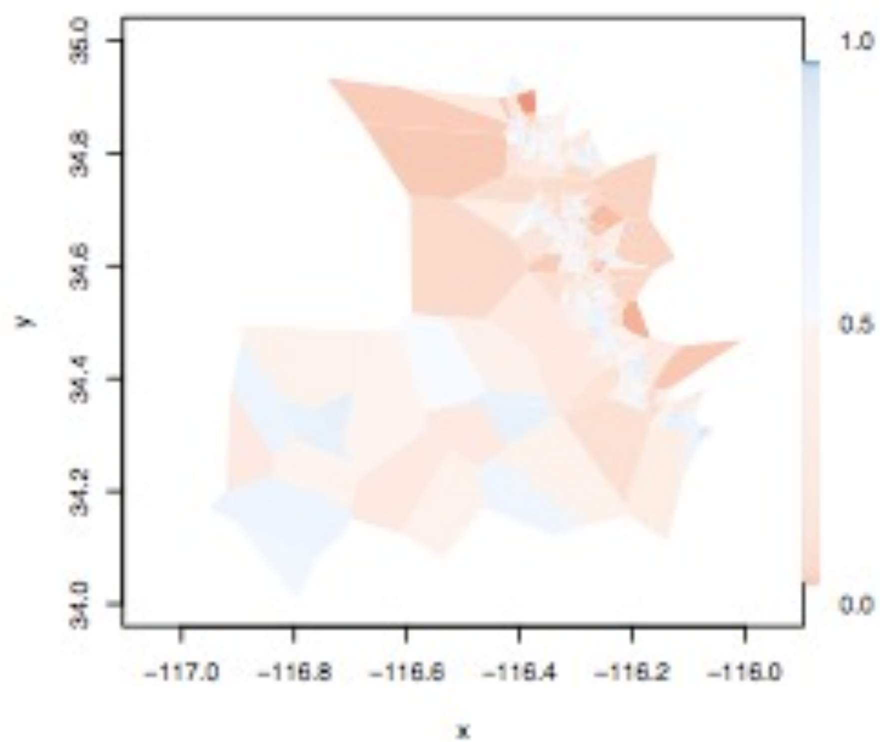
A Voronoi tessellation divides a space into cells C_i, where C_i contains all locations closer to event i than any other observed event.

Within each cell, calculate residuals

$$\hat{r}_i := 1 - \int_{C_i} \hat{\lambda}\, \mathrm{d}\mu = 1 - |C_i|\bar{\lambda},$$

$r \sim 1 - X$; $X \sim \Gamma(3.569, 3.569)$ (Tanemura 2003)





spatially adaptive and nonparametric.

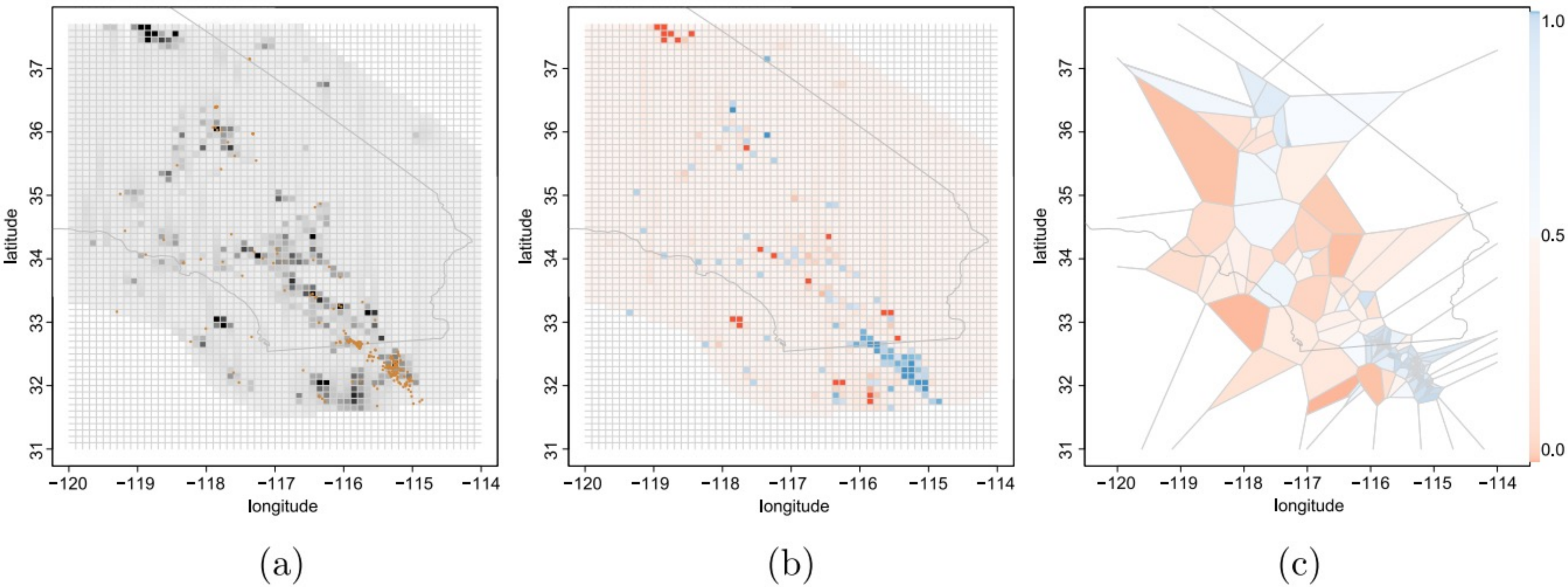
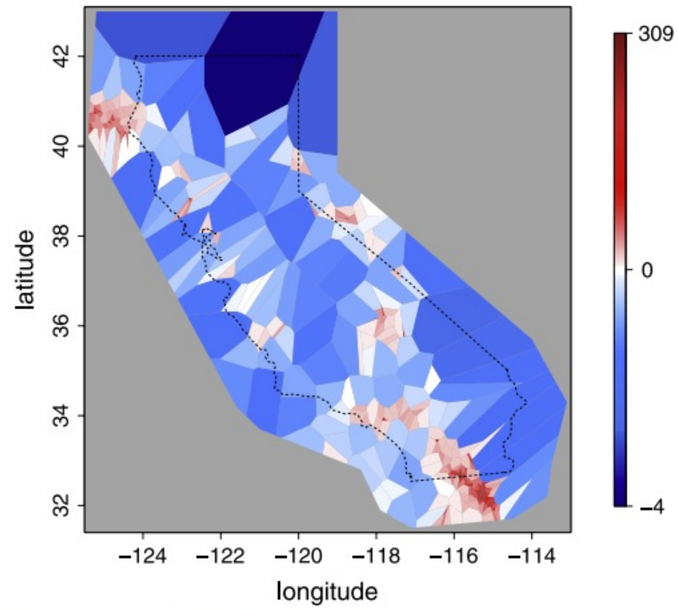
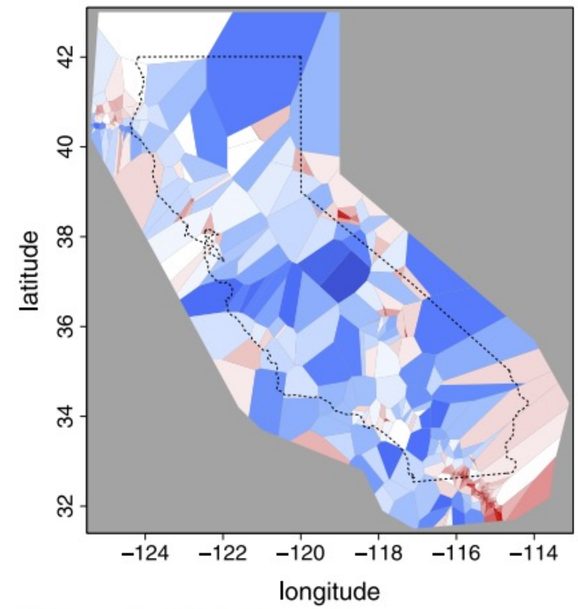


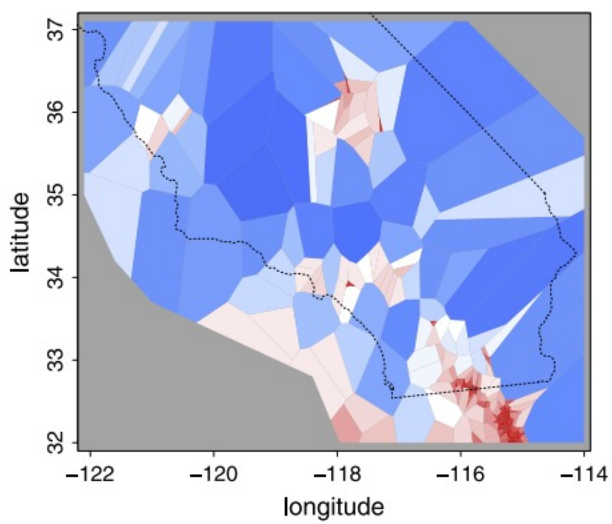
FIG. 2. (a) *Estimated rates under the Helmstetter, Kagan and Jackson* (2007) *model, with epicentral locations of observed earthquakes with $M \geq 4.0$ in Southern California between January* 1, 2006 *and January* 1, 2011 *overlaid.* (b) *Raw pixel residuals for Helmstetter, Kagan and Jackson* (2007) *with pixels colored according to their corresponding p-values.* (c) *Voronol residuals for Helmstetter, Kagan and Jackson* (2007) *with pixels colored according to their corresponding p-values.*



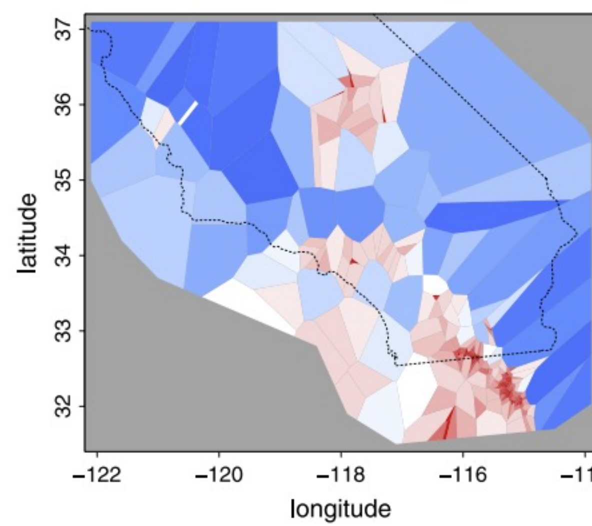
(a) Voronoi residuals for null model.



(b) Voronoi residuals for Model A.



(c) Voronoi residuals for Model B.



(d) Voronoi residuals for Model C.

With 2 models, can compare loglikelihoods across pixels or Voronoi cells.

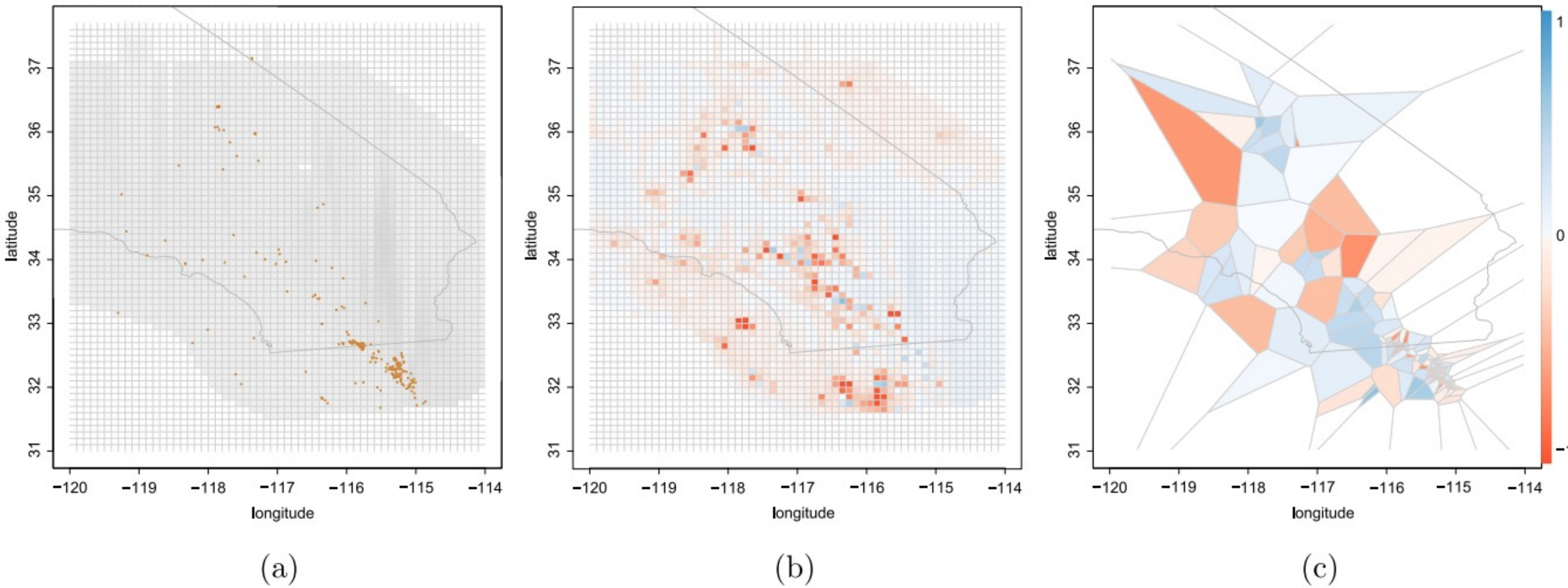


FIG. 3. (a) *Estimated rates under the Shen, Jackson and Kagan* (2007) *model, with epicentral locations of observed earthquakes with* $M \geq 4.0$ *in Southern California between January* 1, 2006 *and January* 1, 2011 *overlaid.* (b) *Pixel deviance plot with blue favoring model A, Helmstetter, Kagan and Jackson* (2007), *versus model B, Shen, Jackson and Kagan* (2007). *Coloration is on a linear scale.* (c) *Voronoi deviance plot with blue favoring model A, Helmstetter, Kagan and Jackson* (2007), *versus model B, Shen, Jackson and Kagan* (2007). *Coloration is on a linear scale.*