Stat 13, Intro. to Statistical Methods for the Life and Health Sciences.

1. Review exercises.
2. Statistical analysis of wildfires.
3. Forecasting earthquakes.
4. Global temperature data.
5. Disease epidemics.

The final is Fri Dec14, 8-11am.
Bring a PENCIL and CALCULATOR and any books or notes you want. No computers.
http://www.stat.ucla.edu/~frederic/13/F18.

3. Review exercises.

5. Suppose the 100 bald men with a mean pulse of 67bpm have a median pulse of 70bpm. Describe the distribution of the pulses of bald men.

a. Left-skewed.

b. Right-skewed.

c. Symmetric.

d. Normal.

e. Confounding factors with heteroskedastic t-test confidence intervals.

f. None of the above.

6. Suppose the IQR is 20bpm, the range is (50bpm, 100bpm), and the 75^{th} percentile is 80bpm. What is the 25^{th} percentile pulse?

7. Find a 95% CI for the difference in mean pulse between the bald men and the long haired men. Assume the 100 bald men have mean 67bpm and sd 10bpm, and the 300 long haired men have mean 72bpm and sd 14bpm.

8. What can we conclude about statistical significance, based on this 95% CI?

3. Review exercises.

5. Left skewed, because the mean < median.

3. Review exercises.

5. Suppose the 100 bald men with a mean pulse of 67bpm have a median pulse of 70bpm. Describe the distribution of the pulses of bald men.

a. Left-skewed.

b. Right-skewed.

c. Symmetric.

d. Normal.

e. Confounding factors with heteroskedastic t-test confidence intervals.

f. None of the above.

6. Suppose the IQR is 20bpm, the range is (50bpm, 100bpm), and the 75^{th} percentile is 80bpm. What is the 25^{th} percentile pulse?

7. Find a 95% CI for the difference in mean pulse between the bald men and the long haired men. Assume the 100 bald men have mean 67bpm and sd 10bpm, and the 300 long haired men have mean 72bpm and sd 14bpm.

8. What can we conclude about statistical significance, based on this 95% CI?

3. Review exercises.

6. 60bpm.

3. Review exercises.

5. Suppose the 100 bald men with a mean pulse of 67bpm have a median pulse of 70bpm. Describe the distribution of the pulses of bald men.

a. Left-skewed.

b. Right-skewed.

c. Symmetric.

d. Normal.

e. Confounding factors with heteroskedastic t-test confidence intervals.

f. None of the above.

6. Suppose the IQR is 20bpm, the range is (50bpm, 100bpm), and the 75^{th} percentile is 80bpm. What is the 25^{th} percentile pulse?

7. Find a 95% CI for the difference in mean pulse between the bald men and the long haired men. Assume the 100 bald men have mean 67bpm and sd 10bpm, and the 300 long haired men have mean 72bpm and sd 14bpm.

8. What can we conclude about statistical significance, based on this 95% CI?

3. Review exercises.

7. -5 +/- 1.96 SE, where SE = $\sqrt{(100/100 + 14^2/300)} = 1.28582$. So the 95% CI is -5 +/- 1.96(1.28582) = -5 +/- 2.53.

3. Review exercises.

5. Suppose the 100 bald men with a mean pulse of 67bpm have a median pulse of 70bpm. Describe the distribution of the pulses of bald men.

a. Left-skewed.

b. Right-skewed.

c. Symmetric.

d. Normal.

e. Confounding factors with heteroskedastic t-test confidence intervals.

f. None of the above.

6. Suppose the IQR is 20bpm, the range is (50bpm, 100bpm), and the 75^{th} percentile is 80bpm. What is the 25^{th} percentile pulse?

7. Find a 95% CI for the difference in mean pulse between the bald men and the long haired men. Assume the 100 bald men have mean 67bpm and sd 10bpm, and the 300 long haired men have mean 72bpm and sd 14bpm.

8. What can we conclude about statistical significance, based on this 95% CI?

3. Review exercises.

8. 0 is not in the interval, so the difference in pulse between bald men and long haired men is statistically significant.

Forecasting wildfires.

(from Alexander & Cruz, 2013)

Table 1
Classification of surface fire spread models (1946–2000)

Reference	Type	Origin
Foens [4]	Theoretical	United States
Emmons [5]	Theoretical	United States
Hottel et al. [6]	Theoretical	United States
McArthur [7]	Empirical[a]	Australia
Van Wagner [8]	Theoretical	Canada
Thomas [9]	Theoretical	United Kingdom
McArthur [10]	Empirical[a]	Australia
Anderson [11]	Theoretical	United States
Frandsen [12]	Semiempirical	United States
Rothermel [13]	Semiempirical[a]	United States
Pagni and Peterson [14]	Theoretical	United States
Telisin [15]	Theoretical	Russia
Steward [16]	Theoretical	United States
Konev and Sukhinin [17]	Theoretical	Russia
Cekirge [18]	Theoretical	United States
Fujii et al. [19]	Theoretical	Japan
Grishin et al. [20]	Theoretical	Russia
Griffin and Allan [21]	Semiempirical	Australia
Huang and Xie [22]	Theoretical	United States
Sneeuwjagt and Peet [23]	Semiempirical	Australia
Albini [24,25]	Theoretical	United States
De Mestre et al. [26]	Theoretical	Australia
Weber [27]	Theoretical	Australia
Borrows et al. [28]	Semiempirical	Australia
Forestry Canada Fire Danger Group [29]	Empirical[a]	Canada
Croba et al. [30]	Theoretical	Greece
Marsden-Smedley and Catchpole [31]	Semiempirical	Australia
Grishin [32]	Theoretical	Russia
Dupuy [33]	Theoretical	France
Santoni and Balbi [34]	Theoretical	France
Linn [35]	Theoretical	United States
Catchpole et al. [36]	Semiempirical	Australia
Catchpole et al. [37]	Semiempirical	Australia
Fernandes [38]	Semiempirical	Portugal
Vega [39]	Semiempirical	Spain
McCaw [40]	Semiempirical	Australia
Viegas et al. [41]	Empirical	Portugal
Cheney et al. [42]	Empirical	Australia
Larini et al. [43]	Theoretical	France
Margerit and Guillaume [44]	Theoretical	France
Burrows [45,46]	Semiempirical	Australia
Hargrove et al. [47]	Empirical[a]	United States

[a] Models that constitute the basis of operating tools actually used in forestry agencies.

(from Pastor et al., 2003)

2) Currently used evaluation methods for spread models.

a) Is the problem of modeling spread solved?

Many simplifying assumptions in current models.

Simulations can do very well retrospectively,
on data used to make the sims.

b) Methods for model assessment

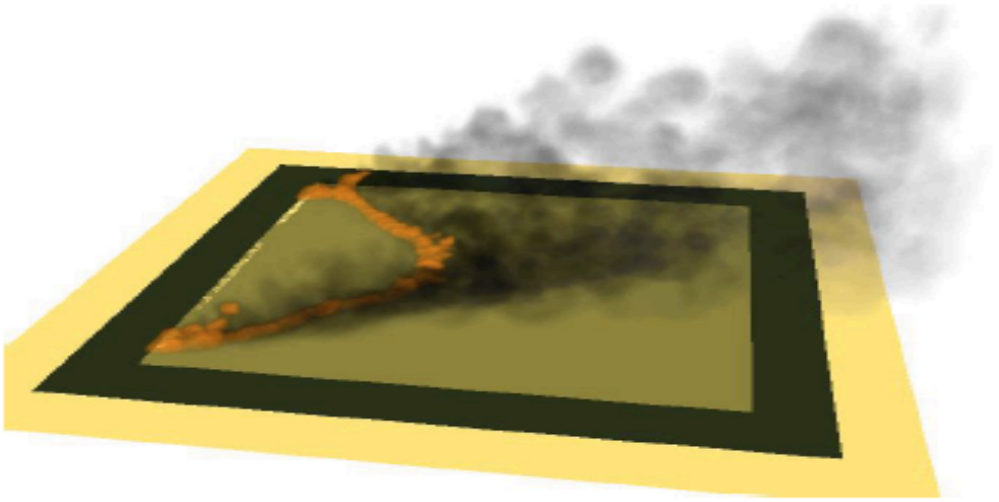
Evaluate a particular statistic, like surface temperature or spread rate.

Comparison of final burn areas.

Error matrices, Cohen's *K,* Sørensen's *Q*.

Mell et al. (2007), **retrospective.** The experiment shown here was used in model construction.

(a)



(b)

Figure 4: (a) Photograph of experimental fire F19 at $t = 56$ s. (b) Snapshot of WFDS simulation of experimental fire F19 at $t = 56$ s.

Observed vs. Predicted spread rate (Cruz & Alexander, 2013), **retrospective:**

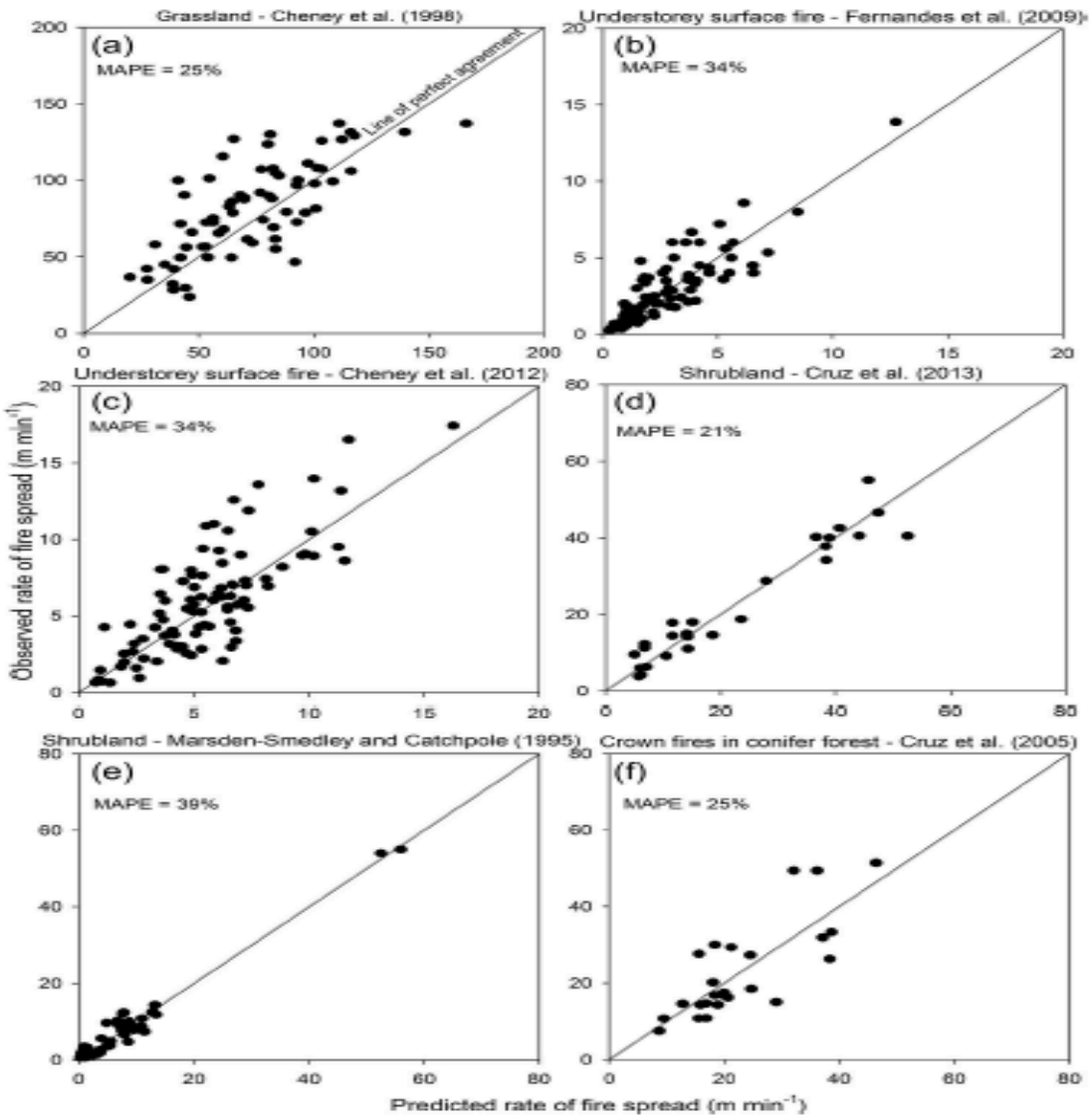


Fig. 3. Observed rates of spread used in model development versus model predictions for experimental surface fires in (a) grasslands fires (Cheney et al., 1998), (b) surface fire in maritime pine forest (Fernandes et al., 2009), (c) surface fire in dry eucalypt forest (Cheney et al., 2012), (d) surface and crown fires in mallee-heath shrublands (Cruz et al., 2013), (e) buttongrass moorlands (Marsden-Smedley and Catchpole, 1995), and (f) crown fires in conifer forests (Cruz et al., 2005). MAPE = mean absolute percent error.

Observed vs. Predicted spread rate (Cruz & Alexander, 2013), **prospective:**

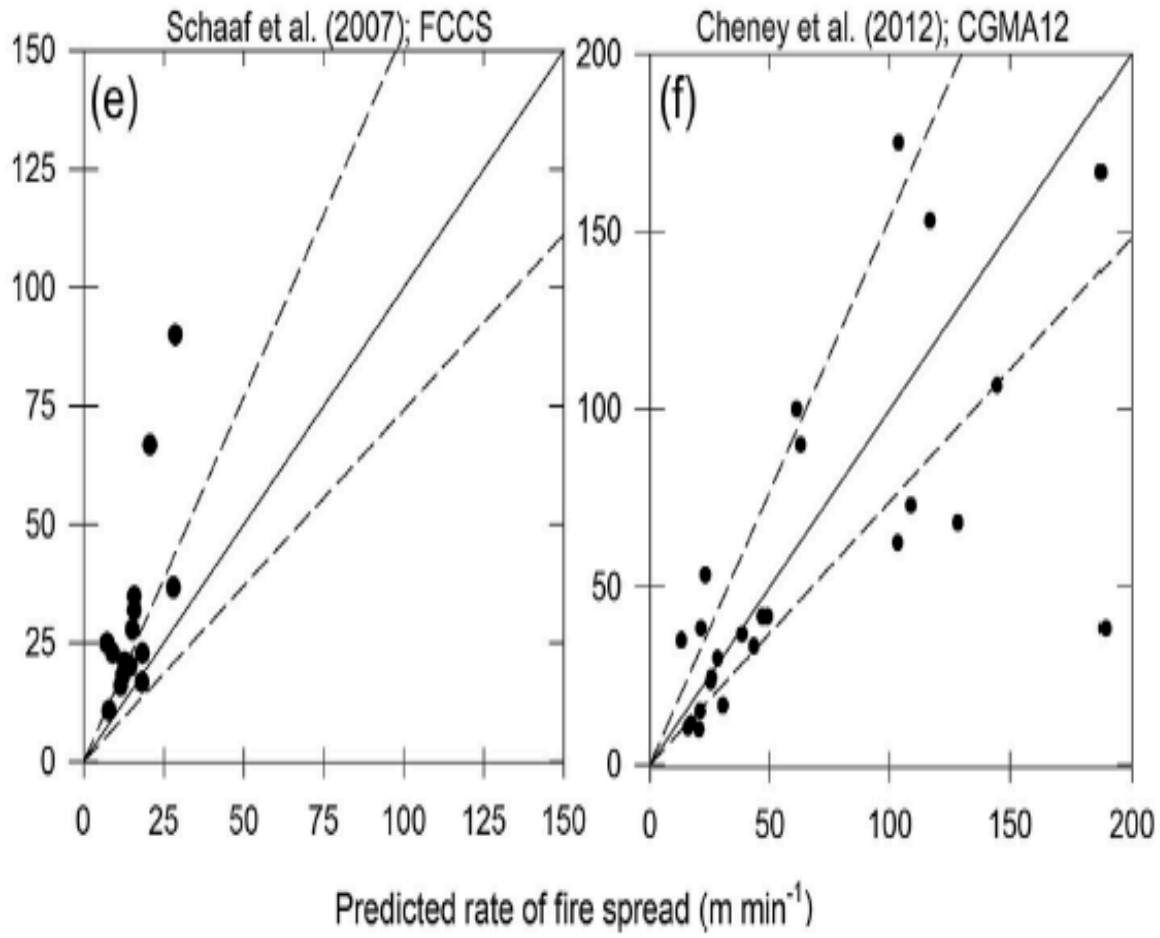


Fig. 6. Observed rates of spread versus model predictions for a selection of studies presented in Table 1 featuring both 'moderately slow' and 'exceedingly fast' spreading fires (i.e. upwards of ~150 m min^{-1}). The dashed lines around the line of perfect agreement indicate the ±35% error interval. Refer to Table 2 for the mean of the model abbreviations (e.g. RCR72 = Rothermel, 1972). MAPE = mean absolute percent error.

Observed and Predicted surface temperature (de Mestre et al., 1989, from Weber 1991)

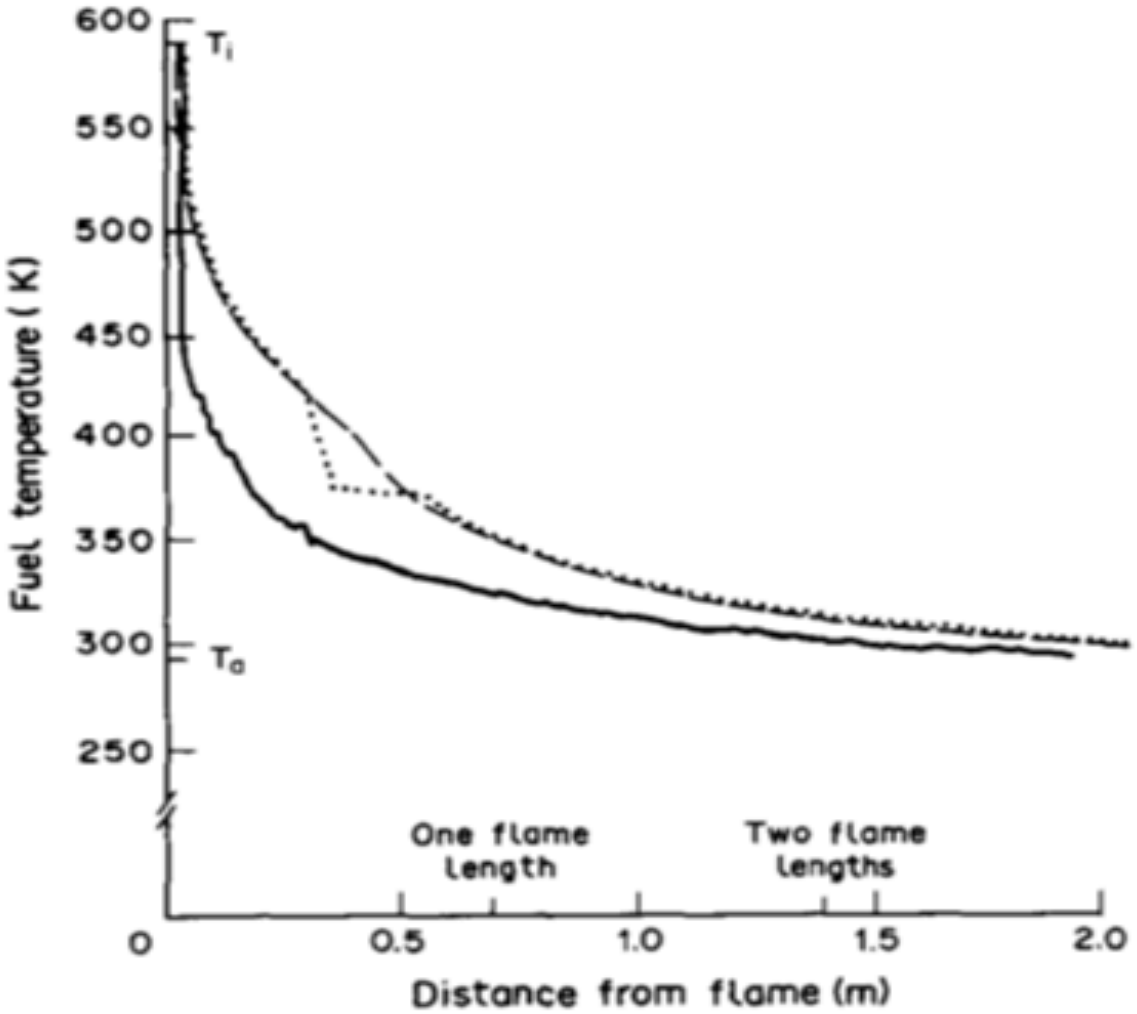


FIG. 2. Fuel surface temperature ahead of the fire front. Experimental (solid line), 'frozen' moisture model (dotted line), and averaged moisture model (dashed line) are shown. From de Mestre *et al.*[3]

Errors in windspeed measurements (Sullivan and Knight, 2001)

Table 2. Errors in estimating 5-min wind averages at a fire front as a percentage of the measured wind within 68% confidence limits (1 SD) assuming a standard deviation of 27% of the mean wind speed.

No. of anemometers	Fire width			
	0–40 m	≈80 m	≈160 m	≈300 m
1	±38%	±33%	±30%	±28%
2	±33%	±27%	±23%	±21%
4	±30%	±23%	±19%	±16%
8	±28%	±21%	±16%	±13%

One anemometer can predict the wind speed at another location (say at the front of a small fire) with a probability curve associated with the convolution of two 26.7% curves, i.e., from eq. 8, $s_f = 37.8\%$ of the mean wind speed. In other words there is a 68% chance of the wind speed at the fire front being within ±37.8% of the measured wind speed. This is not particularly good.

Observed and Simulated burn polygons (Arca et al., 2007)

Time step	Observed ROS	Simulated ROS
1	7.0	6.5
2	12.4	10.3
3	6.6	7.4
whole area	8.1	8.1

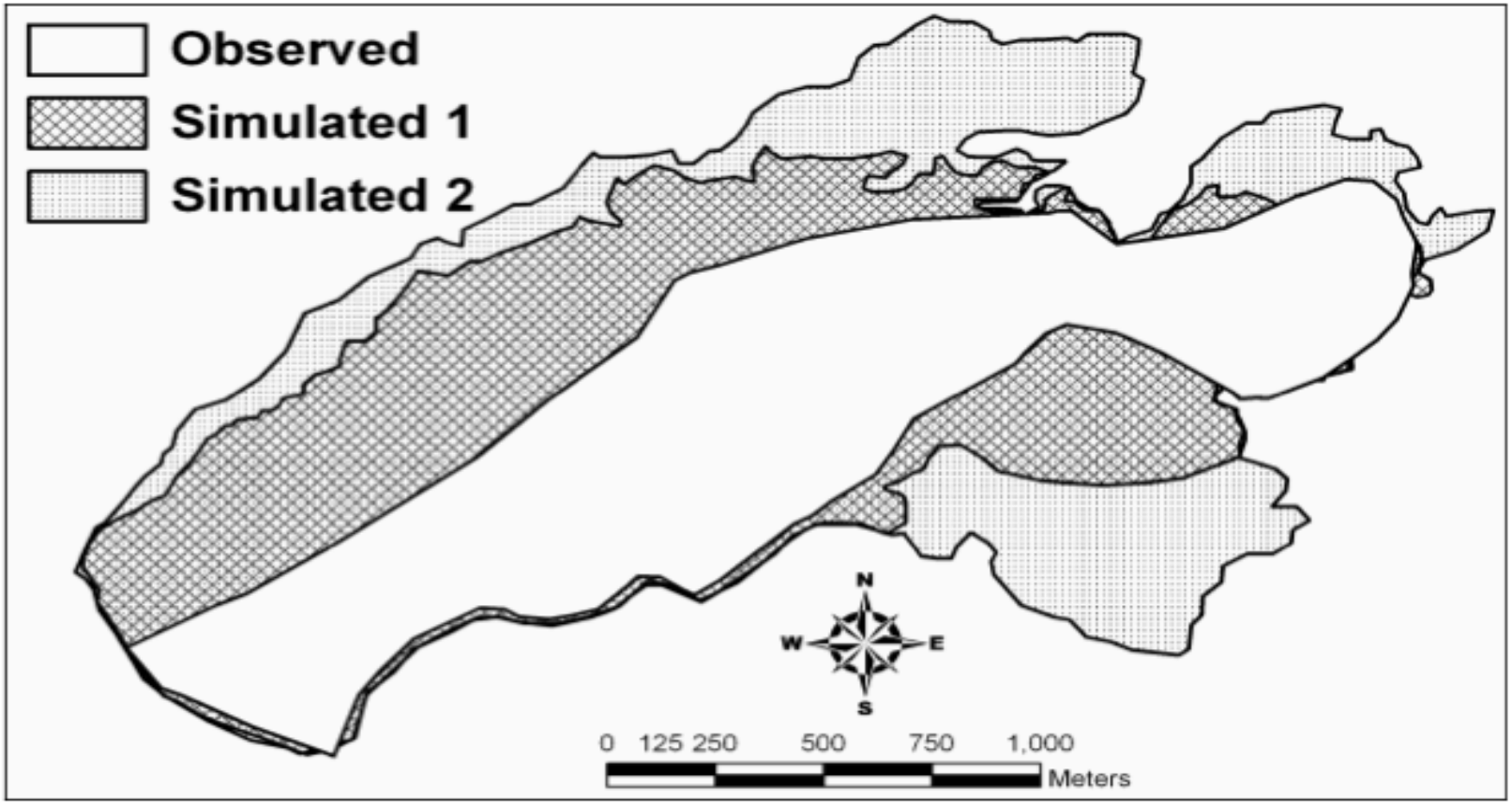


Figure 2 – Comparison between observed and simulated fire areas from the simulation n. 4 (custom fuel model CM28) using raster wind maps (Simulated 1) and constant wind field (Simulated 2).

Observed and Simulated burn polygons (Fujioka, 2002)

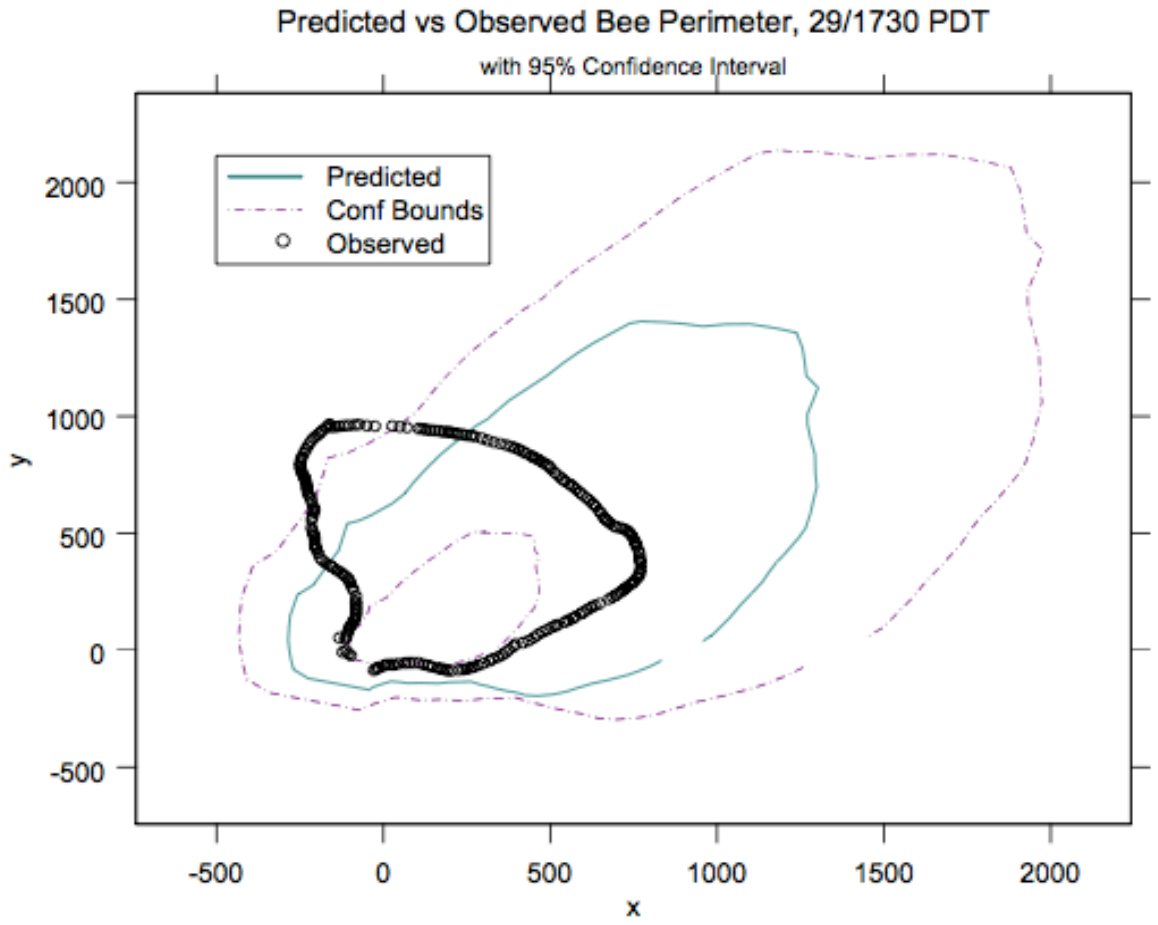


Figure 7. Position and error corrected perimeter prediction for the Bee Fire, 29 June 1996, 1730 PDT, and approximate 95% confidence interval for the true perimeter.

Remote Sensing techniques.

Error matrix (Congalton 1991)

Table 1. An Example Error Matrix

	Reference Data				
	D	C	BA	SB	row total
D	65	4	22	24	115
C	6	81	5	8	100
BA	0	11	85	19	115
SB	4	7	3	90	104
column total	75	103	115	141	434

Land Cover Categories

D = deciduous

C = conifer

BA = barren

SB = shrub

OVERALL ACCURACY = 321/434 = 74%

PRODUCER'S ACCURACY

D = 65 / 75 = 87%
C = 81 / 103 = 79%
BA = 85 / 115 = 74%
SB = 90 / 141 = 64%

USER'S ACCURACY

D = 65 / 115 = 57%
C = 81 / 100 = 81%
BA = 85 / 115 = 74%
SB = 90 / 104 = 87%

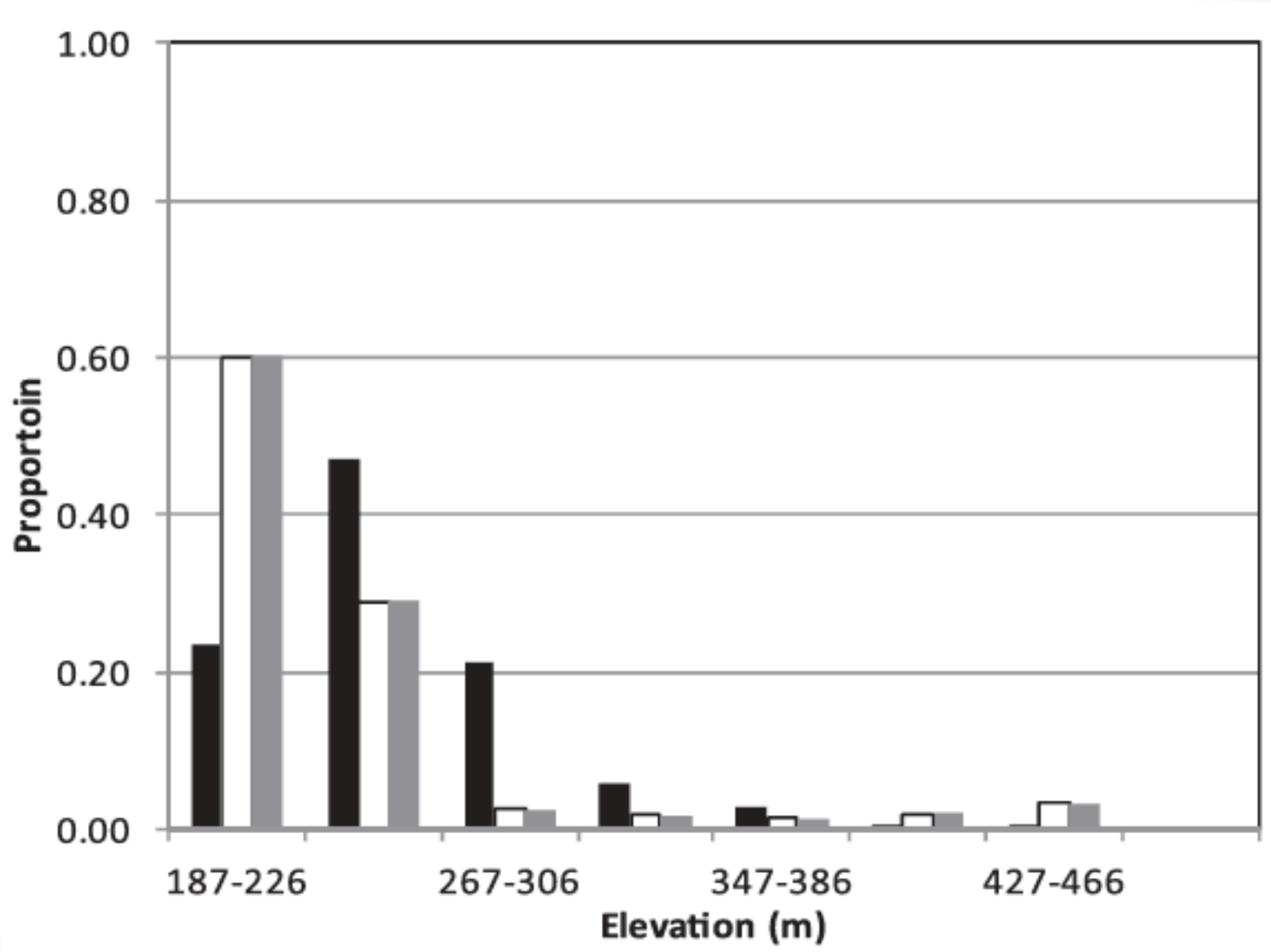
Cohen's *K* (Congalton 1991)

$$\hat{K} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} * x_{+i})},$$

Sørensen's original formula was intended to be applied to presence/absence data, and is

$$QS = \frac{2C}{A+B} = \frac{2|A \cap B|}{|A| + |B|}$$

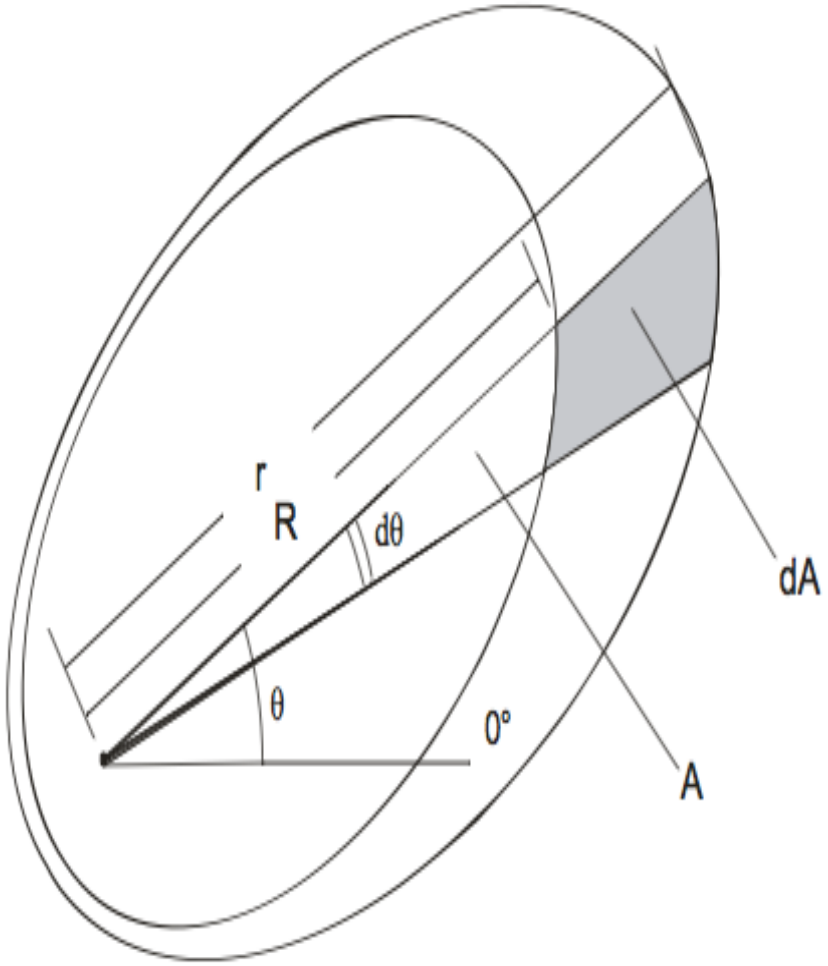
Overlap proportion (Duff et al., 2013). Black = overlapping burn area,
white = burn area predicted but not observed, grey = burn area observed but not predicted.

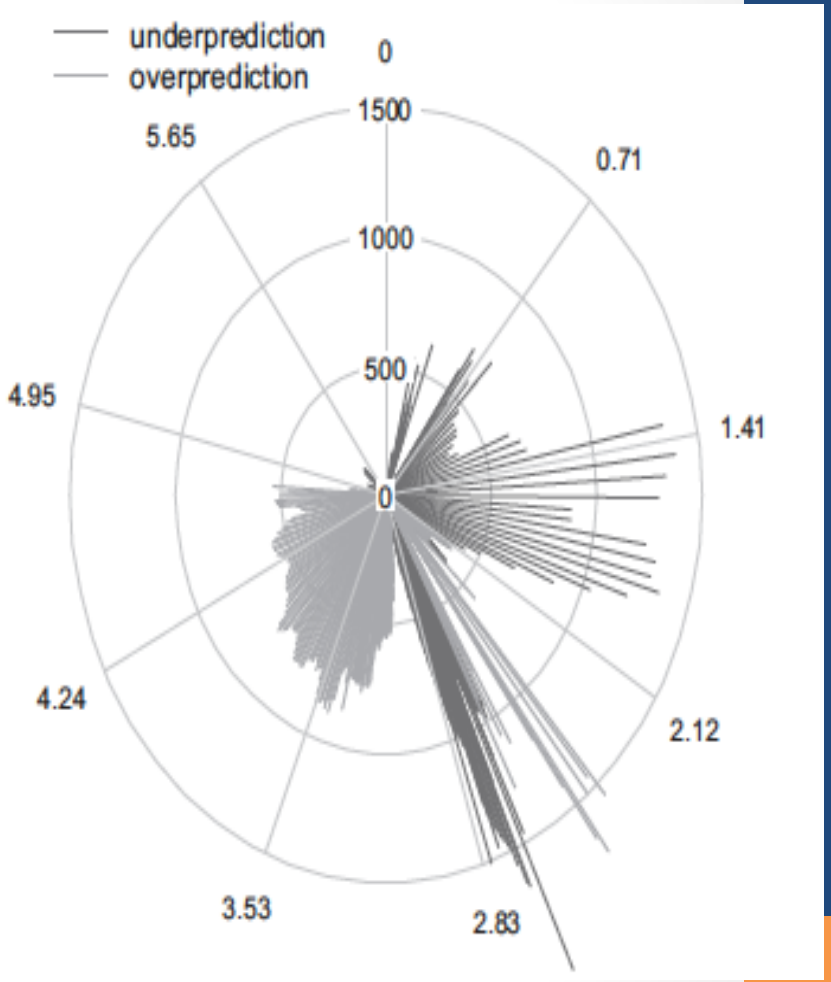


Difference between burn areas, evaluated radially as a function of θ

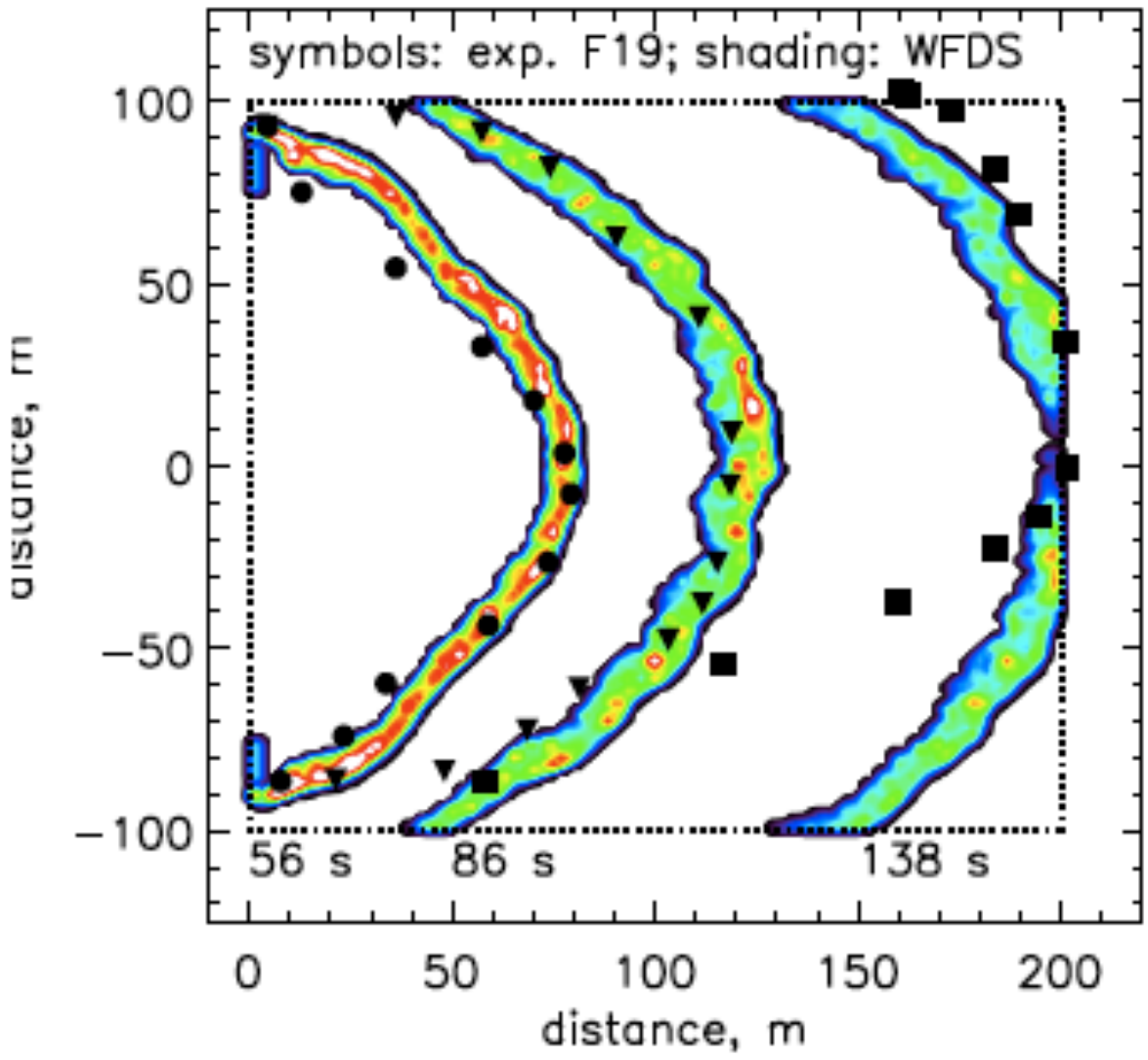
(Cui & Perera, 2010)

(Duff et al., 2013)





Few studies look at evolution of predicted & observed burn areas over time (Mell et al. 2007)



Burning Index (BI)

NFDRS:

Spread Component (SC) and Energy Release Component (ERC), each based on dozens of equations.

$$BI = [10.96 \times SC \times ERC]^{0.46}$$

- Uses daily weather variables, drought index, and vegetation info. Human interactions excluded.
- Predicts: flame length,
  … area/fire? # of fires? # of fires? Total burn area?

Percentage of days with wildfires vs. Index (Viegas et al., 1999)

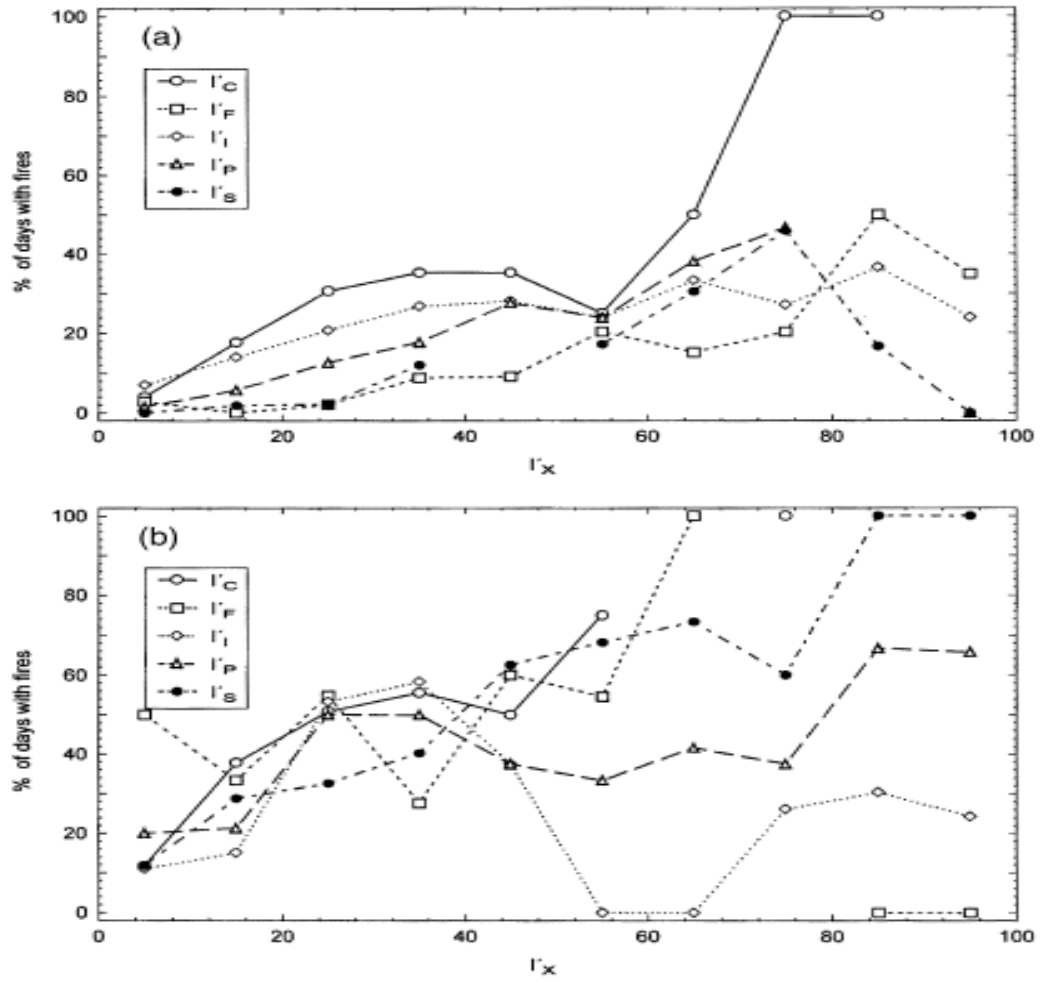


Figure 3. Percentage y_6 of days with fires: (a) A. H. Provence; (b) Veneto.

Mean area burned per day vs. Index (Viegas et al., 1999)

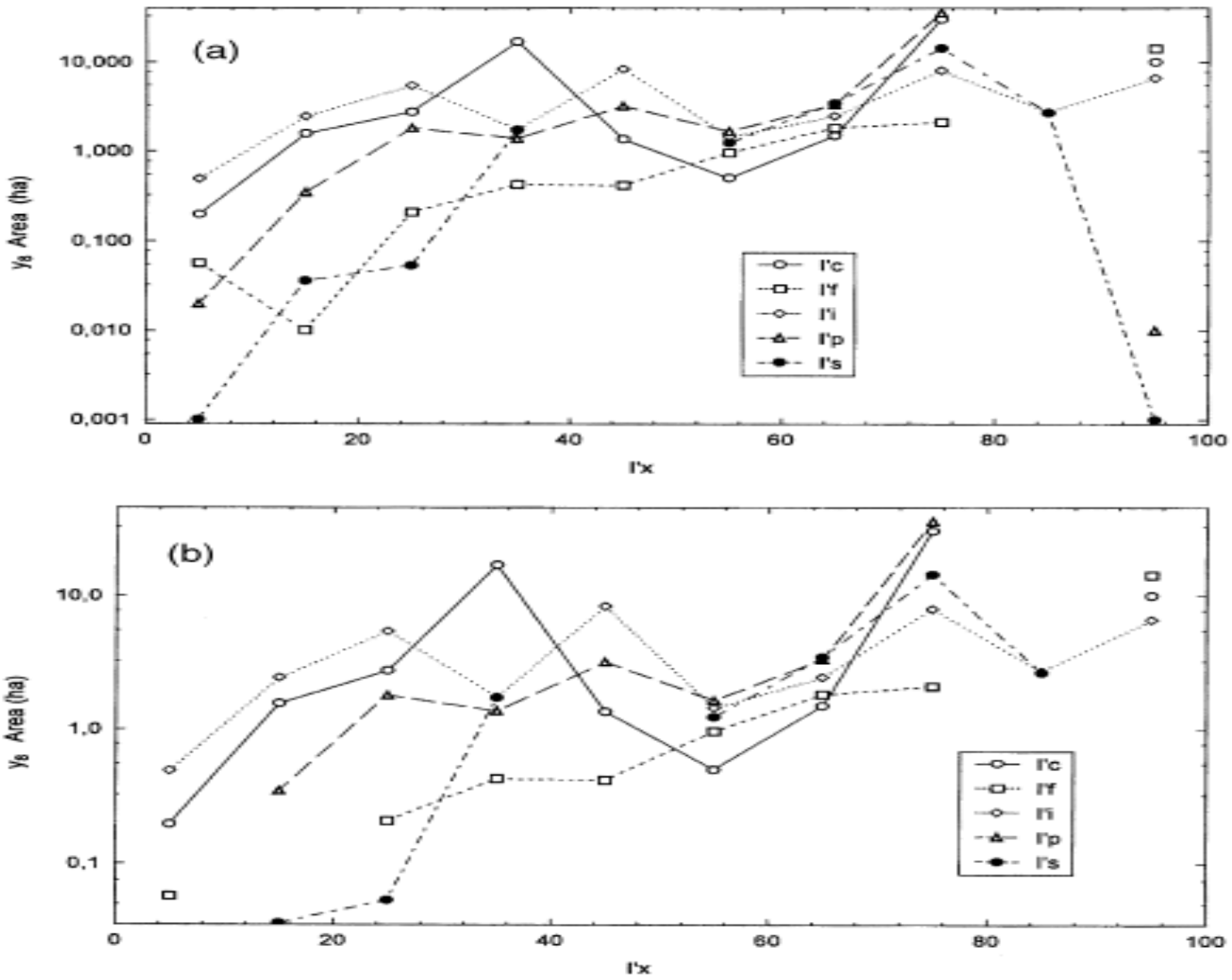


Figure 4. Average area y_8 (ha) burned daily: (a) A. H. Provence; (b) Veneto.

Some BI equations: (From Pyne et al., 1996:)

Rate of spread: $R = I_R \, \xi \, (1 + \phi_w + \phi_s) / (\rho_b \varepsilon \, Q_{ig})$.

Oven-dry bulk density: $\rho_b = w_0 / \delta$.

Reaction Intensity: $I_R = \Gamma' \, w_n \, h \, \eta_M \eta_s$.

Effective heating number: $\varepsilon = \exp(-138/\sigma)$.

Optimum reaction velocity: $\Gamma' = \Gamma'_{max} \, (\beta / \beta_{op})^A \exp[A(1 - \beta / \beta_{op})]$.

Maximum reaction velocity: $\Gamma'_{max} = \sigma^{1.5} \, (495 + 0.0594 \, \sigma^{1.5})^{-1}$.

Optimum packing ratios: $\beta_{op} = 3.348 \, \sigma^{-0.8189}$. $A = 133 \, \sigma^{-0.7913}$.

Moisture damping coef.: $\eta_M = 1 - 259 \, M_f/M_x + 5.11 \, (M_f/M_x)^2 - 3.52 \, (M_f/M_x)^3$.

Mineral damping coef.: $\eta_s = 0.174 \, S_e^{-0.19}$ (max = 1.0).

Propagating flux ratio: $\xi = (192 + 0.2595 \, \sigma)^{-1} \exp[(0.792 + 0.681 \, \sigma^{0.5})(\beta + 0.1)]$.

Wind factors: $\sigma_w = C U^B \, (\beta/\beta_{op})^{-E}$. $C = 7.47 \exp(-0.133 \, \sigma^{0.55})$. $B = 0.02526 \, \sigma^{0.54}$.
$E = 0.715 \exp(-3.59 \times 10^{-4} \, \sigma)$.

Net fuel loading: $w_n = w_0 \, (1 - S_T)$.

Heat of preignition: $Q_{ig} = 250 + 1116 \, M_f$.

Slope factor: $\phi_s = 5.275 \, \beta^{-0.3} \, (\tan \phi)^2$.

Packing ratio: $\beta = \rho_b / \rho_p$.

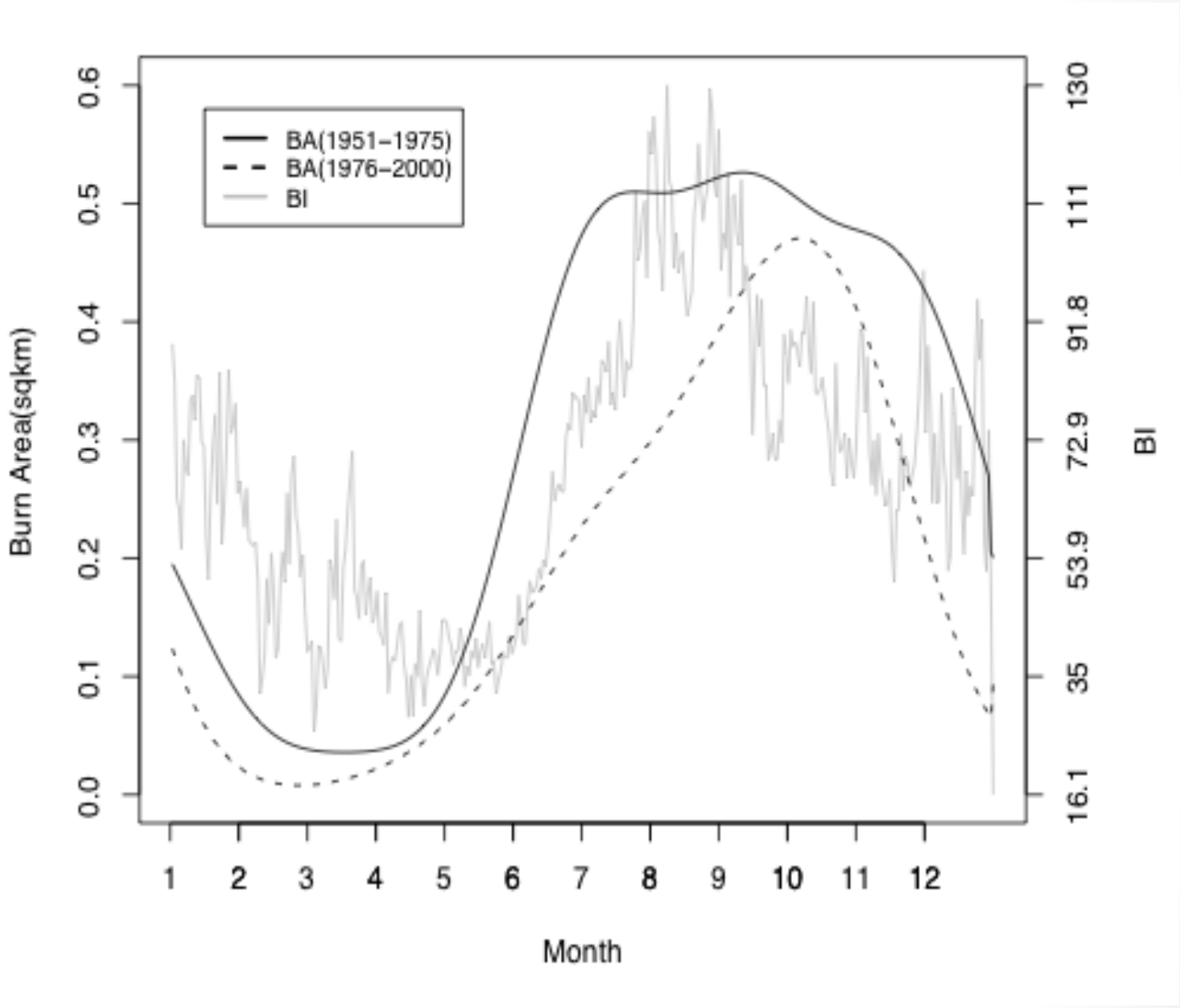
Good news for BI:

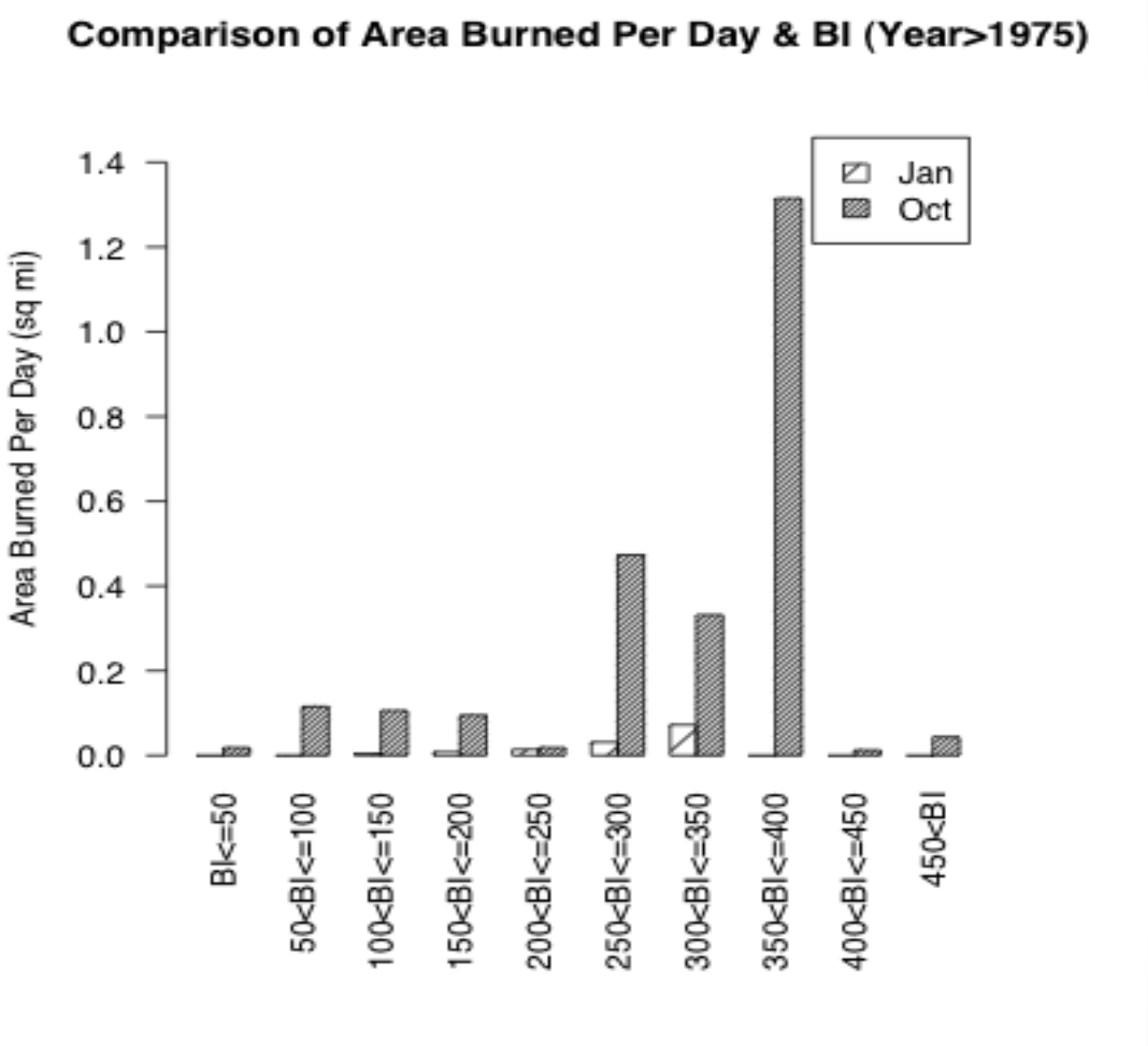
- BI is positively associated with wildfire occurrence.
- Positive correlations with number of fires, daily area burned, and area per fire.
- Properly emphasizes windspeed, relative humidity.

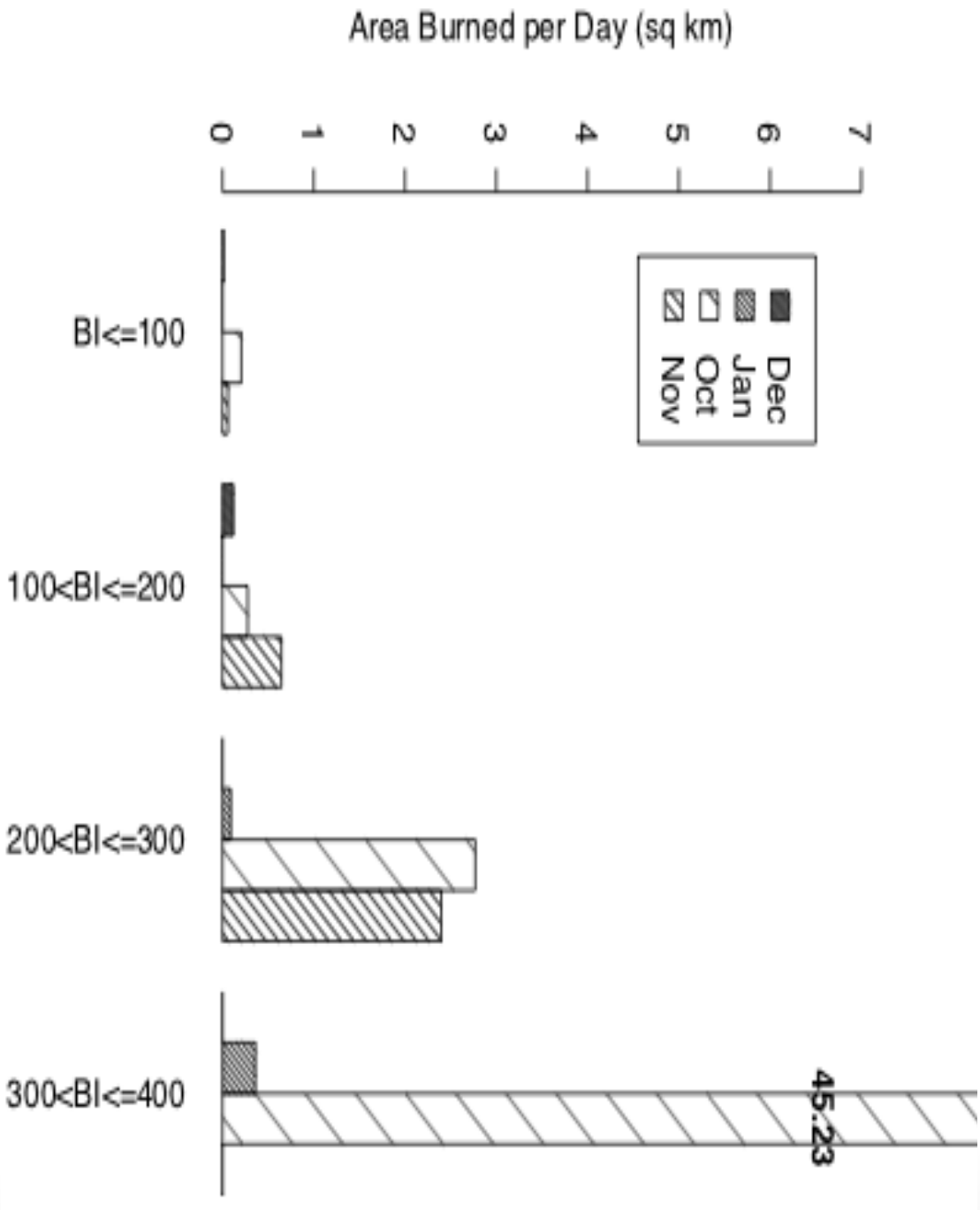
Some problems with BI

- Correlations are low.
  - Corr(BI, area burned) = 0.09
  - Corr(BI, # of fires) = 0.13
  - Corr(BI, area per fire) = 0.076
  - ! Corr(date, area burned) = 0.06
  - ! Corr(windspeed, area burned) = 0.159

- Too high in Winter (esp Dec and Jan)
  Too low in Fall (esp Sept and Oct)

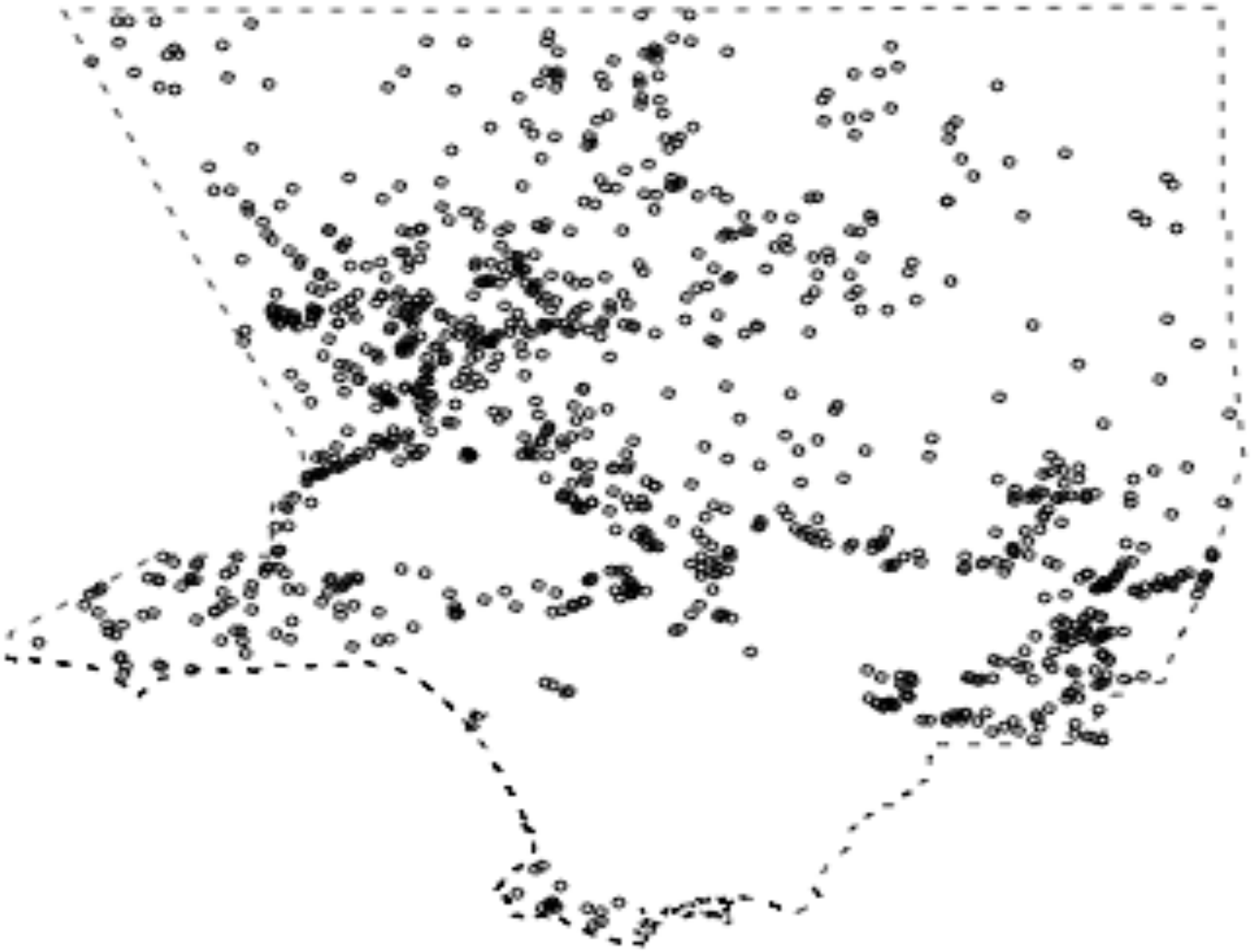

BA(1951–1975)
BA(1976–2000)
BI
Burn Area(sqkm)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
BI
16.1
35
53.9
72.9
91.8
111
130
1
2
3
4
5
6
7
8
9
10
11
12
Month

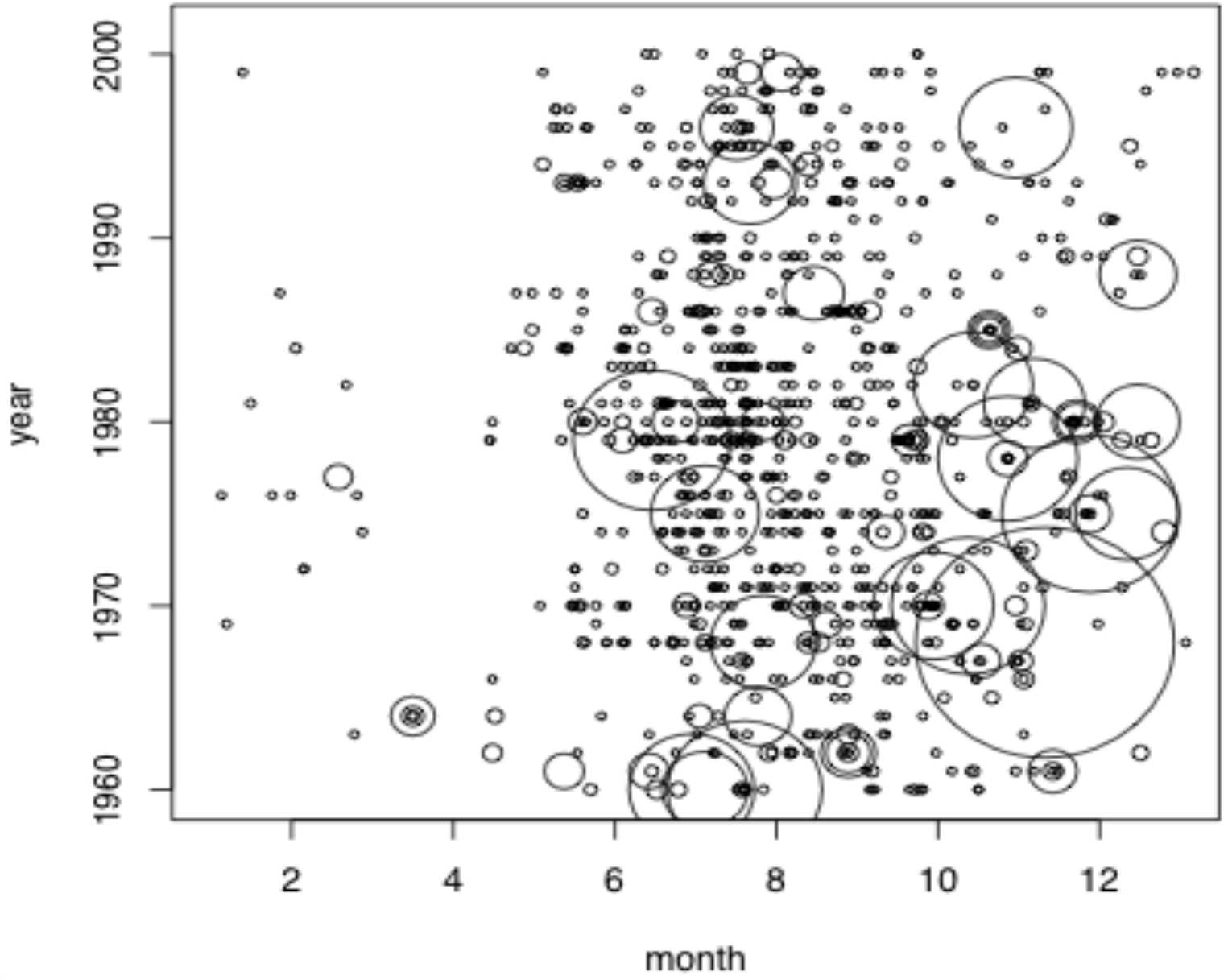

Comparison of Area Burned Per Day & BI (Year>1975)
Area Burned Per Day (sq mi)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Jan
Oct
BI<=50
50<BI<=100
100<BI<=150
150<BI<=200
200<BI<=250
250<BI<=300
300<BI<=350
350<BI<=400
400<BI<=450
450<BI

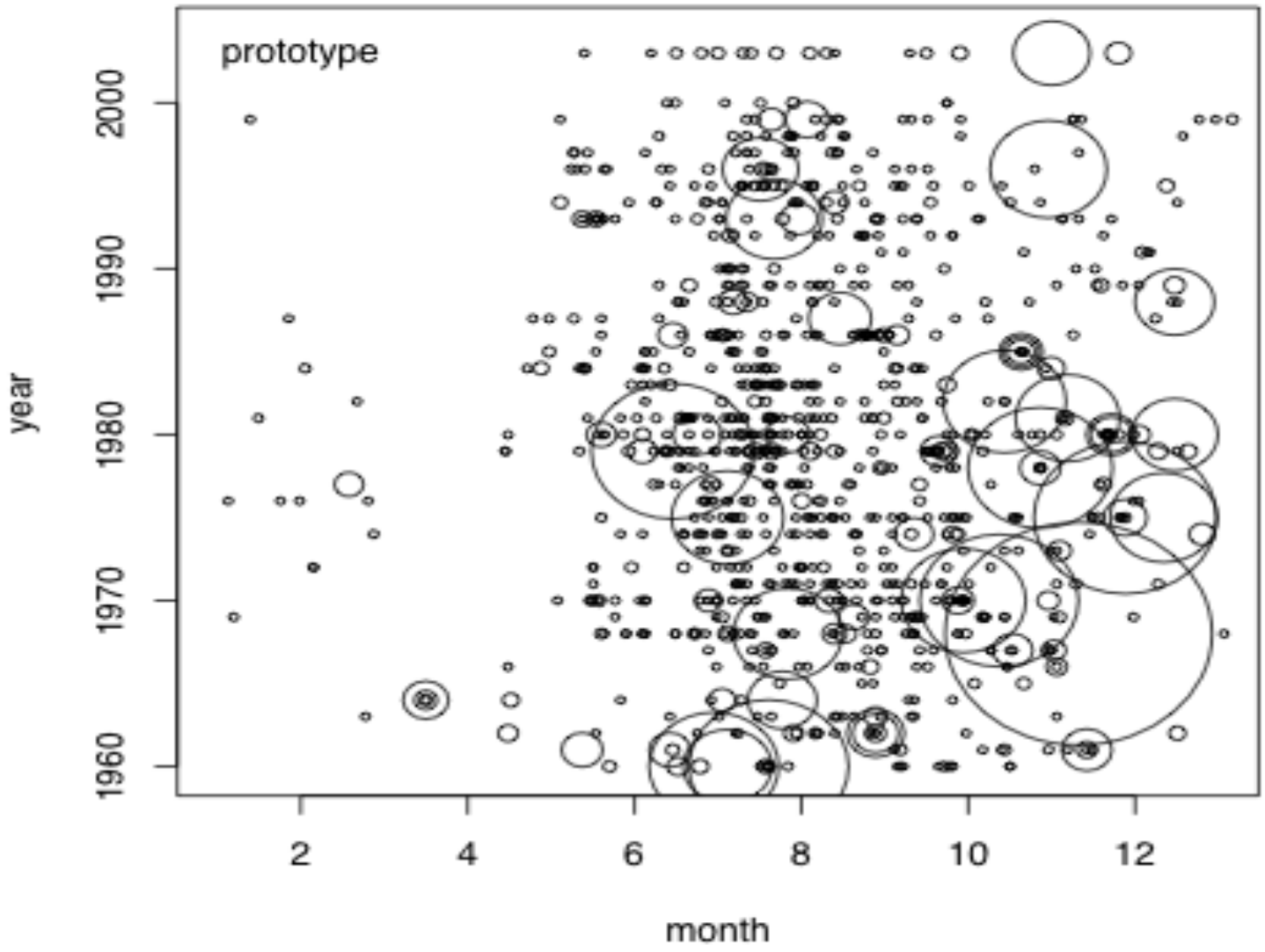

Area Burned per Day (sq km)
0
1
2
3
4
5
6
7
BI<=100
100<BI<=200
200<BI<=300
300<BI<=400
Dec
Jan
Oct
Nov
45.23

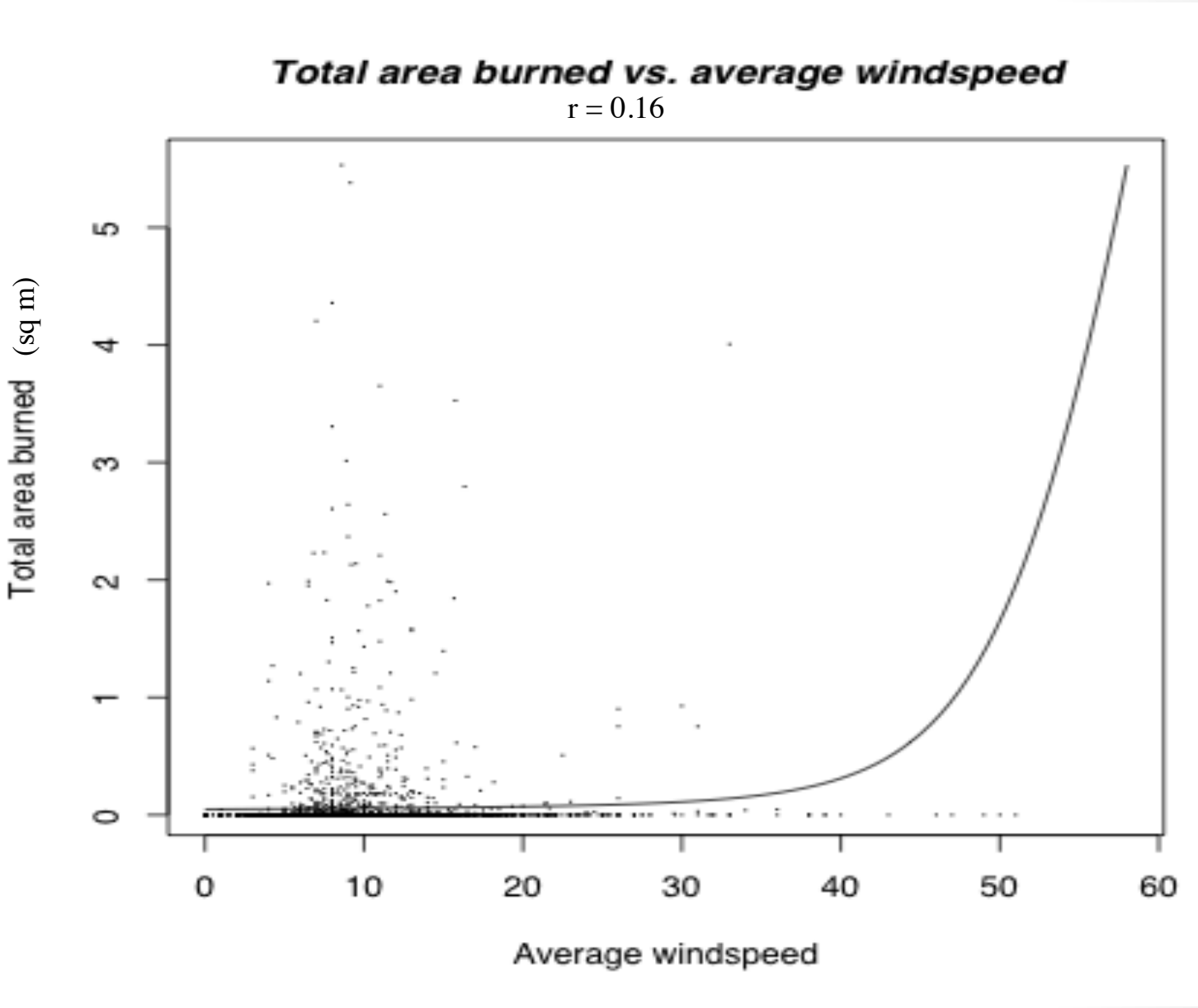
Separable Estimation for Point Processes

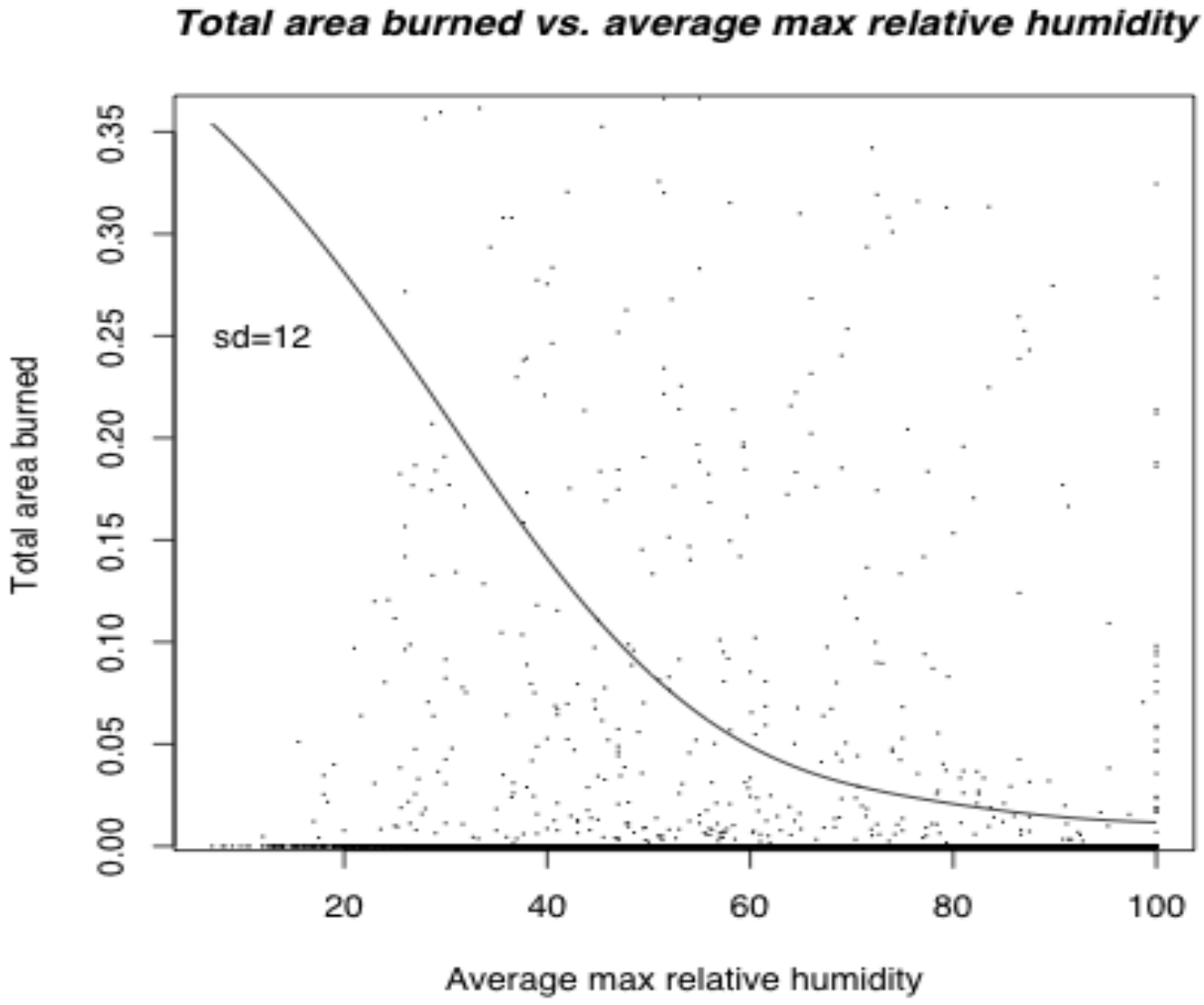
- Consider $\lambda(t, x_1, \ldots, x_k; \theta)$. [For fires, x_1=location, x_2 = area.]

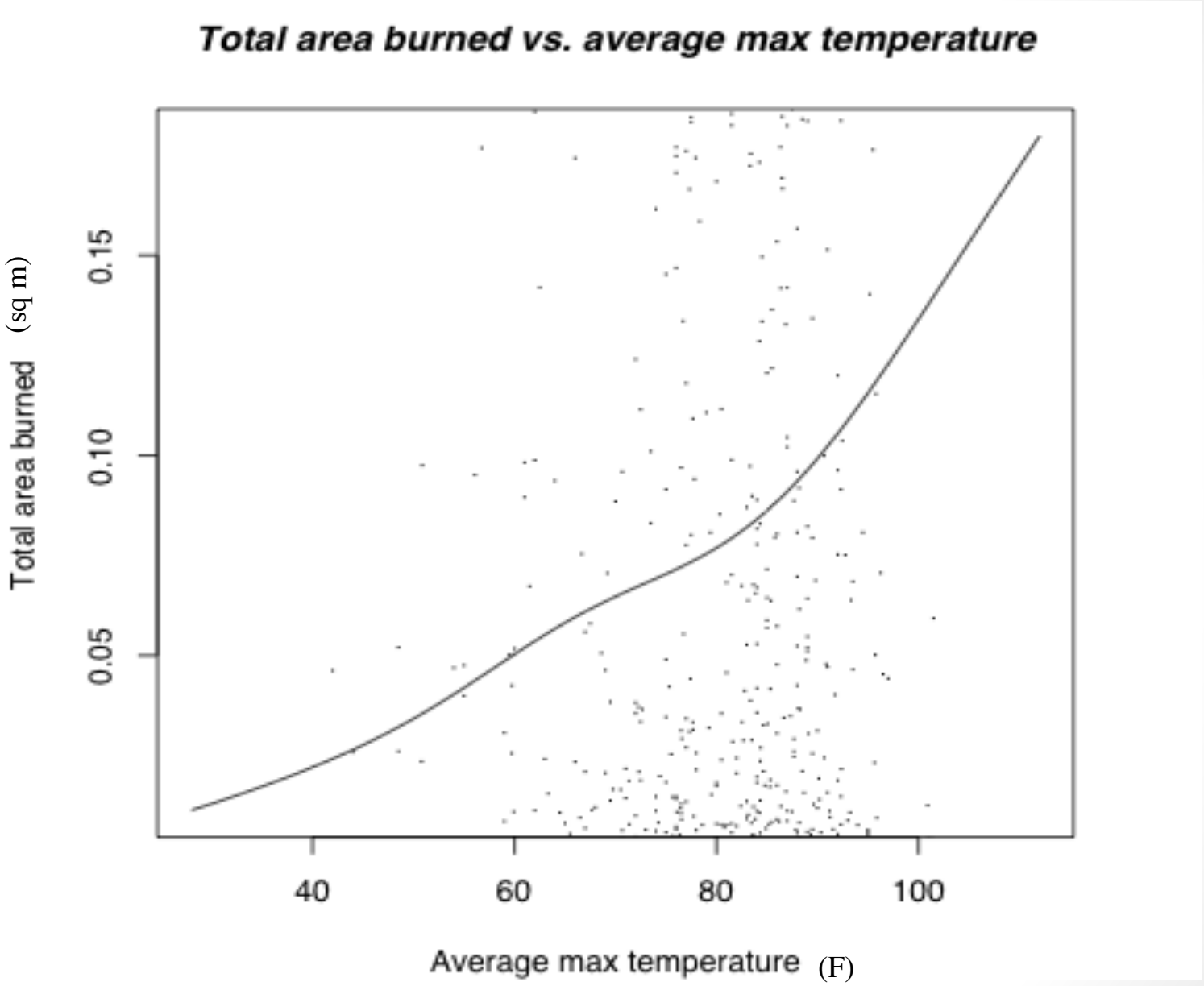


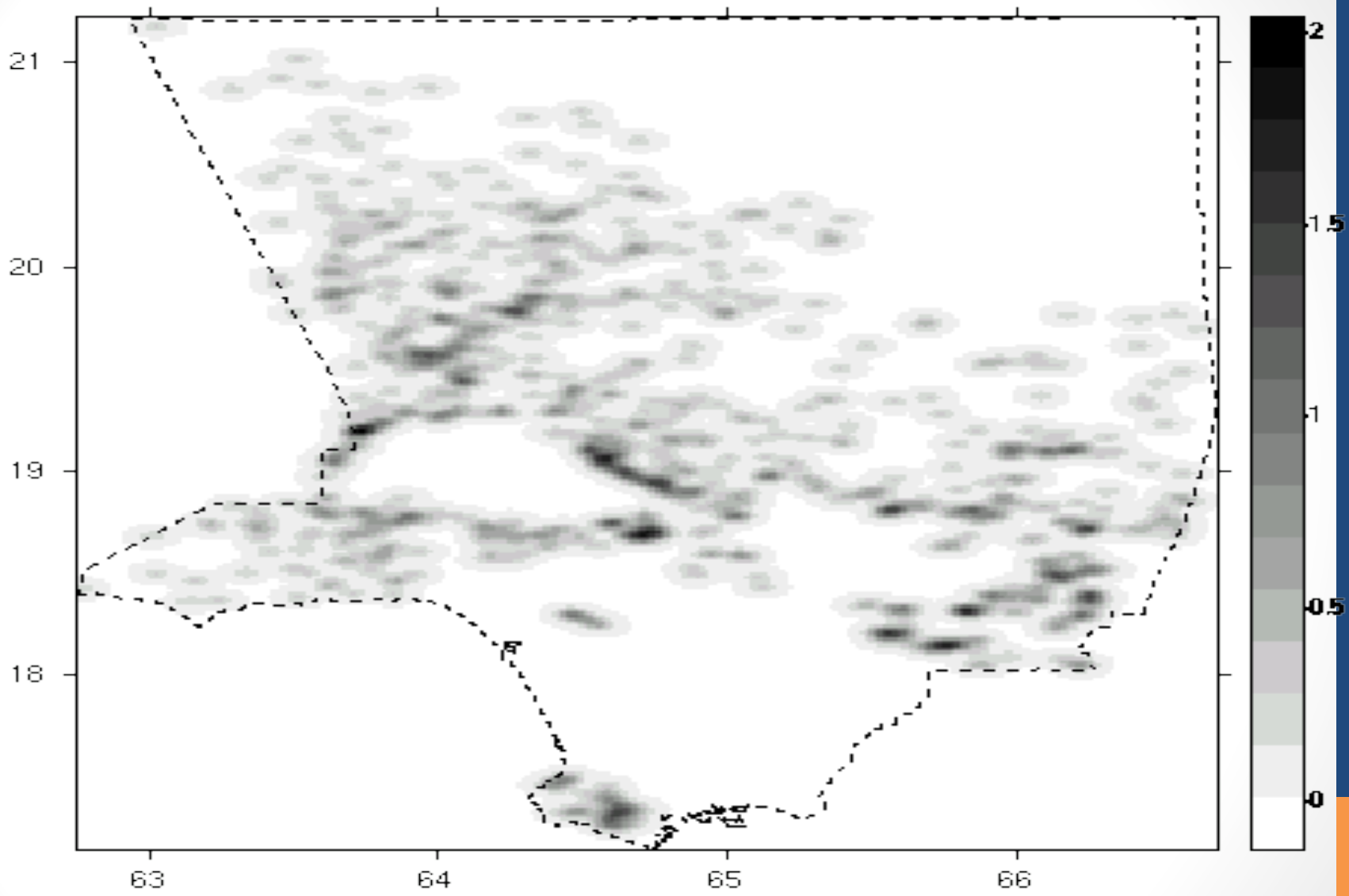

year
1960
1970
1980
1990
2000
2
4
6
8
10
12
month


prototype
year
1960
1970
1980
1990
2000
2
4
6
8
10
12
month

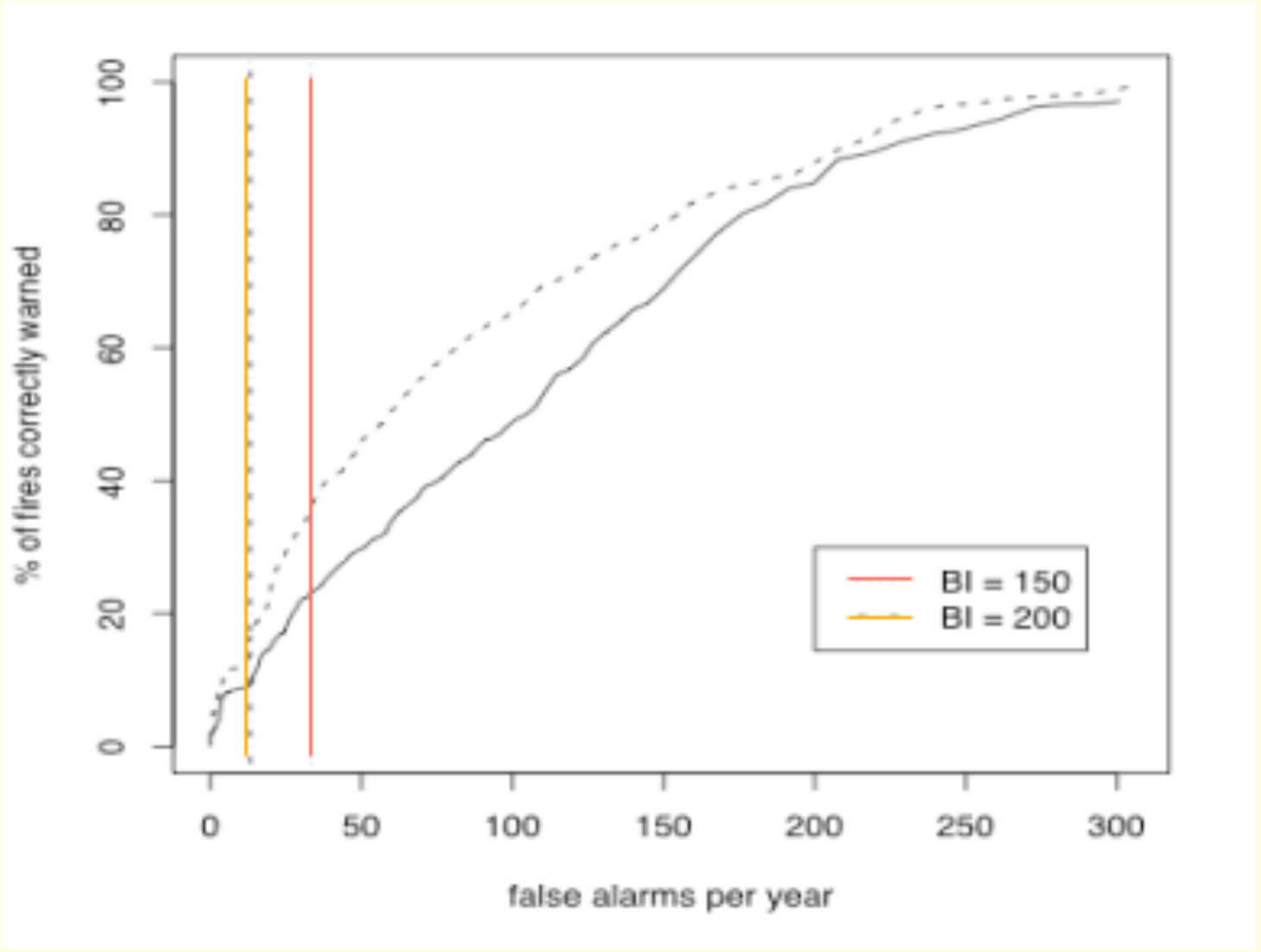

Total area burned vs. average windspeed
r = 0.16
Total area burned (sq m)
0
1
2
3
4
5
0
10
20
30
40
50
60
Average windspeed


Total area burned vs. average max relative humidity
Total area burned
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
sd=12
20
40
60
80
100
Average max relative humidity


Total area burned vs. average max temperature
Total area burned (sq m)
0.05
0.10
0.15
40
60
80
100
Average max temperature (F)


21
20
19
18
63
64
65
66
2
1.5
1
0.5
0

Error diagrams (Molchan 1990, Xu and Schoenberg 2011)



Model Construction

- Wildfire incidence seems roughly multiplicative.
  (only marginally significant in separability test)
- Windspeed. RH, Temp, Precip all matter.
- Tapered Pareto size distribution f, smooth spatial background μ.

[*] $\lambda(t,\mathbf{x},a) =$

$f(a)\,\mu(\mathbf{x})\,\beta_1 \exp(\beta_2 RH + \beta_3 WS)\,(\beta_4 + \beta_5 Temp)(\max\{\beta_6 - \beta_7 Prec, \beta_8\})$

Relative AICs (Model - Constant, so lower is better):

Constant	RH	BI	Model [*]
0	-262.9	-302.7	-601.1

Comparison of Predictive Efficacy

	False alarms per year	% of fires correctly alarmed
BI 150:	32	22.3
Model [*]:	32	34.1
BI 200:	13	8.2
Model [*]:	13	15.1

Forecasting Earthquakes.

The Hawkes process (Hawkes 1971) is a useful form for modeling clustered processes like earthquakes, where

$$\lambda(t,x,y,m) = \mu(x,y,m) + \sum_{i:ti<t} g(t-t_i, x-x_i, y-y_i, m_i).$$

An example is the Epidemic-Type Aftershock Sequence (ETAS) model of Ogata (1988, 1998), which has been used for earthquakes as well as invasive species (Balderama et al. 2012) and crime (Mohler et al. 2011). With ETAS, $\mu(x,y,m) = \mu(x,y)\, f(m)$, and e.g.

$$f(m) \propto \exp\{-\beta(m - m_0)\},$$

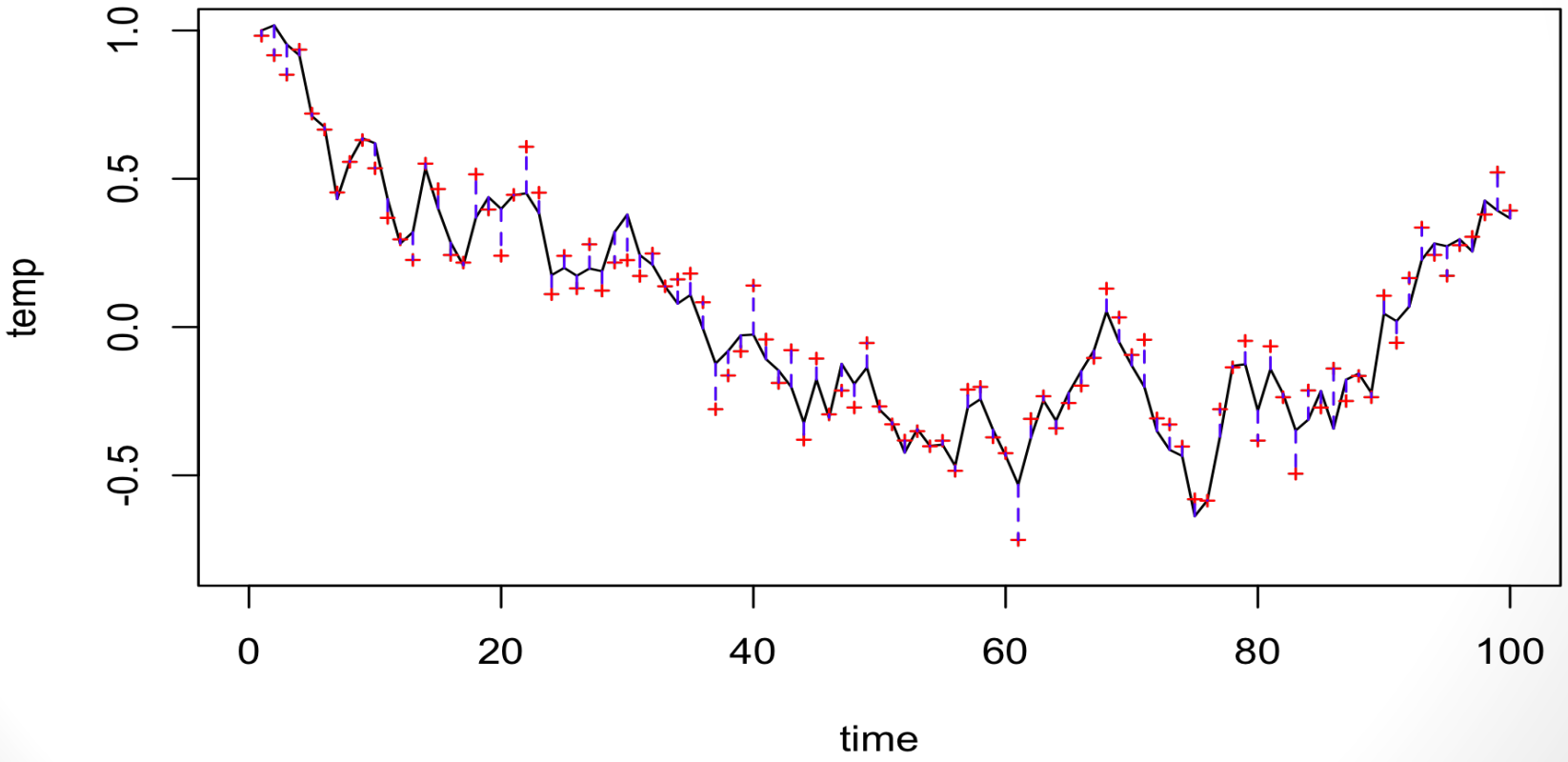
$$g(t,x,y,m) = \frac{K_0 \exp\{\alpha(m - m_0)\}}{(t+c)^p (x^2 + y^2 + d)^q},$$

Difficulty of point process model evaluation.

With most types of data, such as regression style data, time series data, or observations on a spatial-temporal grid, can just look at residuals

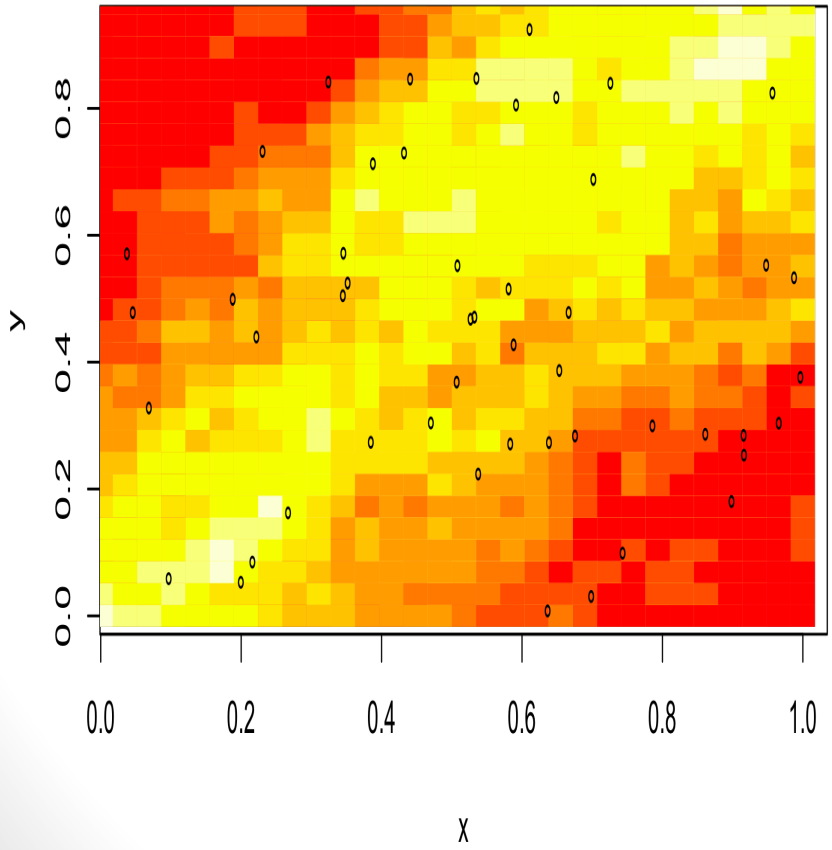
observed – predicted

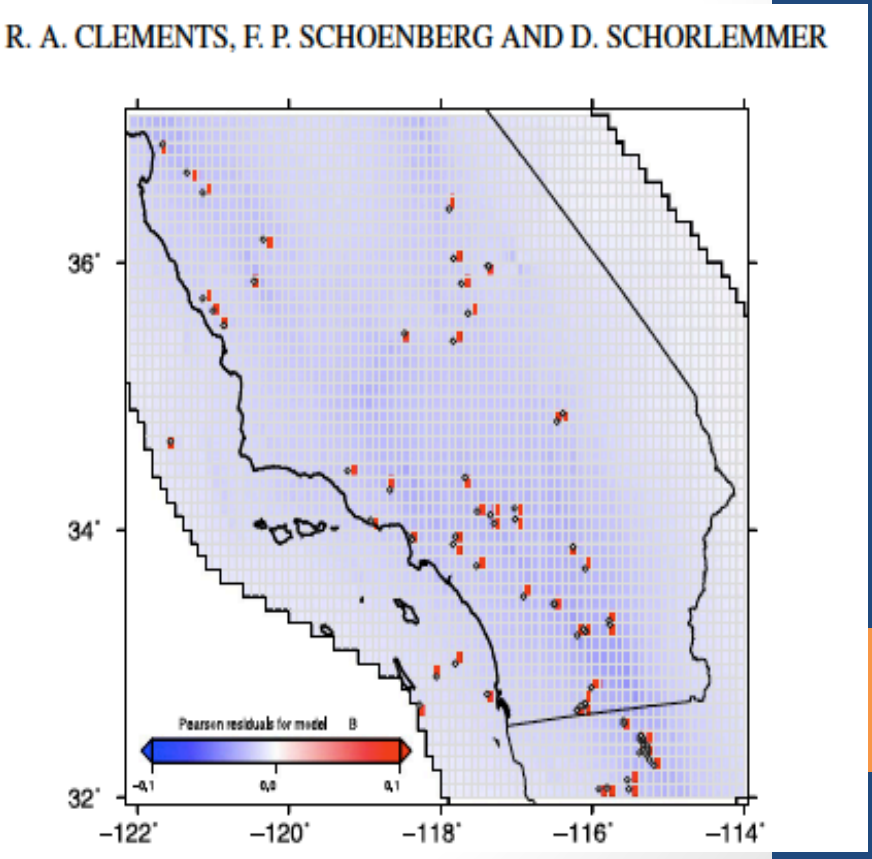
for each observation. Closer to 0 = better.



With point processes, this doesn't work.

With point processes, the observations are a collection of points indicating where and when the phenomenon occurred, and typically the model output is an intensity function
$\lambda(u)$ = anticipated rate of points around spatial-temporal location u.
Can compare number observed and number predicted over each bin, but there are problems.
large pixels yield low power.
with small pixels, residuals are mostly 0 or 1. Highly skewed.
For many models, the residual plot simply looks like a plot of the points themselves.





2. RELM and CSEP.

The Regional Earthquake Likelihood Models (RELM) project [Field (2007)] led to the Collaboratory for the Study of Earthquake Predictability (CSEP) [Jordan (2006)].

RELM tested earthquake forecast models for California.

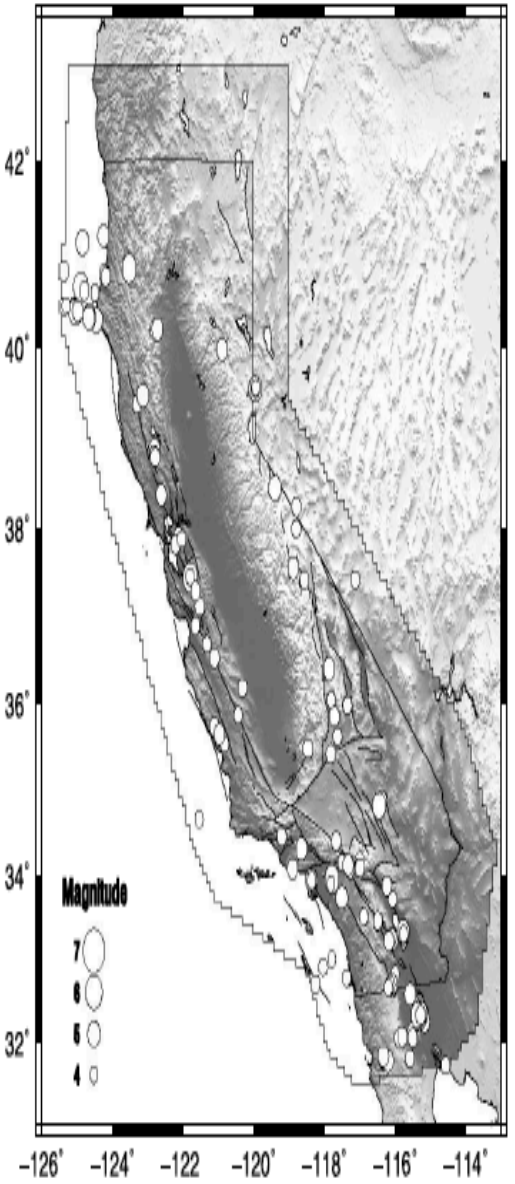
Rigorous, prospective 5 year testing in a dedicated testing center, 1/1/06-1/1/11. [Schorlemmer and Gerstenberger (2007)].

CSEP expanded to regional earthquake forecasts around the world, including California, Japan, New Zealand, Italy, the Northwest Pacific, the Southwest Pacific and the entire globe. Testing centers in Japan, Switzerland, New Zealand and the United States.

CSEP models are five-year, one-day, or recently three-month forecasts. Forecast an expected number of events in each space time magnitude bin. For 1 day models, bins are 0.1◦ lon by 0.1◦ lat by 0.1M from M3.95 to 8.95.

For M8.95-10, one bin of size 0.1◦ by 0.1◦ by 1.05M.

The U.S. testing center is located at the So. California Earthquake Center (SCEC) and hosts forecast experiments for California, the Northwest and Southwest Pacific, and the global experiments.



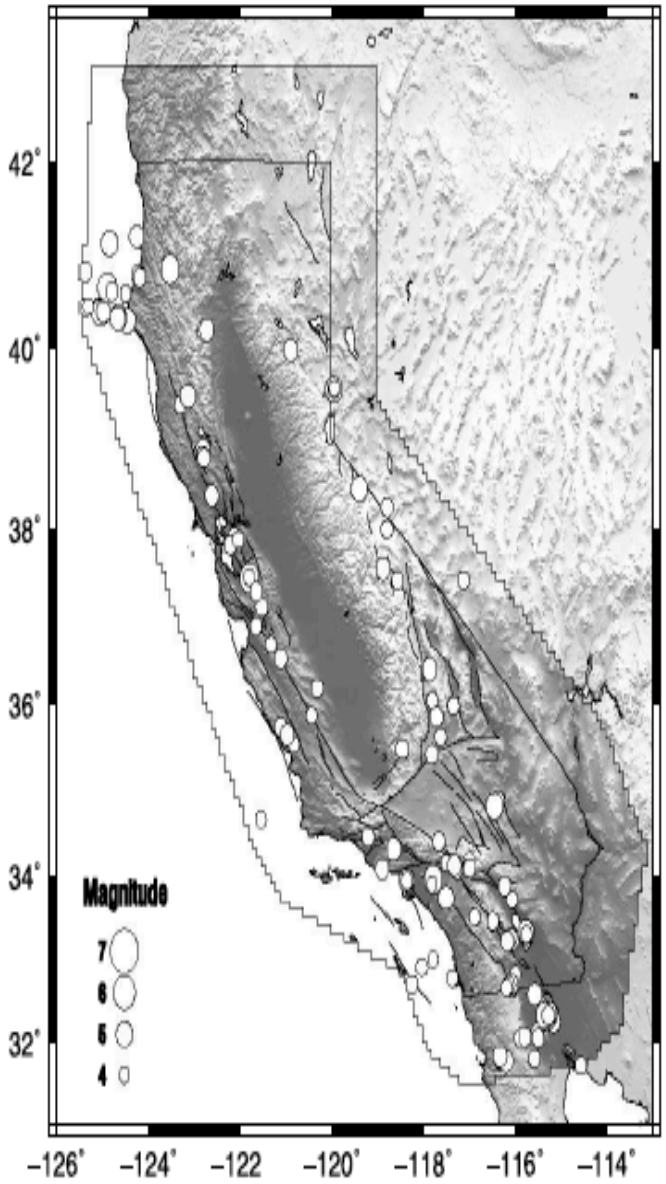
Some models in CSEP.

A. Helmstetter, Kagan and Jackson (2007)
B. Kagan, Jackson and Rong (2007).
C. Shen, Jackson and Kagan (2007).
Epidemic-Type Aftershock Sequence (ETAS) model [Zhuang, Ogata and Vere-Jones (2004), Ogata and Zhuang (2006)].
Short-Term Earthquake Probabilities (STEP) model [Gerstenberger et al. (2005)].

All based exclusively on previous seismicity except C, which uses geodetic and geological info.

Earthquake catalogs were obtained from the Advanced National Seismic System (ANSS).

142 shallow earthquakes with a M≥3.95 occurred in RELM's spatial temporal window.
Note that each RELM model does not necessarily produce a forecasted seismicity rate for every pixel in the space–time region.

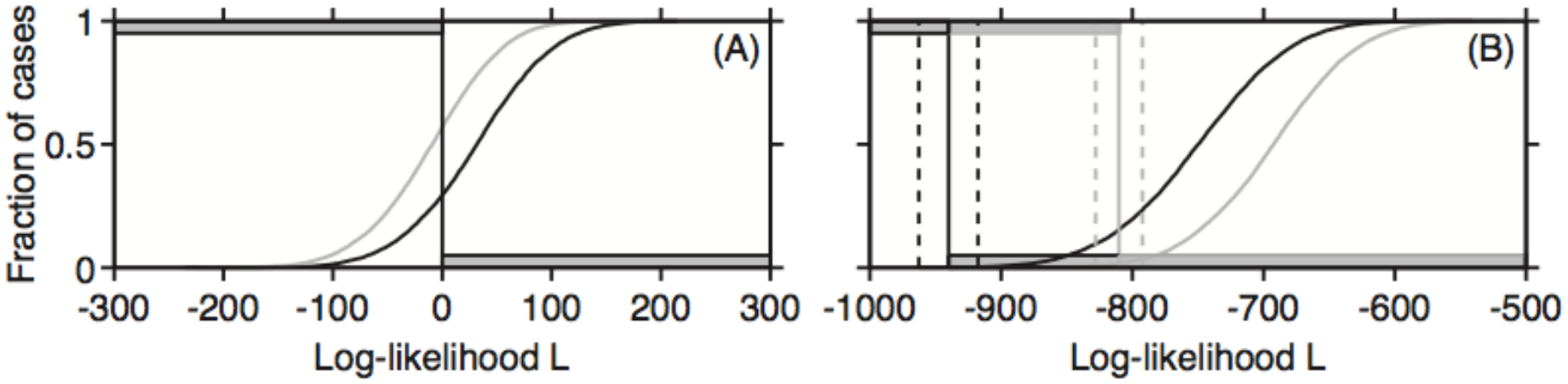


Numerical summaries.

In RELM, consensus was reached that all models would be tested using a certain suite of numerical tests [Jackson and Kagan (1999), Schorlemmer et al. (2007)].

N-Test compares total number of earthquakes for model and observations.
L-Test compares likelihood.
R-Test compares likelihood ratio for two models.

L and N test can be used to see if discrepancies for 1 model are stat. significant.
Simulate, and see if the likelihood for the data is in the middle 95% range.



Schorlemmer et al. 2007

Some drawbacks of these tests were discovered (Schorlemmer et al. 2010).

The N-test and L-test have low power and are typically unable to discern significant lack of fit unless the overall rate of the model fits very poorly.

Further, even when the tests do reject a model, they do not typically indicate where or when the model fits poorly, or how it could be improved.

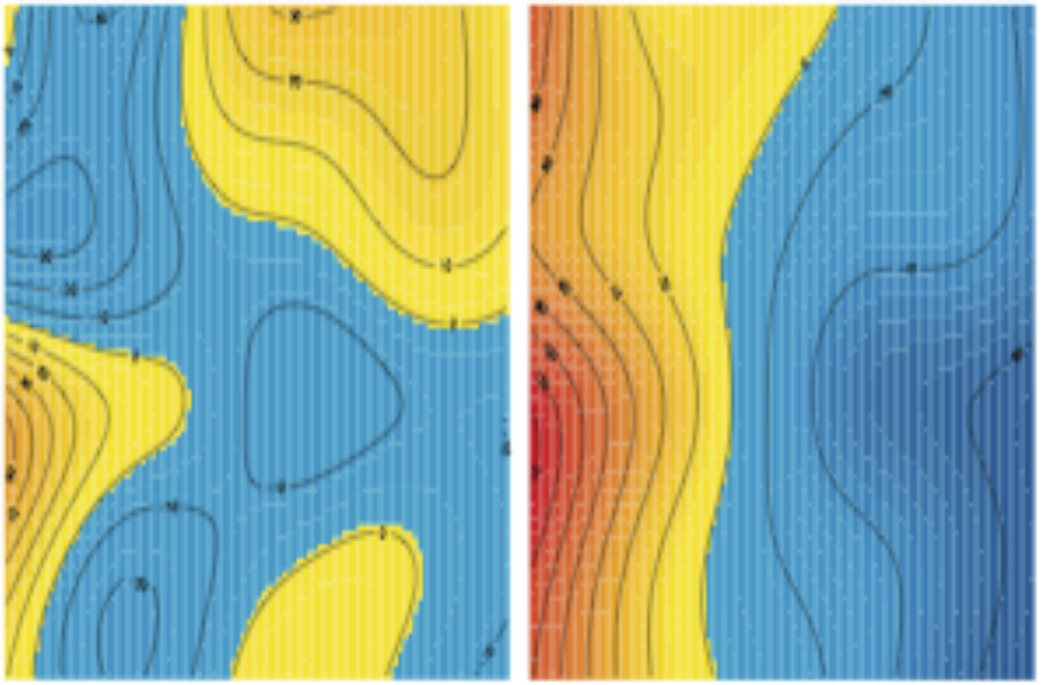
Model	ℓ_{obs}	γ	N_{obs}	δ
		Mainshock+Aftershock		
A. Helmstetter	−22881.46	0.000	142	**0.000**
B. Kagan	−10765.43	0.008	81	**0.001**
C. Shen	−10265.20	0.002	86	**0.043**
		Daily		
ETAS	−387.69	1.00	85	**0.00**
STEP	−50.43	0.00	83	**0.99**

Clements et al. 2011

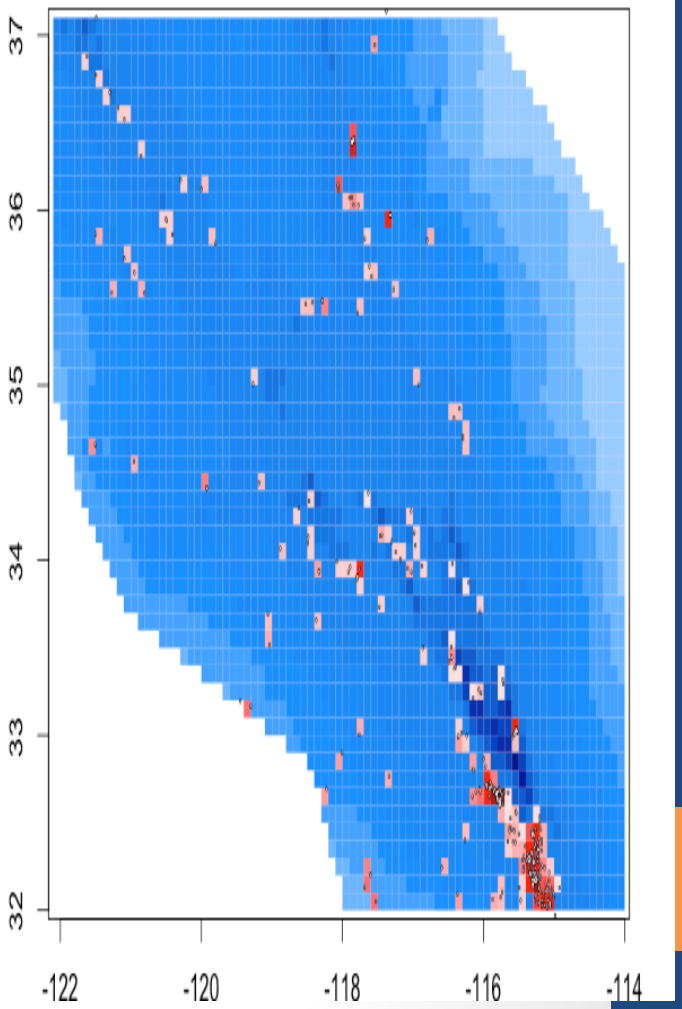
Pixel-based residuals.
Compare $N(A_i)$ with $\int_A \lambda(t, x)\, dt\, dx$, on pixels A_i.
(Baddeley, Turner, Møller, Hazelton, 2005)

Problems:

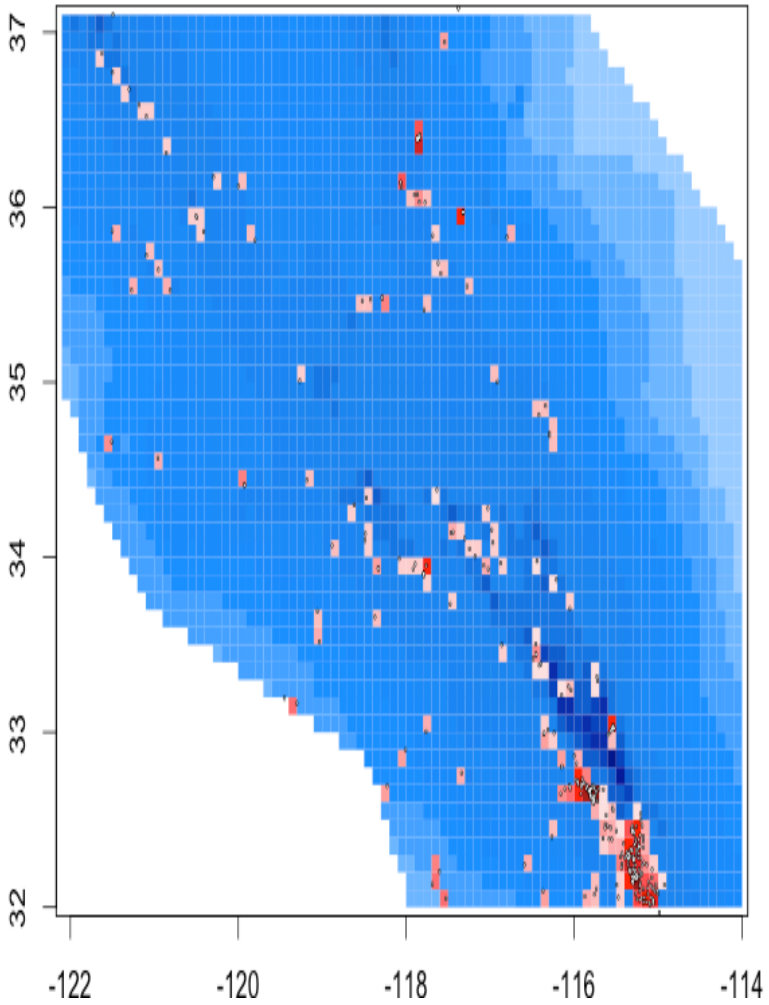
* If pixels are large, lose power.
* If pixels are small, residuals are mostly ~ 0,1.
  -- non-normality after standardization.
* Smoothing reveals only gross features.



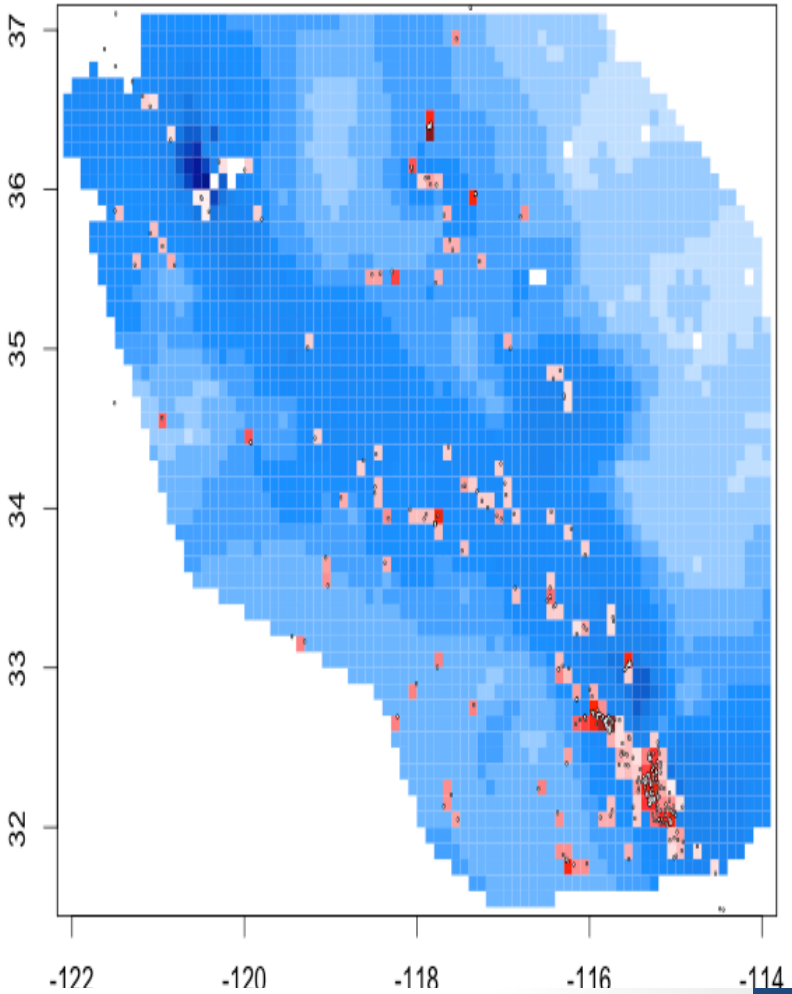
Pearson residuals for Model B:



Pearson residuals for Model B:



Pearson residuals for Model C:



How can you see how well the model fits?

a) Deviance residuals

b) Superthinned residuals

c) Voronoi residuals.

-- Given two competing models, can consider the difference between residuals, number of observed fires – number expected, over each pixel.

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?

-- Better: consider difference between *log-likelihoods*, in each pixel. The result may be called ***deviance residuals*** (Clements et al. 2011), ~ resids from gen. linear models.

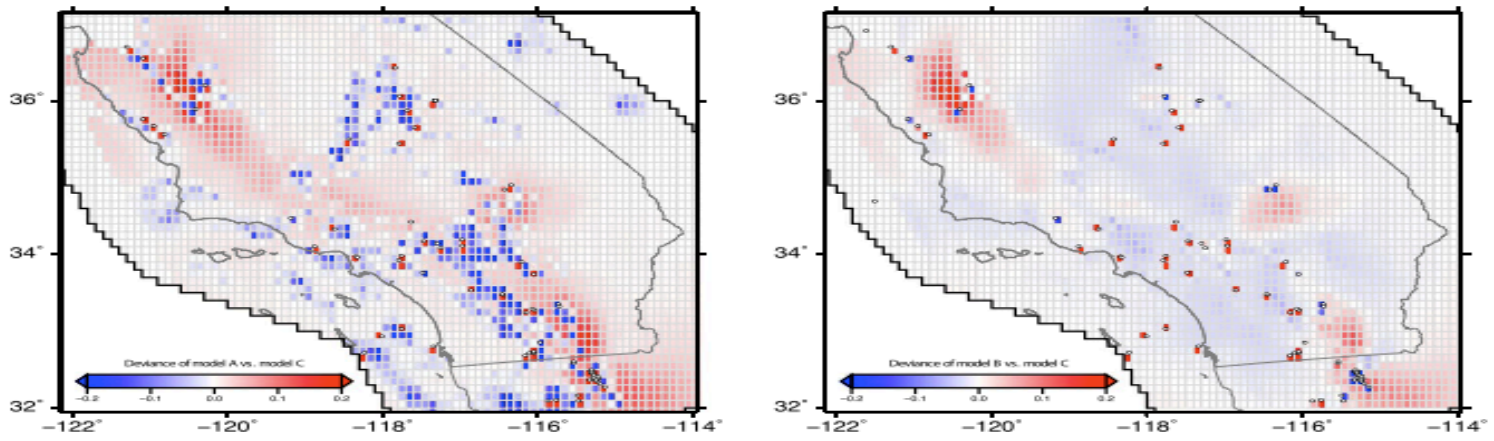


FIG. 4. *Left panel* (a): *deviance residuals for model A versus C. Sum of deviance residuals is* 86.427. *Right panel* (b): *deviance residuals for model B versus C. Sum of deviance residuals is* −7.468.

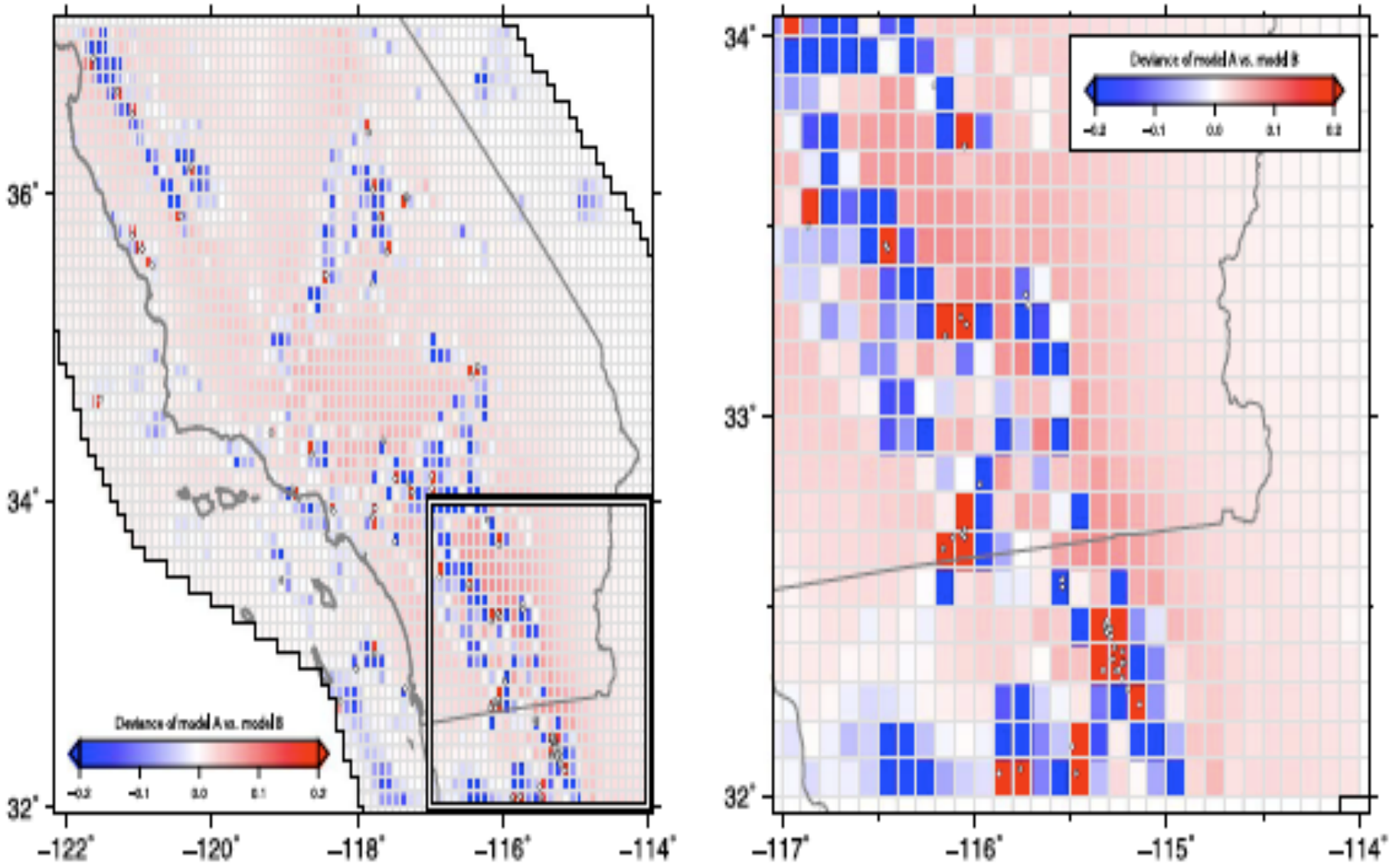


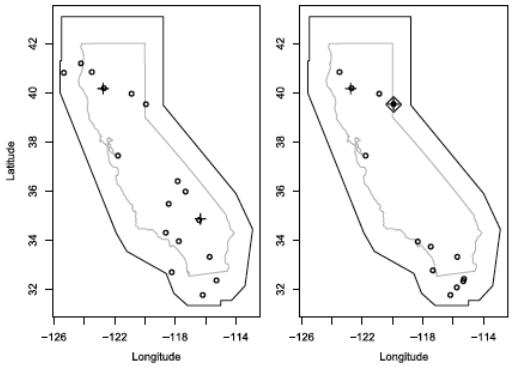
FIG. 3. *Left panel* (a): *deviance residuals for model A versus B. Sum of deviance residuals is* 84.393. *Right panel* (b): *close-up of deviance residuals for model A versus B near the Imperial fault.*

Superthinning (Clements et al., 2012)

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

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



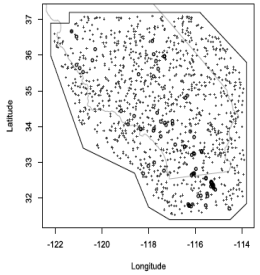


FIG. 9. *Superposed residuals for model C. Simulated points make up 90.7% of all points.*

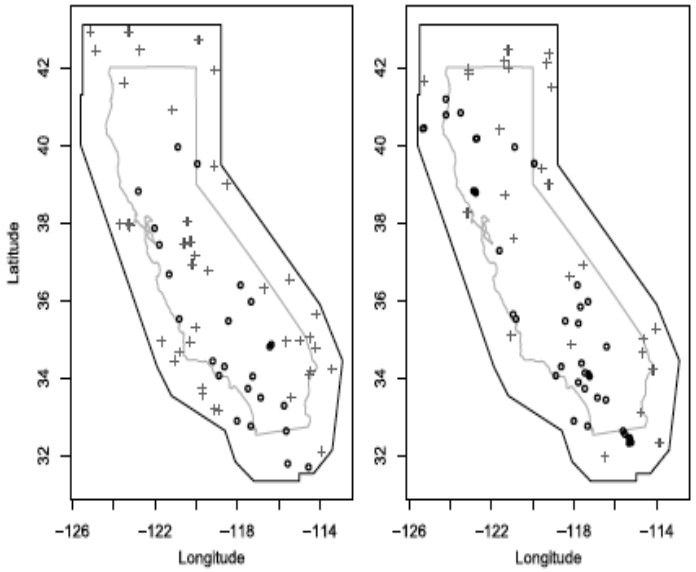
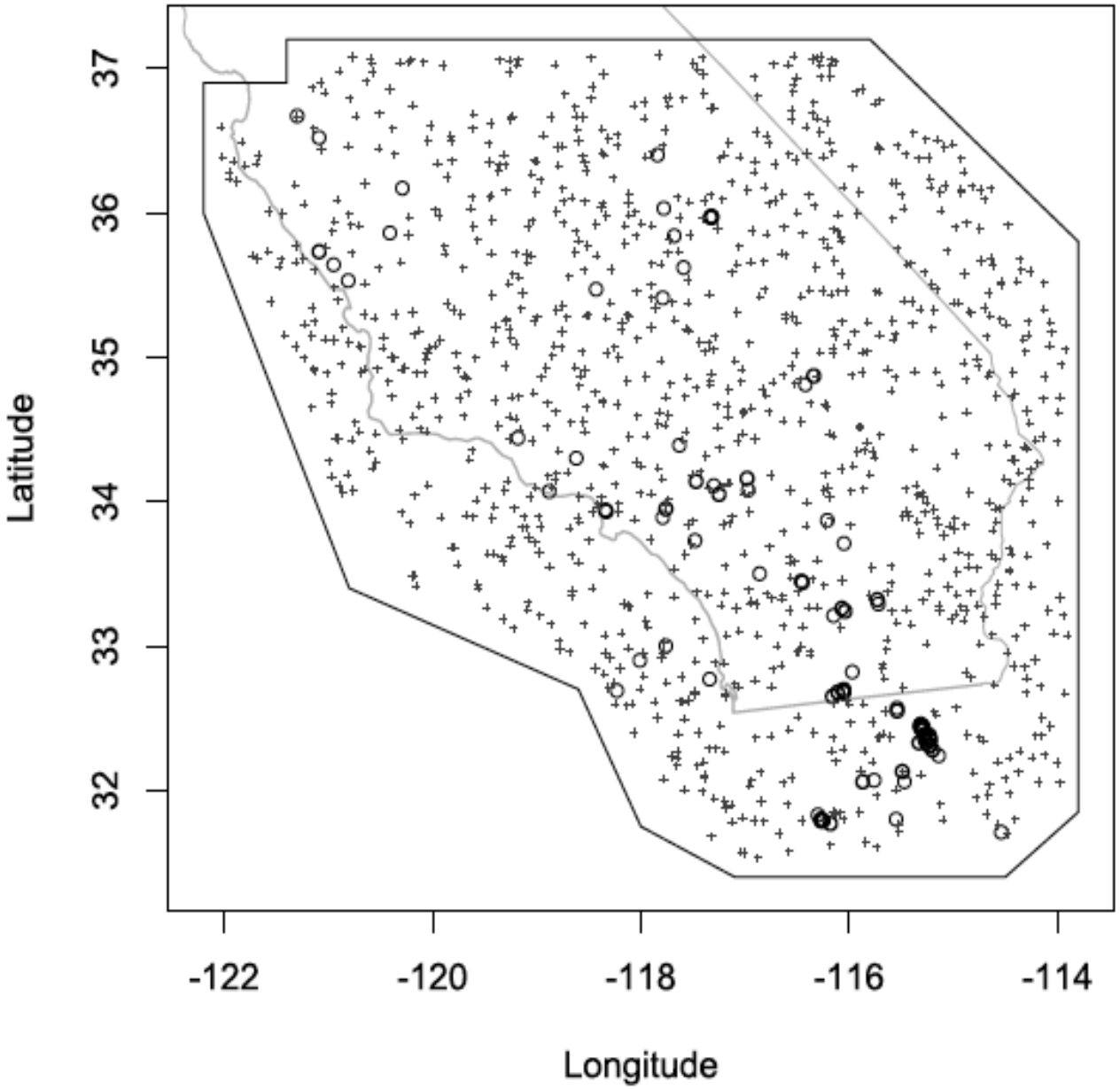


FIG. 11. *One realization of super-thinned residuals for the five models considered (circles = observed earthquakes; plus signs = simulated points). Top-left panel* (a): *model A* ($k = 2.76$). *Top-center panel* (b): *model B* ($k = 2.95$). *Top-right panel* (c): *model C* ($k = 2.73$). *Bottom-left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



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

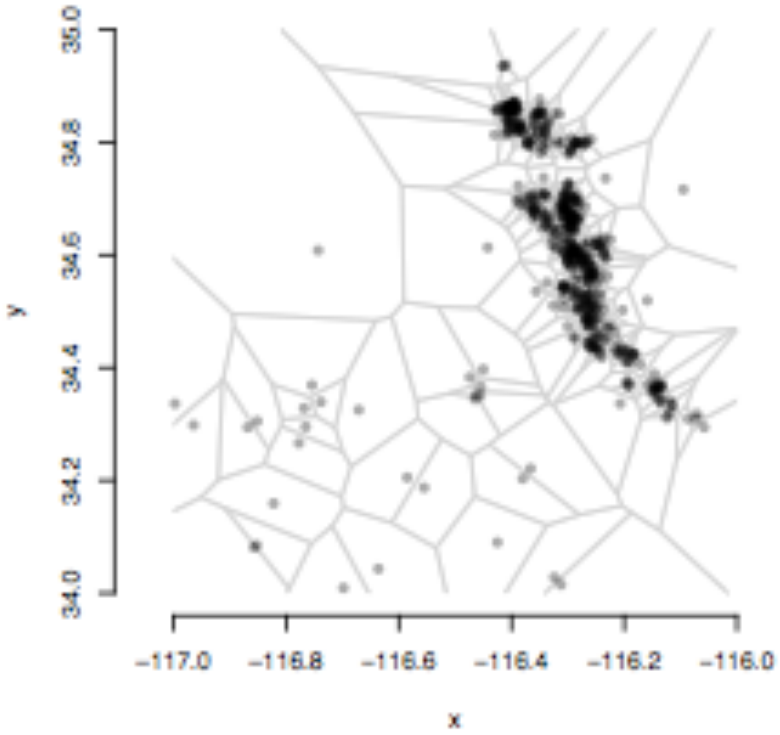
c) Voronoi residuals (Bray et al. 2013)

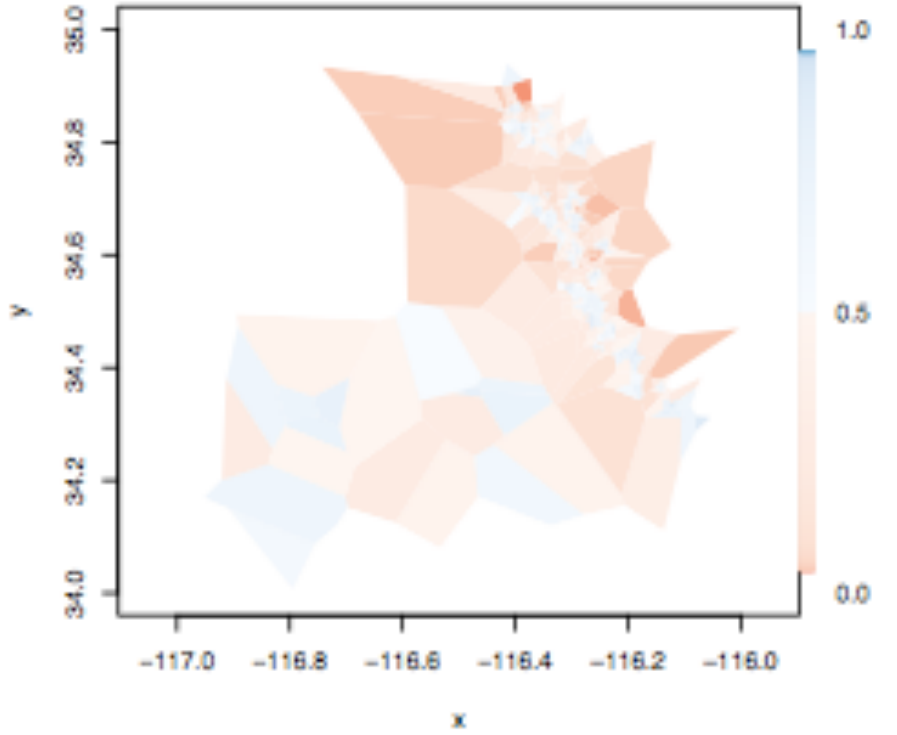
A Voronoi tessellation divides a space into cells Ci where Ci contains all locations closer to event i than any other observed event.

Within each cell, calculate residuals

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

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



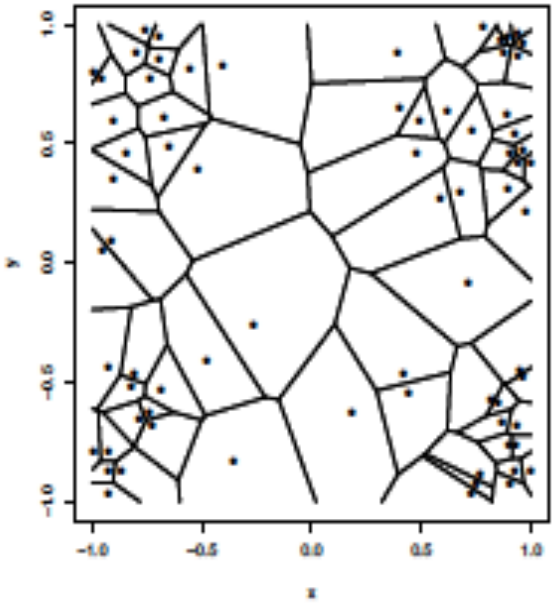


spatially adaptive and nonparametric.

Voronoi Residuals

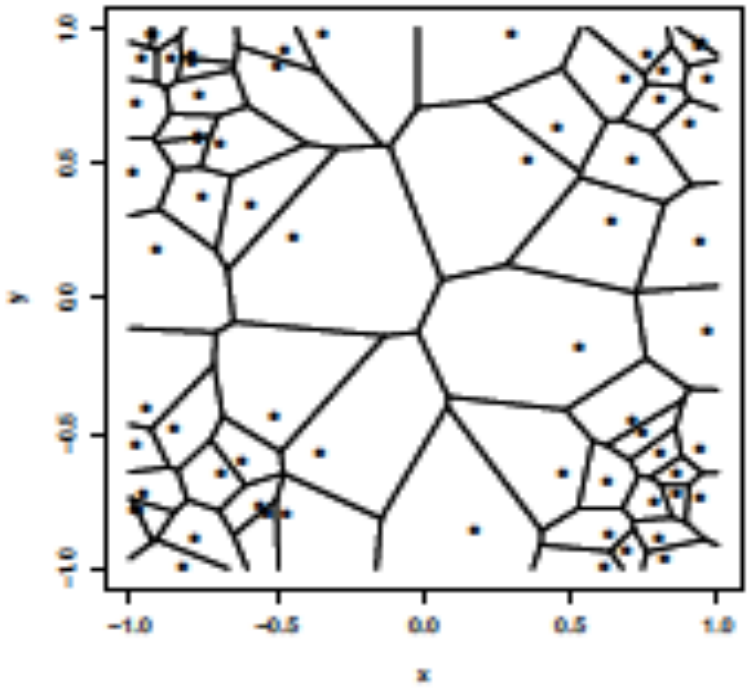
Voronoi Tessellation: (Okabe, 2000) a partitioning of S into n convex polygons (tiles)

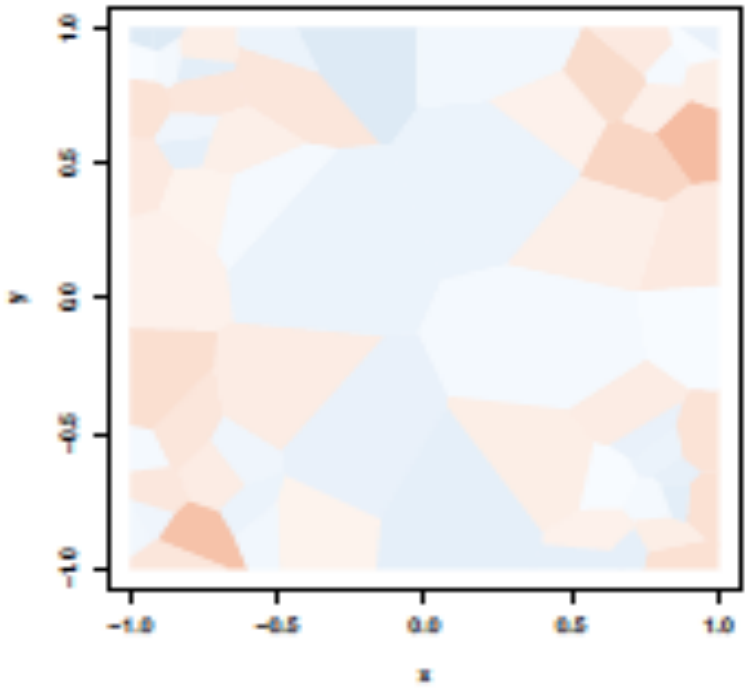
$$D_i = \{x \in X : ||x - x_i|| \leq ||x - x_j||, \forall j \neq i\}$$



Voronoi Residual:

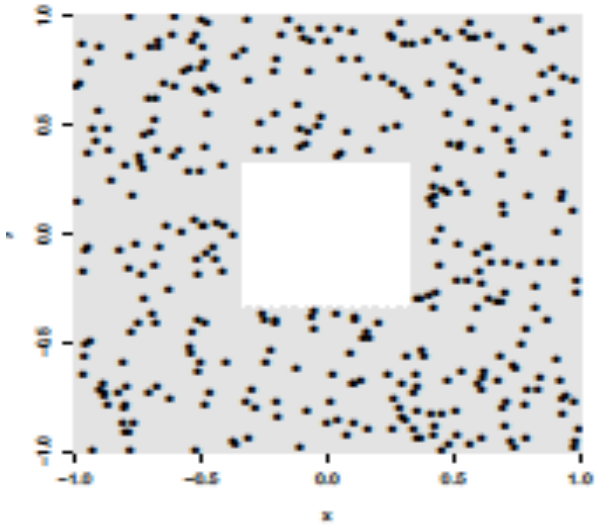
$$\hat{r}_i = \frac{N(D_i) - \int_{D_i} \hat{\lambda}(x_i)\,\mathrm{d}x_i}{SE\left(\int_{D_i} \hat{\lambda}(x_i)\,\mathrm{d}x_i\right)} = \frac{1 - |D_i|\bar{\lambda}}{SE\left(|D_i|\bar{\lambda}\right)}$$

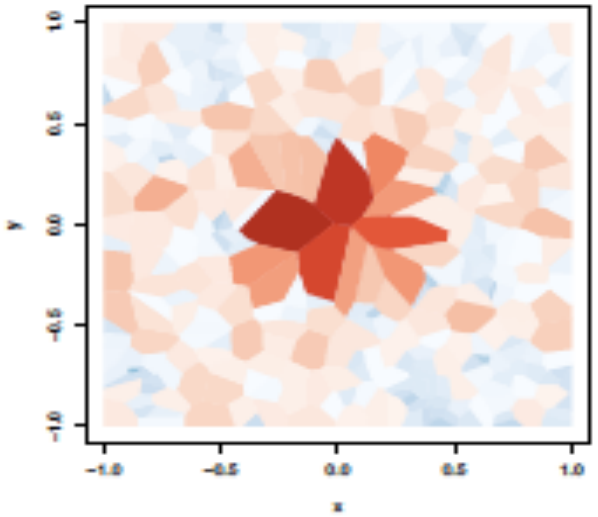




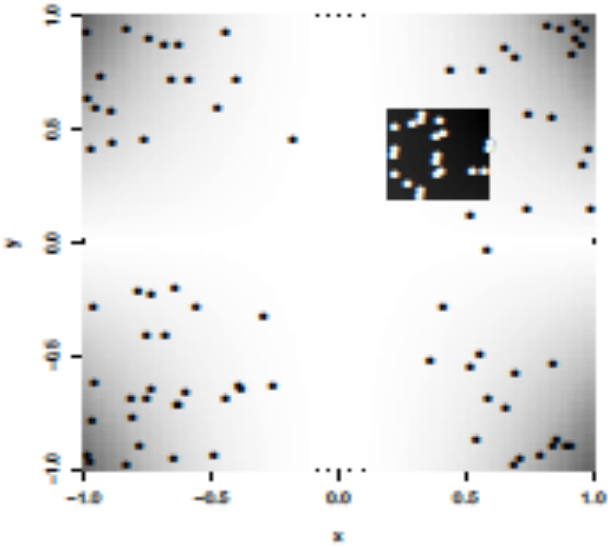
- generating model = fitted model = $100x^2|y|$
- less-skewed residuals with narrower range

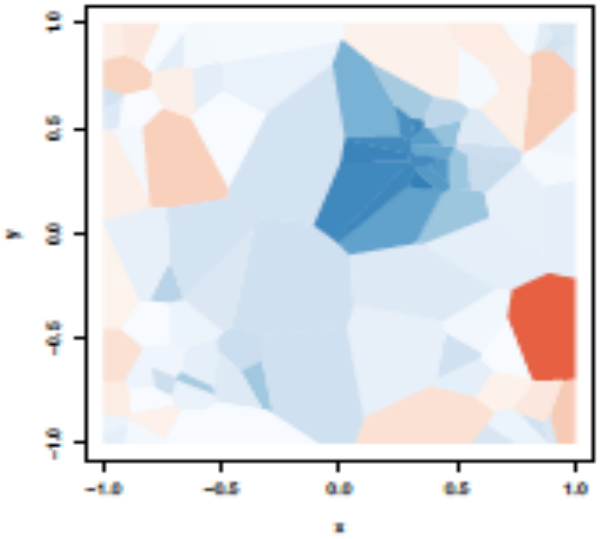
overprediction

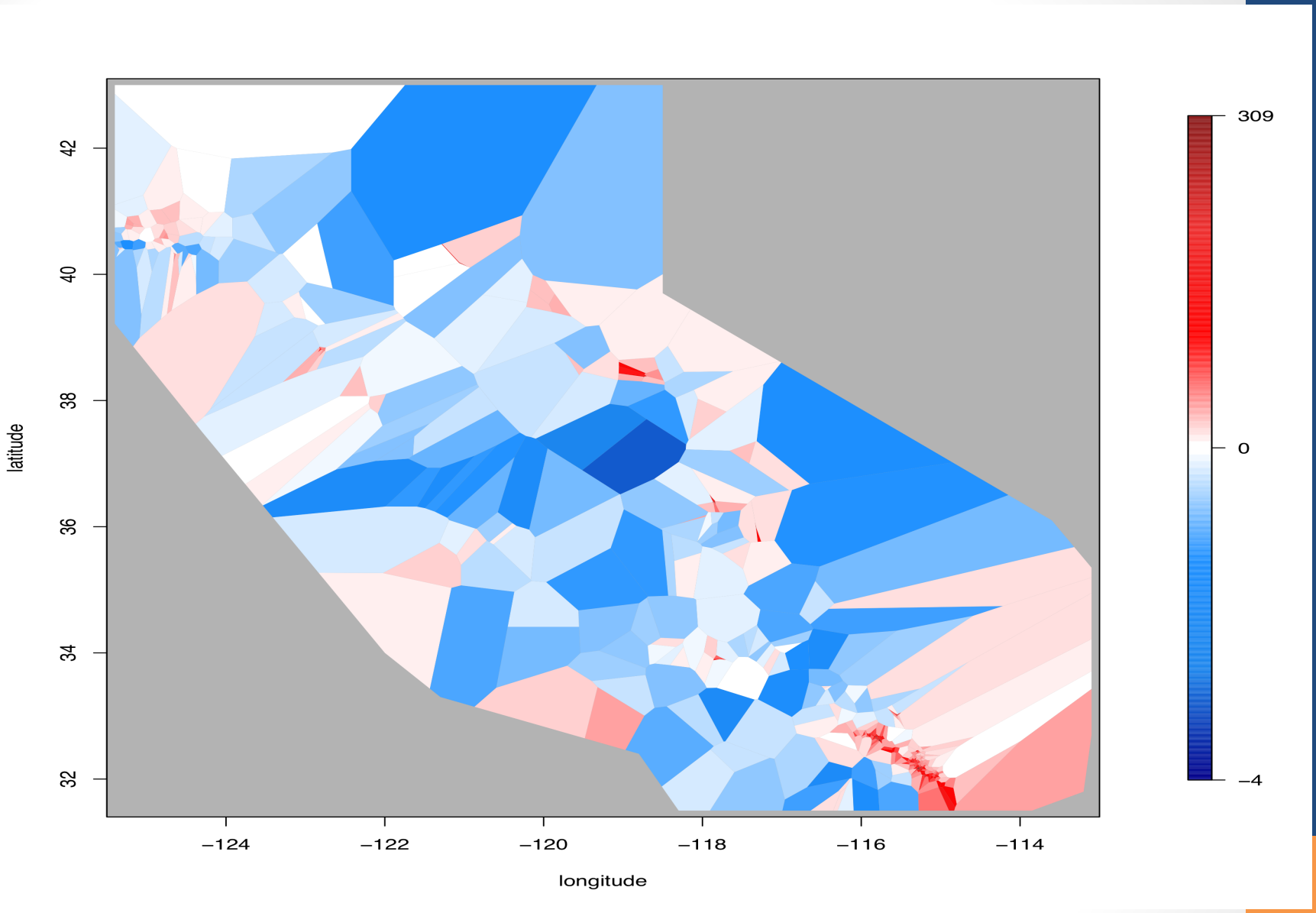




underprediction

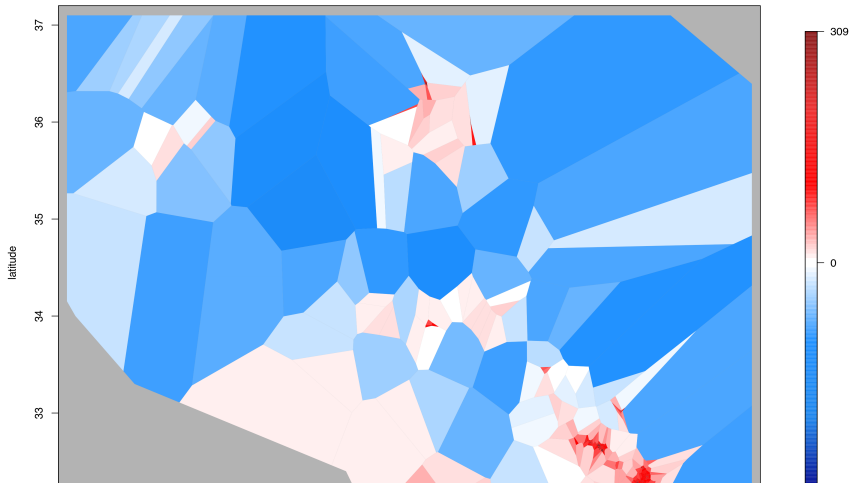




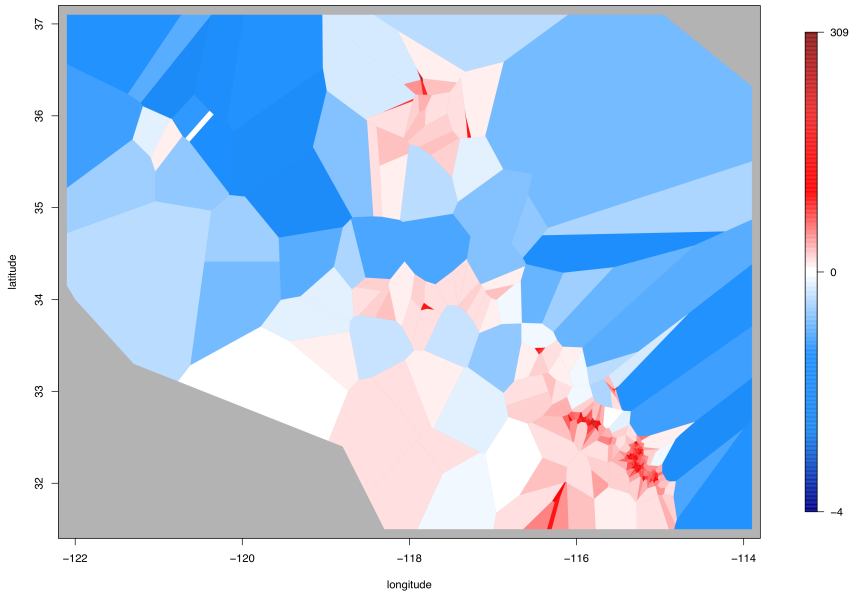

latitude
42
40
38
36
34
32
−124
−122
−120
−118
−116
−114
longitude
309
0
−4

Model A

Model B



Model C



Global temperature data. Hawkins and Sutton 2017.

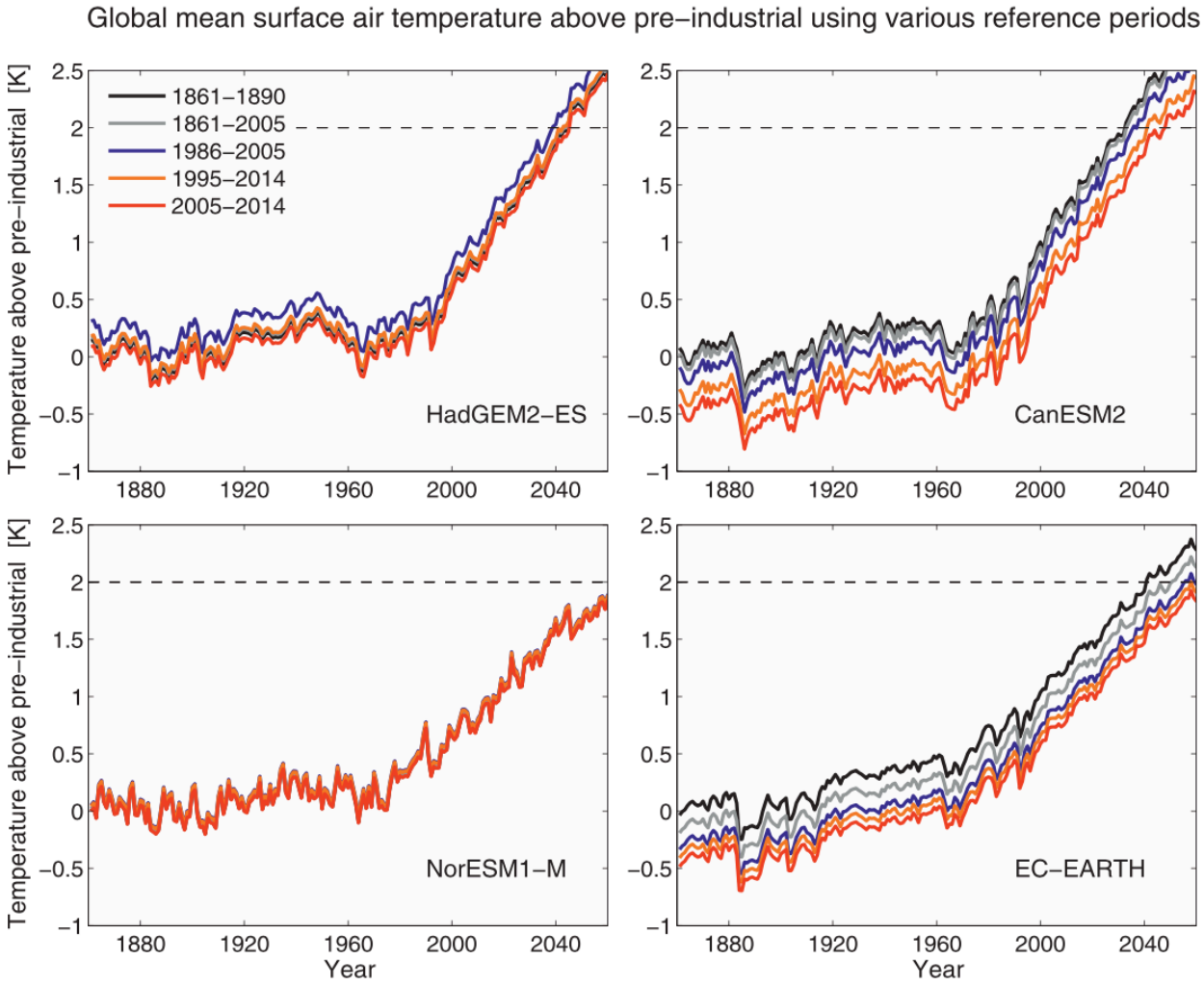


FIG. 5. Comparing projections (RCP4.5) using five different reference periods for four example GCMs, showing the ensemble means. Note that the observed change is used as the anomaly from 1850 to 1900 to the chosen reference period. Most of the CMIP5 GCMs behave more like CanESM2 and EC-EARTH than NorESM1-M.

Global temperature data. Hawkins and Sutton 2017.

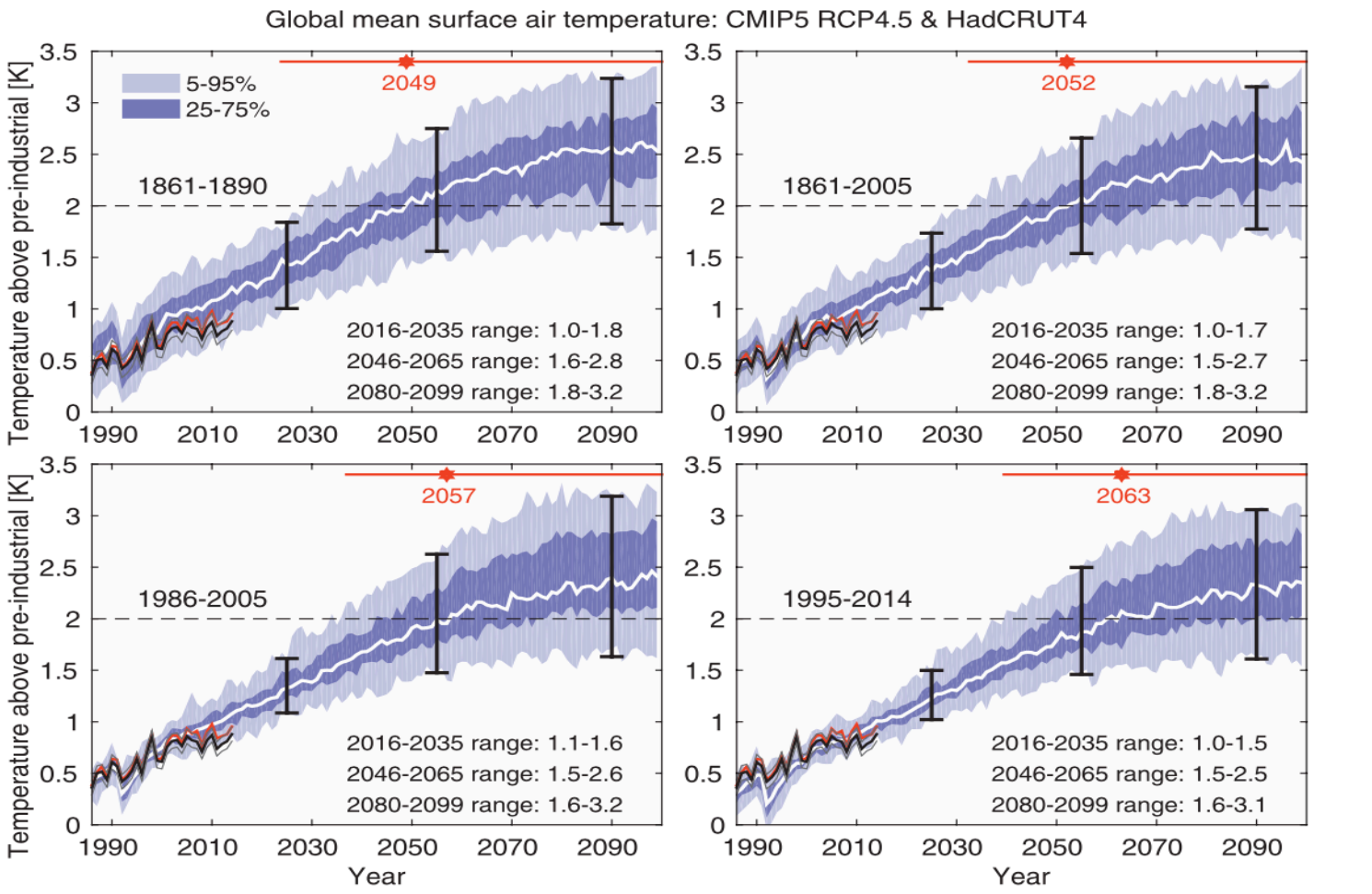


FIG. 7. Comparing CMIP5 projections (RCP4.5, 42 models) for the future using four different reference periods. Red stars indicate projected time of crossing 2 K above preindustrial (defined as 1850–1900) with red bars representing the 5%–95% range. Black error bars show 5%–95% temperature ranges for defined periods as indicated in legend. Observations (black) and observational uncertainties (gray) shown for HadCRUT4.3 (Morice et al. 2012). The CW14 observations are shown in red.

Global temperature data. Kay 2015.

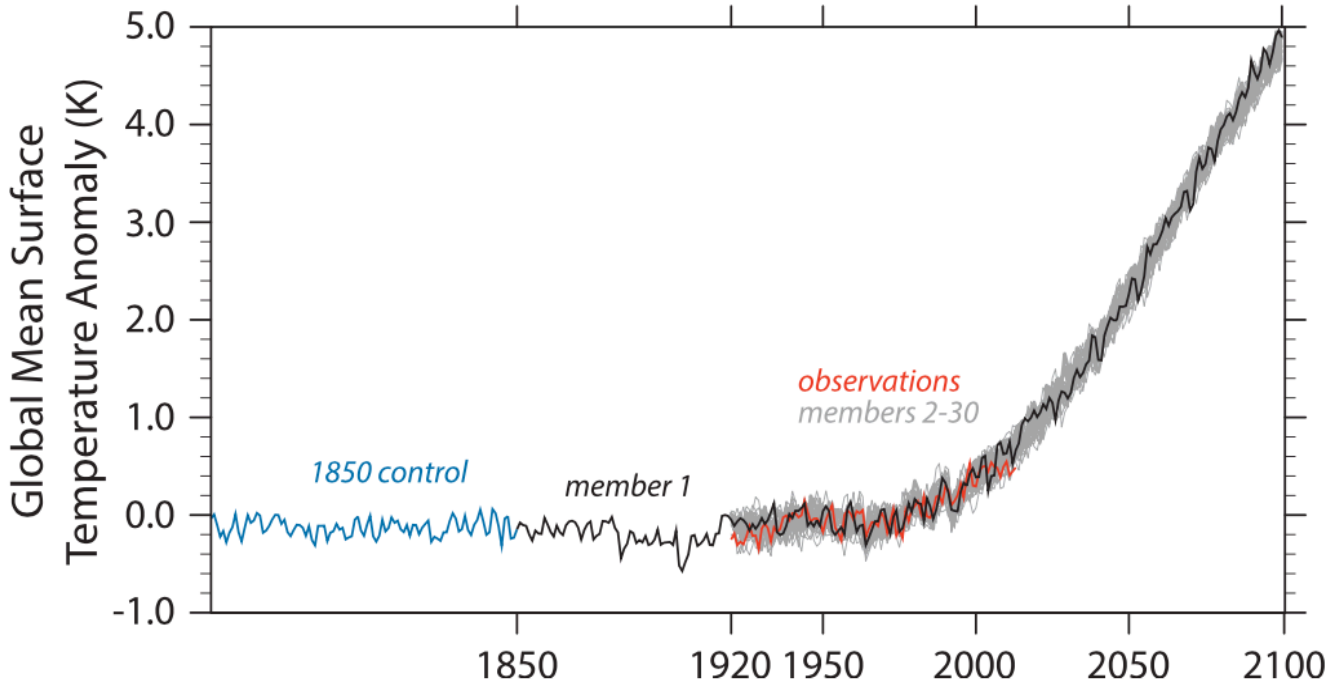
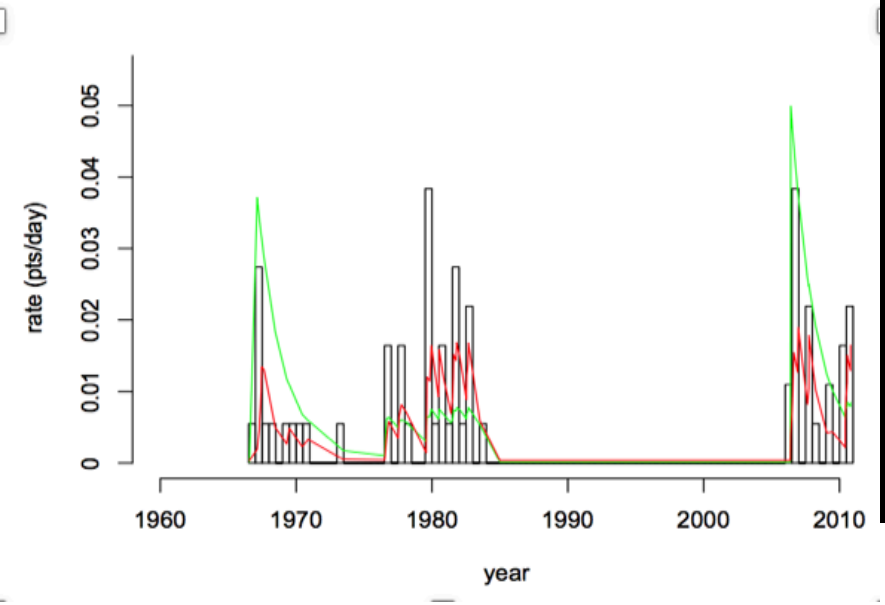


FIG. 2. Global surface temperature anomaly (1961–90 base period) for the 1850 control, individual ensemble members, and observations (HadCRUT4; Morice et al. 2012).

AMERICAN METEOROLOGICAL SOCIETY

Forecasting the spread of infectious diseases.



* Epidemics have traditionally been modeled by grid based compartmental models like SIR and SEIR models (e.g. Althaus 2014).

* More recently, point process models have been used, especially Hawkes models, offering greater forecast precision (Law et al. 2009, Chaffee 2017).

* With Hawkes processes, the productivity or expected number of disease transmissions triggered directly by a given infected persion, is static. In the case of Hawkes models applied to earthquakes (e.g. Ogata 1988, Ogata 1998), the basic Hawkes model was extended to allow the productivity of an earthquake to depend on its magnitude, but still not to depend on the time or location of the event, nor on the number of previously occurring events.

* With diseases, one may want the productivity to vary.

Hawkes models and ETAS.

Probabilistic models for point processes typically involve modeling the conditional rate

$\lambda(t,x,y,m)$ = expected rate of accumulation of points at time t, location (x,y), and magnitude m, given the history of all previous events.

Hawkes (1971) modeled $\lambda(t) = \mu + K\sum_{i:ti<t} g(t-t_i)$.

$\mu \geq 0$ is the background rate, K is the productivity, $0 \leq K \leq 1$, and g is the triggering density satisfying $\int_0^\infty g(u)du \doteq 1$

Ogata (1988) proposed the Epidemic-Type Aftershock Sequence (ETAS) model, which is like a Hawkes model but where the productivity can depend on magnitude.

$\lambda(t) = \mu + K\sum_{i:ti<t} g(t-t_i; m_i)$, with $g(u; m_i) = (u+c)^{-p} \exp\{a(m_i-M_0)\}$.

When the prevalence of a disease is low in a region, as is the case when the epidemic has never struck before or has not struck in considerable time, then the conditional intensity λ is small and one would expect the rate of transmission for each infected person to be quite high. A carrier of the disease may be expected to infect many others.

When the epidemic is at its peak and many subjects have contracted the disease, on the other hand, λ is large and one might expect the rate of transmission to be lower due to human efforts at containment and intervention of the disease, and because many subjects may have already been exposed and thus might be recovered and immune to further infection, or deceased.

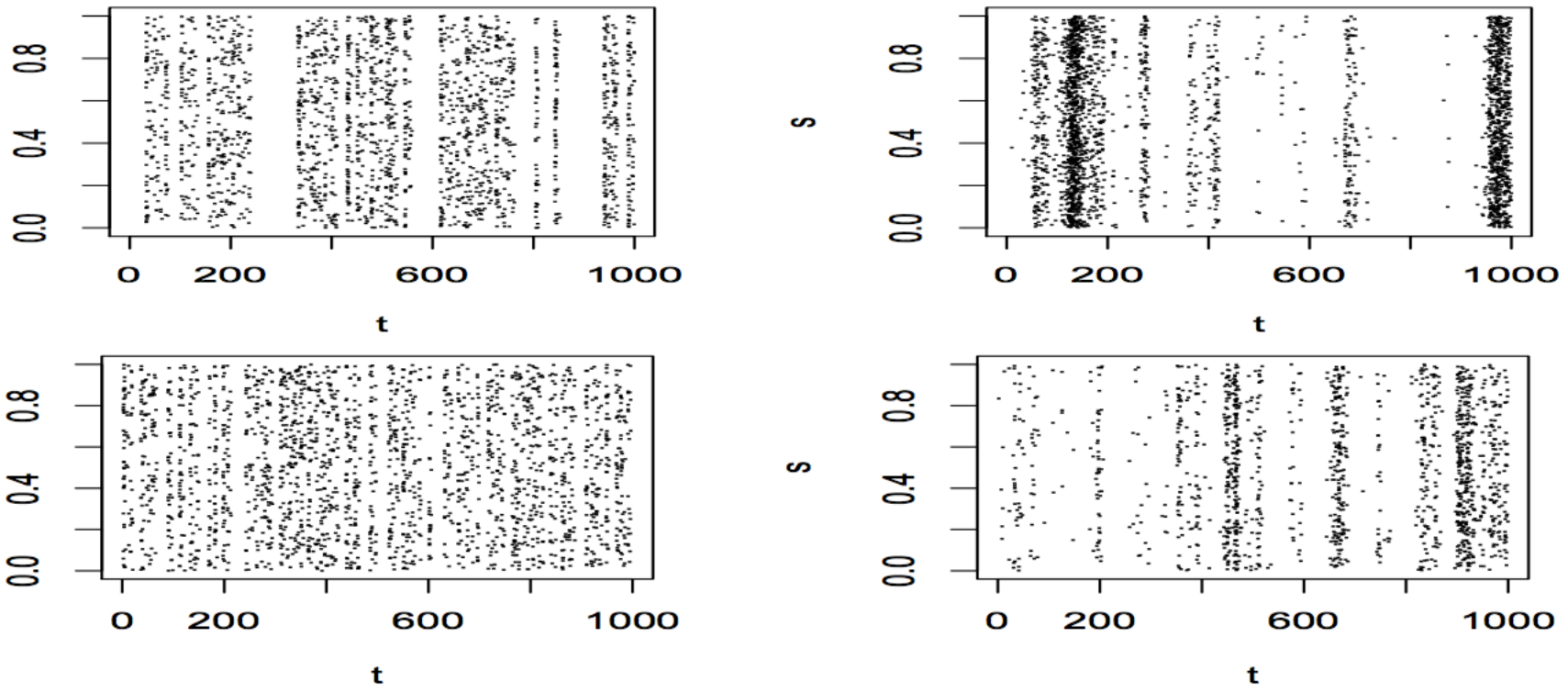
This suggests a model where the productivity for a subject infected at time t is inversely related to the conditional intensity at time t. Since the conditional intensity in turn depends critically on this productivity, we call the model *recursive*.

We may write this model $\lambda(t) = \mu + \int_0^t H(\lambda_{t'})\, g(t - t')\, dN(t')$, where $\mu>0$, $g>0$ is a density function, and $\lambda_{t'}$ means $\lambda(t')$.

We focus in particular in what follows on the case where $H(x) = k\, x^{-\alpha}$, so that $\lambda(t) = \mu + \kappa \int_0^t \lambda_{t'}^{-\alpha}\, g(t - t')\, dN(t')$.

We will refer to the special case where $\alpha = 1$, i.e. where $\lambda(t) = \mu + \kappa \int_0^t g(t - t') / \lambda_{t'}\, dN(t')$, as the *standard recursive model.*

The triggering density g may be given e.g. by an exponential density, $g(u) = \beta \exp(-\beta u)$.

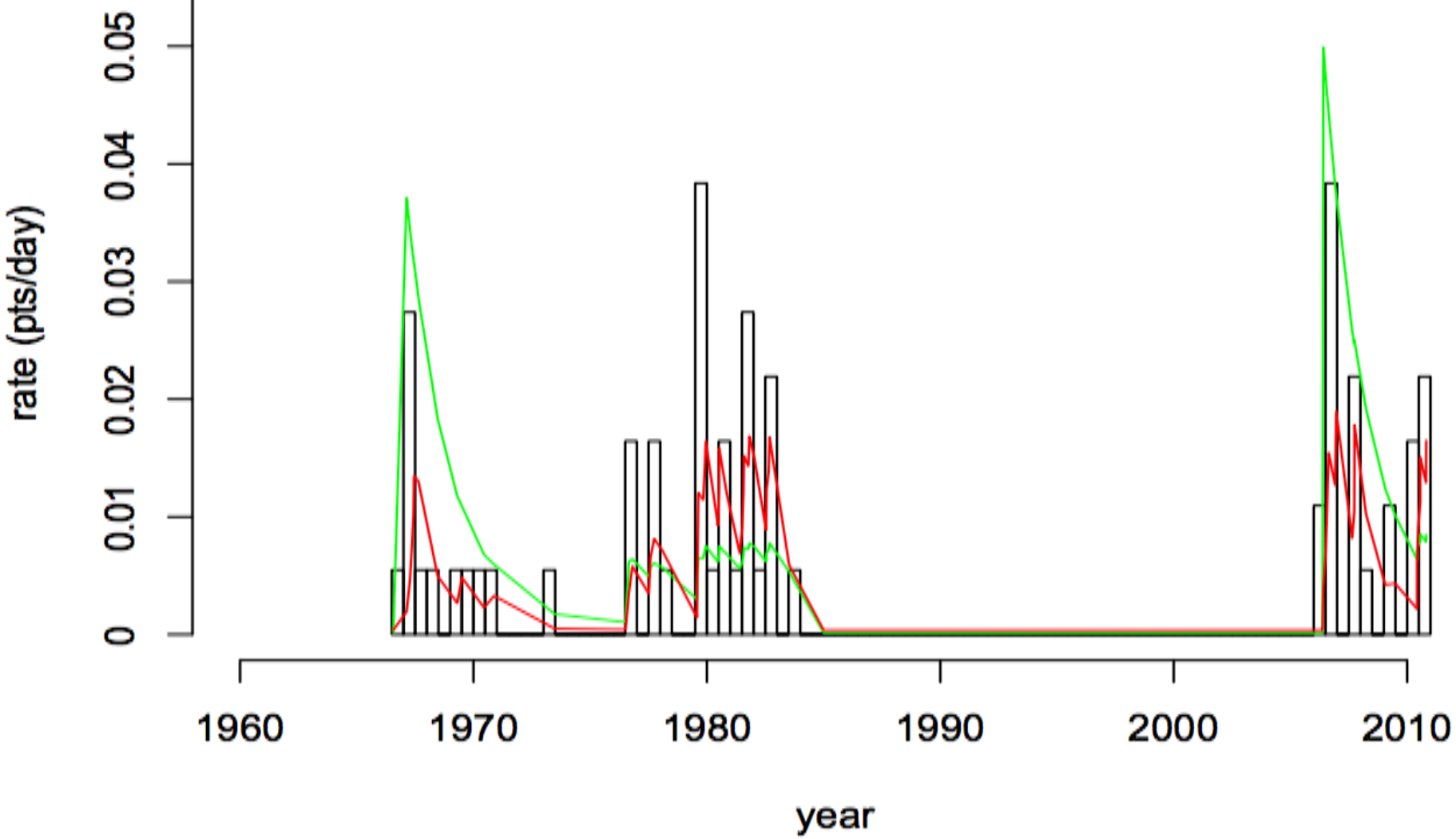


(a) Simulation of a std. recursive model with $\mu = 0.05$, $\kappa = 2$, and g exponential with $\beta = 0.8$.
(b) Simulation of a Hawkes model with the same g and μ as in (a), and with $K = \mu/(\mu+\kappa)$ so that the processes in (a) and (b) have the same expected number of points.
(c) Simulation of a standard recursive model with $\mu=0.1$, $\kappa = 2$, and g exponential with rate 1.
(d) Simulation of a Hawkes model with the same g and μ as in (c), and with $K = \mu/(\mu+\kappa)$ so the processes in (c) and (d) have the same expected number of points.
All 4 simulations are over the same temporal domain [0, 1000]. The points are spread uniformly over the y axis for ease of visualization.

Application to CA Rocky Mountain Spotted Fever cases.

Recorded cases of Rocky Mountain Spotted Fever in California from Jan 1, 1960 to Dec 31, 2011 were collected by the CDC and catalogued by Project Tycho, www.tycho.pitt.edu.

Weeks with no data over this period were treated as having zero confirmed cases.

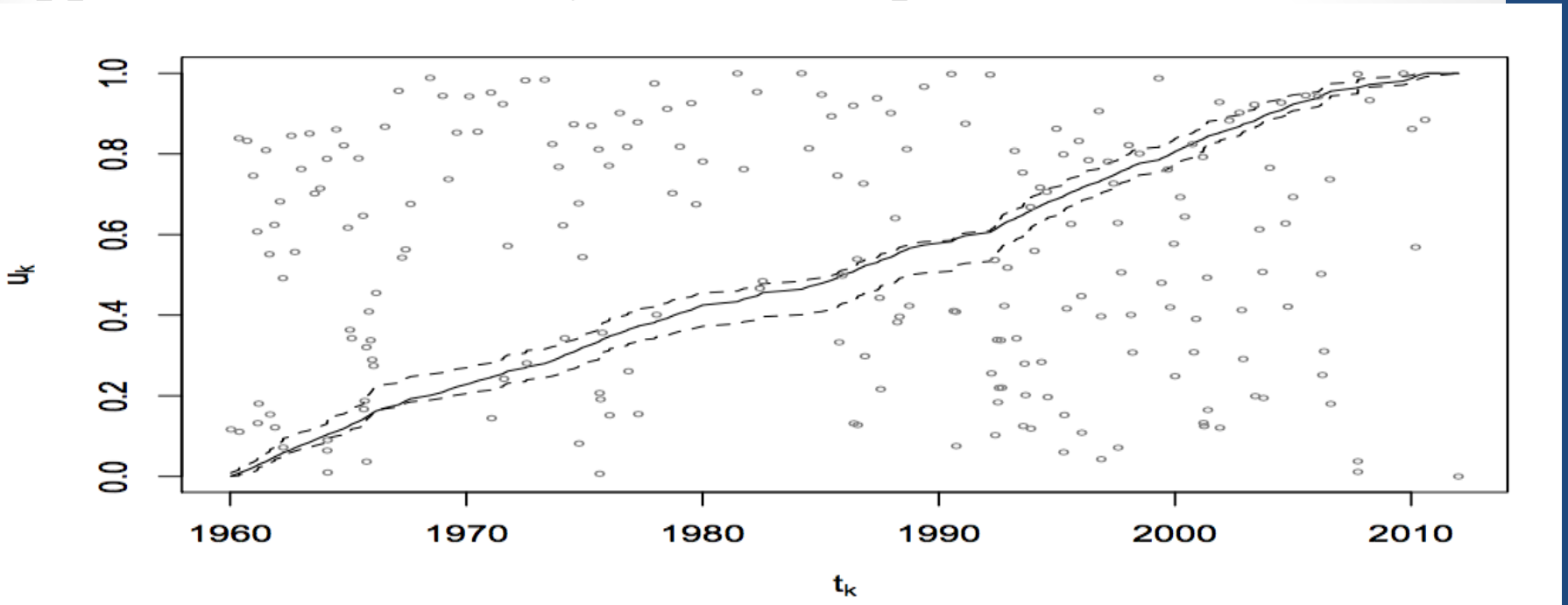


Histogram of confirmed Rocky Mountain Spotted Fever cases in California from 1/1/1960 to 12/31/2011, along with the estimated rate of the recursive model (green) and Hawkes model (red), each with exponential triggering function and fit by maximum likelihood.

Application to CA Rocky Mountain Spotted Fever cases.

In order further to assess the fit of the model, we used super-thinned residuals (Clements et al. 2013). In super-thinning, one selects a constant b, thins the observations by keeping each observed point τ_i independently with probability $b/\lambda(\tau_i)$ if $\lambda(\tau_i) > b$, and superposes points from a Poisson process with rate $(b-\lambda)1_{\lambda \leq b}$, where 1 denotes the indicator function. A default choice for b is the mean of $\hat{\lambda}$ at the observed points, as suggested in Gordon et al. (2015). The resulting super-thinned residuals form a homogeneous Poisson process with rate b iff. $\hat{\lambda}$ is the true conditional rate of the observed point process (Clements et al. 2013).

Application to CA Rocky Mountain Spotted Fever cases.



Super-thinned residuals t_k and their corresponding standardized interevent times u_k. The solid line shows, for each value of t_k, the normalized cumulative sum of u_k. There are fewer small interevent times than expected, especially between 1979 and 1985. Otherwise the interevent times appear to be well scattered.