Statistics 222, Spatial Statistics.

Outline for the day:

1. Deviance residuals.
2. Voronoi deviance residuals.
3. Superthinning.
4. Exercises.
5. Review list.

Recent methods for point process models for occurrences.

1. Deviance residuals

2. Voronoi residuals

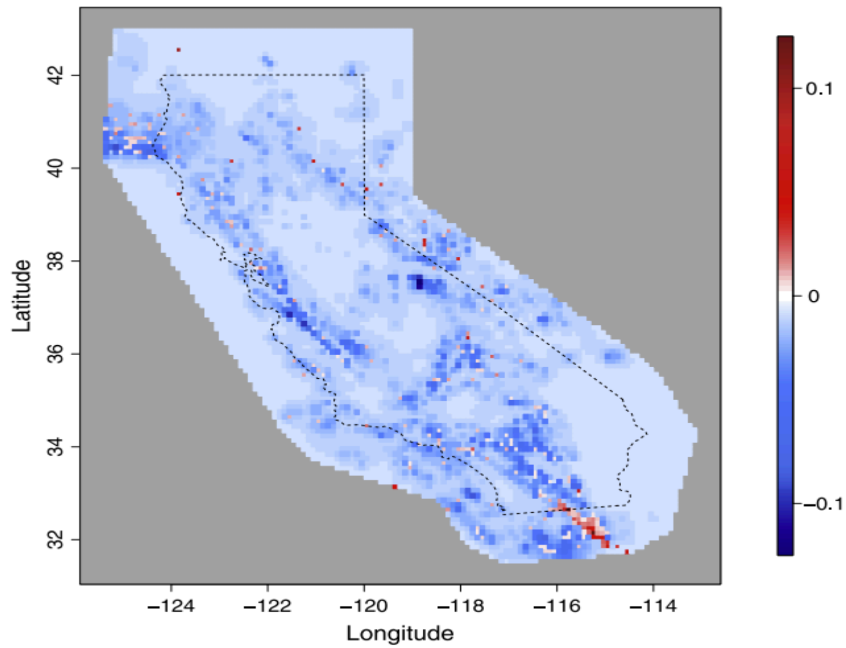
3. Superthinned residuals.

-- 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).

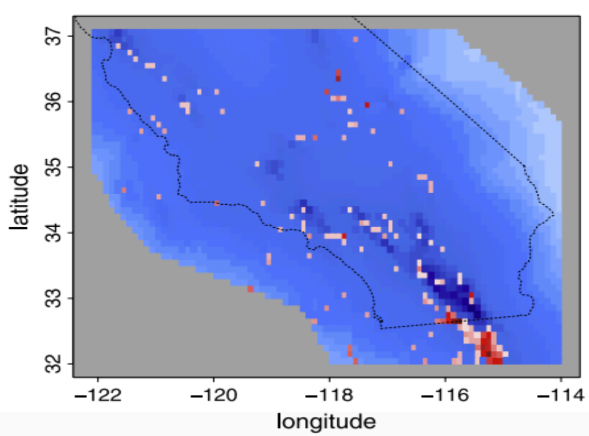
Problem: Hard to interpret. If difference = 3, is this because model A overestimated by 3? Or because model B underestimated by 3? Or because model A overestimated by 1 and model B underestimated by 2?

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

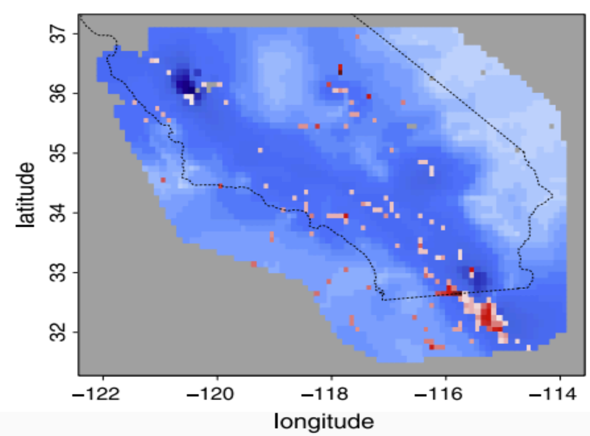
Pearson residuals tend to look just like a map of the points, unless pixels are very large.



(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.

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

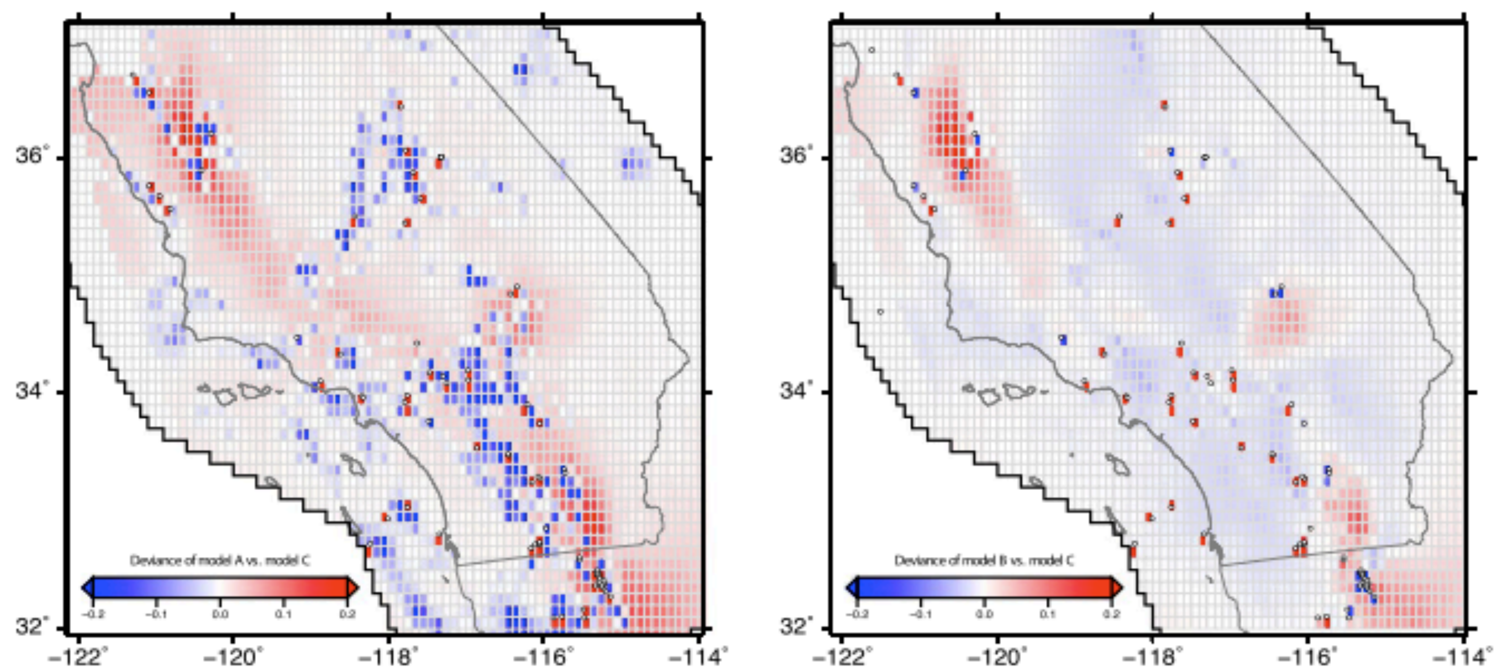


FIG. 4. *Left panel* (a): *deviance residuals for model A versus C. Sum of deviance residuals is* 86.427. *Right panel* (b): *deviance residuals for model B versus C. Sum of deviance residuals is* −7.468.

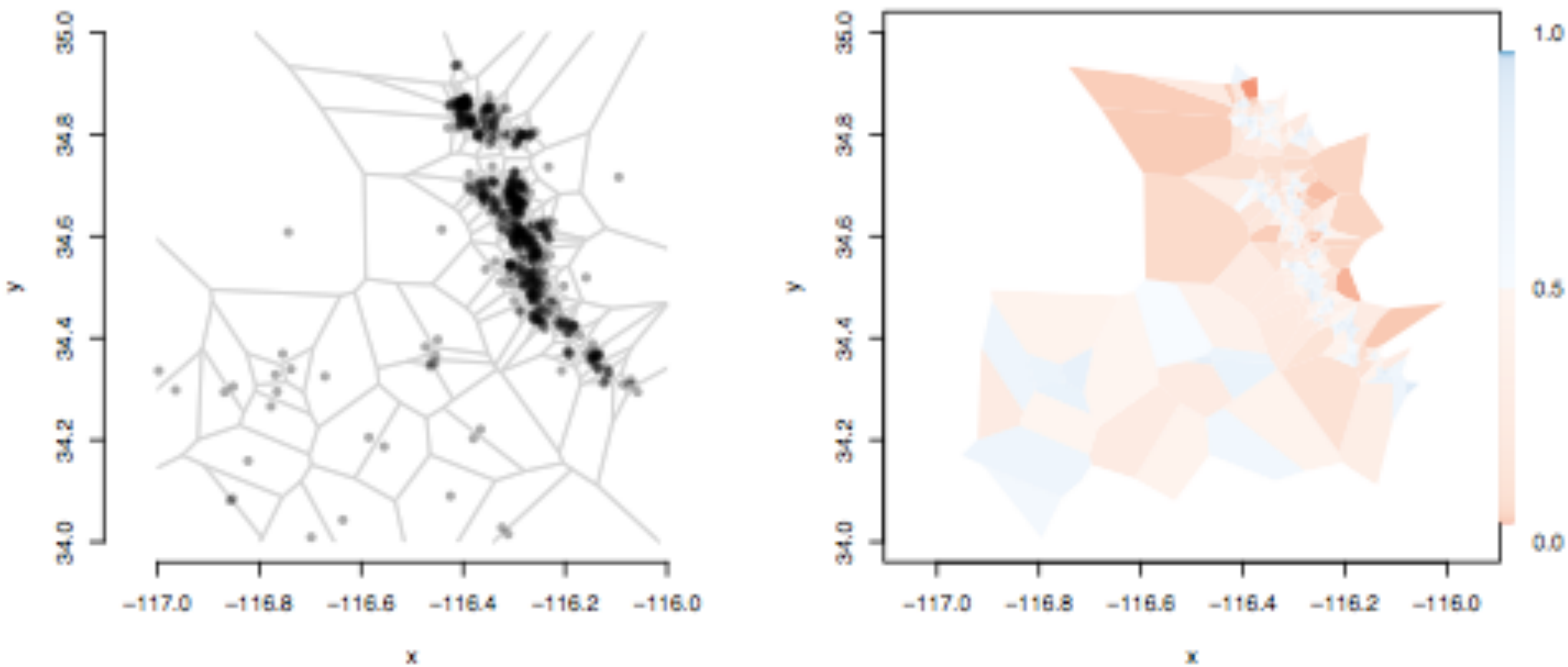
2. Voronoi residuals (Bray et al. 2013)

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

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

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



spatially adaptive and nonparametric.

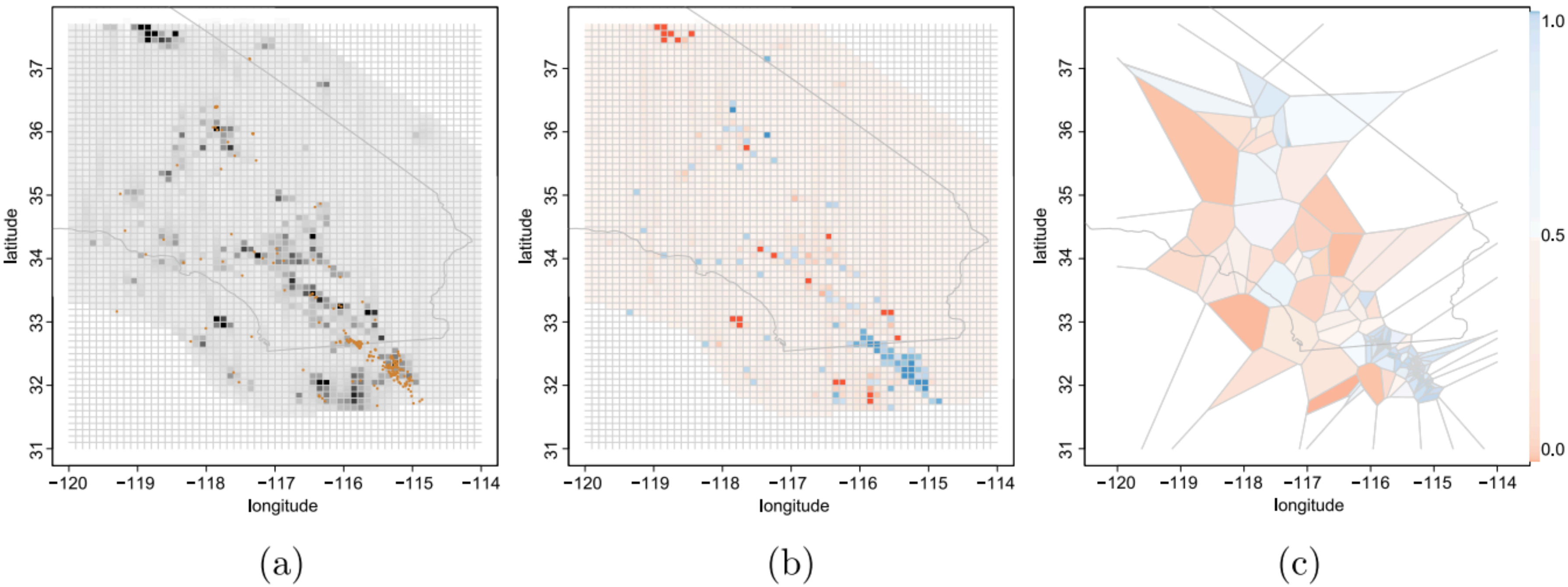
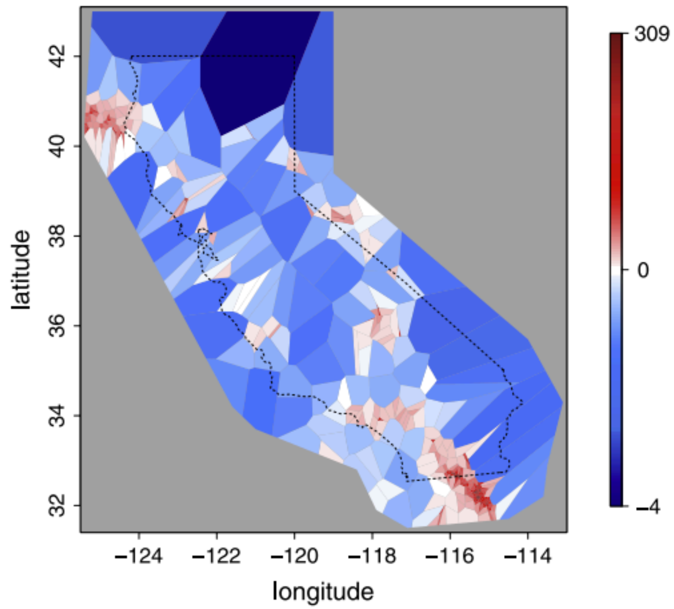
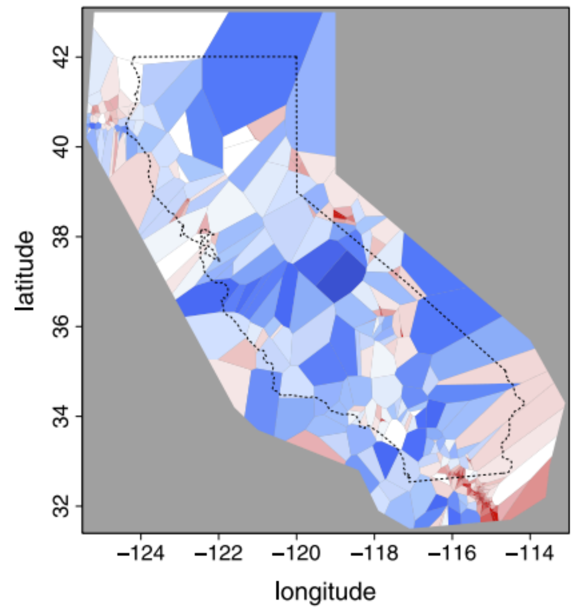


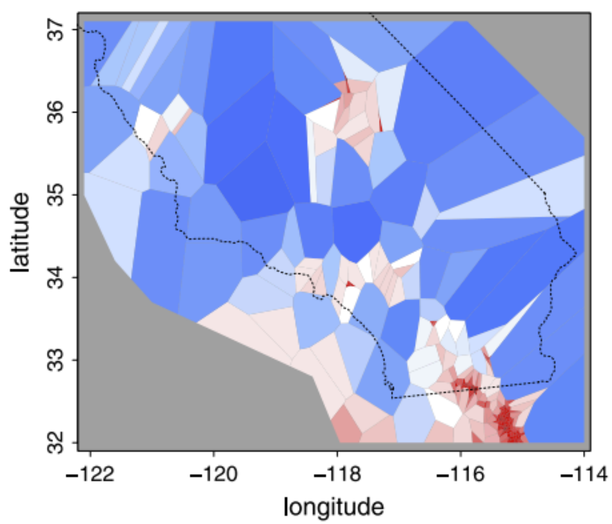
FIG. 2. (a) *Estimated rates under the Helmstetter, Kagan and Jackson* (2007) *model, with epicentral locations of observed earthquakes with* $M \geq 4.0$ *in Southern California between January* 1, 2006 *and January* 1, 2011 *overlaid.* (b) *Raw pixel residuals for Helmstetter, Kagan and Jackson* (2007) *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.*



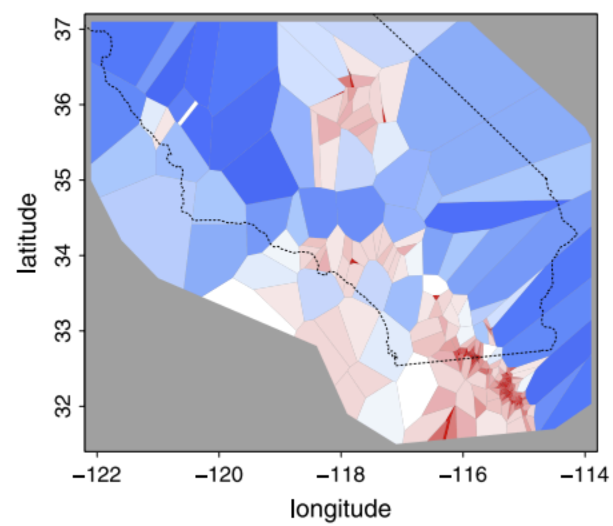
(a) Voronoi residuals for null model.



(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.

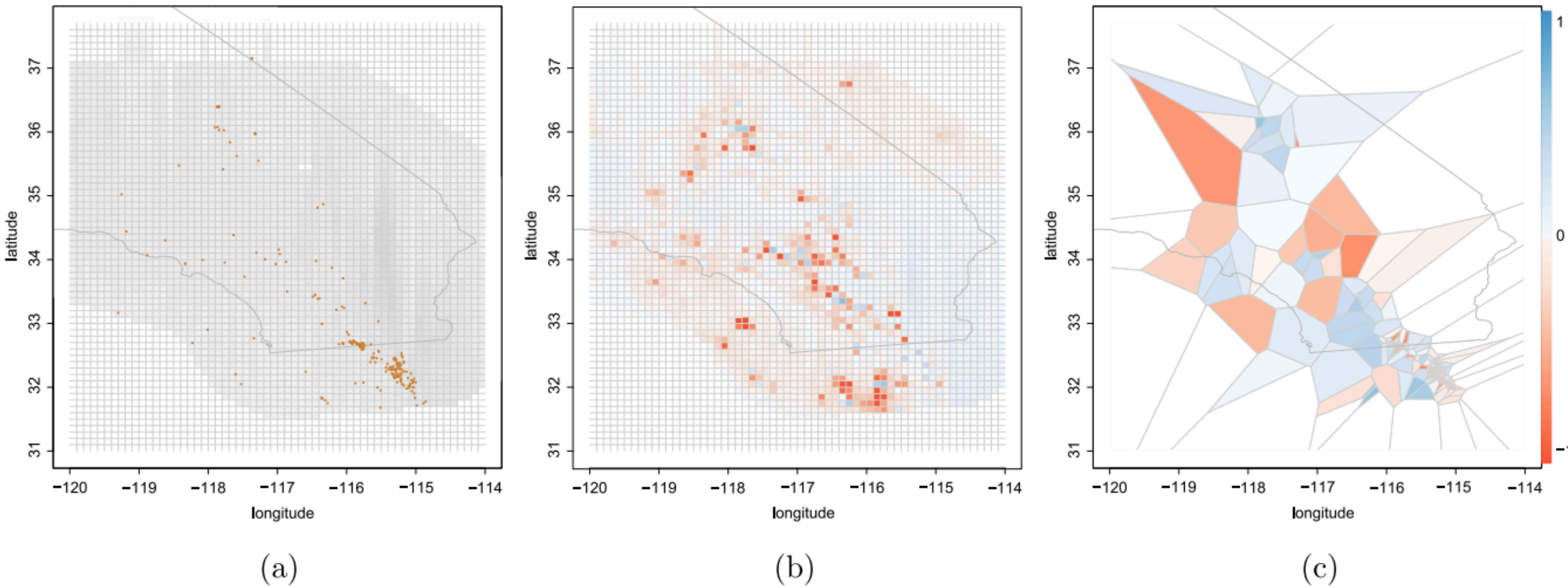


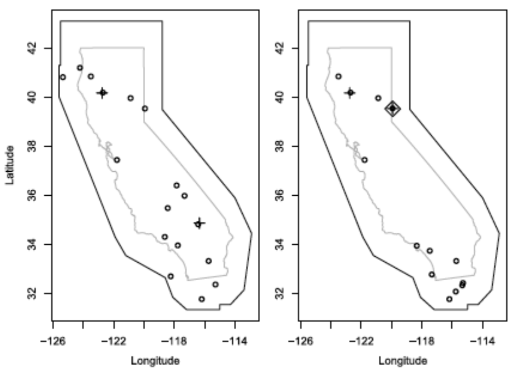
FIG. 3. (a) *Estimated rates under the Shen, Jackson and Kagan* (2007) *model, with epicentral locations of observed earthquakes with* $M \geq 4.0$ *in Southern California between January* 1, 2006 *and January* 1, 2011 *overlaid.* (b) *Pixel deviance plot with blue favoring model A, Helmstetter, Kagan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

3. Superthinning. (Clements et al., 2012)

Choose some number $c \sim \text{mean}(\hat{\lambda})$.

Superpose: where $\hat{\lambda}(t, x, y) < c$, add in points of a simulated Poisson process of rate $c - \hat{\lambda}(t, x, y)$.

Thin: where $\hat{\lambda}(t_i, x_i, y_i) > c$, keep each point (t_i, x_i, y_i) with prob. $c / \hat{\lambda}(t_i, x_i, y_i)$.



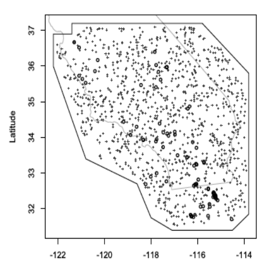


FIG. 9. *Superposed residuals for model C. Simulated points make up 90.7% of all points.*

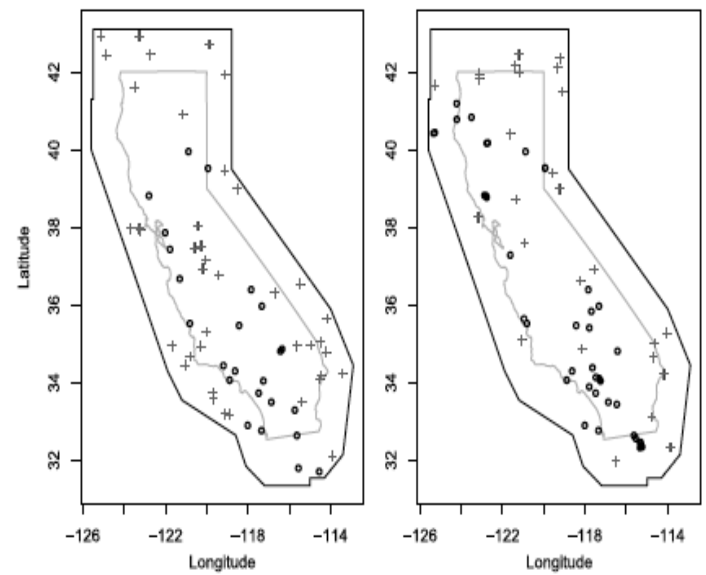


FIG. 11. *One realization of super-thinned residuals for the five models considered (circles = observed earthquakes; plus signs = simulated points). Top-left panel* (a): *model A* ($k = 2.76$). *Top-center panel* (b): *model B* ($k = 2.95$). *Top-right panel* (c): *model C* ($k = 2.73$). *Bottom-left panel* (d): *ETAS* ($k = 1.35$). *Bottom-right panel* (e): *STEP* ($k = 0.75$).

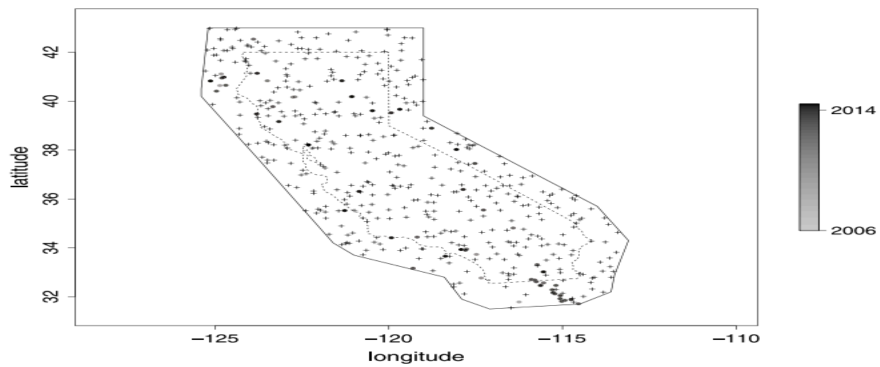
3. Superthinning. (Clements et al., 2013)

Choose some number c. mean($\hat{\lambda}$) at the points is a suggested default.

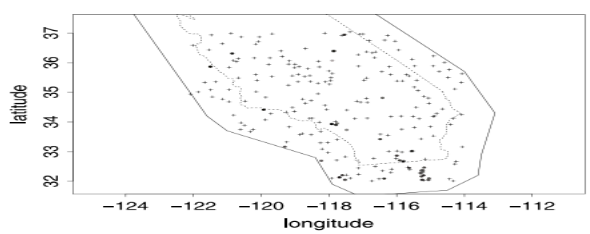
Superpose: where $\hat{\lambda}$ (t ,x,y) < c, add in points of a simulated Poisson process of rate c - $\hat{\lambda}$(t , x, y) .

Thin: where $\hat{\lambda}(t_i,x_i,y_i) > c$, keep each point (t_i,x_i,y_i) with prob. c / $\hat{\lambda}(t_i,x_i,y_i)$.

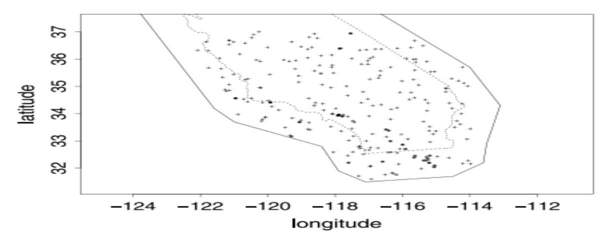
Result is Poisson with rate c, if the model for λ is correct.



(a) Super-thinned residuals Model A.



(b) Super-thinned residuals Model B.



(c) Super-thinned residuals Model C.

Exercises. Superposition.

Suppose N_1 is a Poisson process with rate 3,
and N_2 is a Poisson process with rate $2 + x + 4t$,
independent of N_1, and both are on $[0,10] \times [0,1] \times [0,1]$.
t x y.

Let $M = N_1 + N_2$. Is M a Poisson process? What is its intensity?

Exercises. Superposition.

Suppose N_1 is a Poisson process with rate 3,
and N_2 is a Poisson process with rate $2 + x + 4t$,
independent of N_1, and both are on $[0,10] \times [0,1] \times [0,1]$.
t x y.

Let $M = N_1 + N_2$. Is M a Poisson process? What is its intensity?

For any disjoint measurable sets $B_1, B_2, \ldots,$
$M(B_i) = N_1(B_i) + N_2(B_i)$ is independent of $\{N_1(B_j), j \neq i\}$ and $\{N_2(B_j), j \neq i\}$ and thus is independent of $\{N_1(B_j) + N_2(B_j), j \neq i\}$.

So yes, M is a Poisson process and since $EM(B) = EN_1(B) + EN_2(B)$, M has rate $5 + x + 4t$.

Exercises.

Suppose N is homogeneous Poisson process with rate 1,

and M is a clustered Hawkes process.

Both M and N have 40 points on B = [0,10] x [0,1] x [0,1]

t x y.

Let v1 = the average size of a Voronoi cell in a Voronoi tessellation of N,

and v2 = the average size of a Voronoi cell in a Voronoi tesselation of M.

Which is bigger, v1 or v2, or will they be the same?

Exercises.

Suppose N is homogeneous Poisson process with rate 1,

and M is a clustered Hawkes process.

Both M and N have 40 points on B = [0,10] x [0,1] x [0,1]

t x y.

Let v1 = the average size of a Voronoi cell in a Voronoi tessellation of N,

and v2 = the average size of a Voronoi cell in a Voronoi tesselation of M.

Which is bigger, v1 or v2, or will they be the same?

The same, since v1 = v2 = ¼. Each cell has one point, and the 40 cells occupy an area of size 10.

5. Review list.

1. PP as a random measure.
2. Integration, $\int f(t,x,y)\, dN$.
3. Simple and orderly.
4. Cond. intensity and Papangelou intensity.
5. Poisson processes.
6. Mixed Poisson processes.
7. Compound Poisson processes.
8. Poisson cluster processes.
9. Cox processes.
10. Gibbs and Strauss processes.
11. Matern processes.
12. Hawkes and ETAS processes.
13. Likelihood and MLE.
14. Simulation by thinning.
15. Martingale formula.
16. Kernel smoothing.
17. F,G,J,K, and L functions.
18. Marked G and J functions.
19. Weighted K function.
20. Nonparametric triggering function est.
21. Deviance, Voronoi, and superthinned residuals