Wildfires and Elk and the US Forest Service

David R. Brillinger

Department of Statistics
University of California, Berkeley

brill@stat.berkeley.edu
www.stat.berkeley.edu/~brill

AIF/CIF Salt Lake City, October 9, 2014

The USFS, USGS, and US Census can carry out the big projects, ones that smaller institutions cannot. They also have the resources to make the data conveniently available..

About 15 years ago Ager, Brillinger, Preisler (ABP) began produced a series of papers on animal movement and wildfire risk.

Other USFS researchers involved were: Burgan, Finney, McRoberts, Benoit, Wisdom, Kie, and Fujioka .

Canadian researchers involved too, a succession of meetings at BIRS. Martell, McAlpine, Taylor, Woolford, Wooton.

“Communicating Science in an Uncertain World”

Scientists communicate by talks, papers, teaching and nowadays even creating dotcoms.

Today focus on 4 AGP papers.

An ‘uncertain world”. Indeed and statisticians make it even more uncertain. In part by formalizing and estimating aleatoric (statistical) and epistemic (systematic)uncertainty.

Four papers of our papers will now discussed with graphics and uncertainty highlighted, two concern elk trajectories/paths and two concern wildfire risk,

Overall approach involves exploratory data analysis followed by formal statistical modelling, fitting and model assessment.

The trajectory papers proceed by working with potential functions (i.e. formalizing attraction and repulsion of moving objects).

The risk papers make use of spline functions to bring in smooth relationships.

Elk trajectories.

Elk, deer, and cows,. live in the Starkey Reserve while hikers, rid*ers,* bikers and hunters visit.

Starkey is a major experimental field area set up by the USFS - the largest research enclosure ever built to study usage and interactions.

Trajectories are defined by position estimates and times.

Source of data, $(x(t_j), y(t_j), t_j)$, $j = 1,\ldots,J$

http://www.fs.fed.us/pnw/starkey/introduction.shtml

Motivation: differential equation based on Newton's second law of motion

r = (x,y)

$$d\mathbf{r}(t) = \mathbf{v}(t)dt$$

$$d\mathbf{v}(t) = -\beta\mathbf{v}(t)dt - \beta\nabla H(\mathbf{r}(t), t)dt$$

with $\mathbf{r}(t)$ the particle's location at time t, $\mathbf{v}(t)$ the particle's velocity and $-\beta\nabla H$ the external force field acting on the particle, β being the coefficient of friction,

In the case that β is large, the equations are approximately

$$d\mathbf{r}(t) = -\nabla H(\mathbf{r}(t), t)dt$$

Earliest paper

Brillinger, D. R., Preisler, H. K., Ager, A. A. and Kie, J. G. "The use of potential functions in modelling animal movement". Data Analysis from Statistical Foundations (Ed. A. K. M. E. Saleh). Nova Science, Huntington (2001).

LORAN-C

$$d\mathbf{r}(t) = \mu(\mathbf{r}(t), t)\,dt + \Sigma(\mathbf{r}, t)\,d\mathbf{B}(t)$$

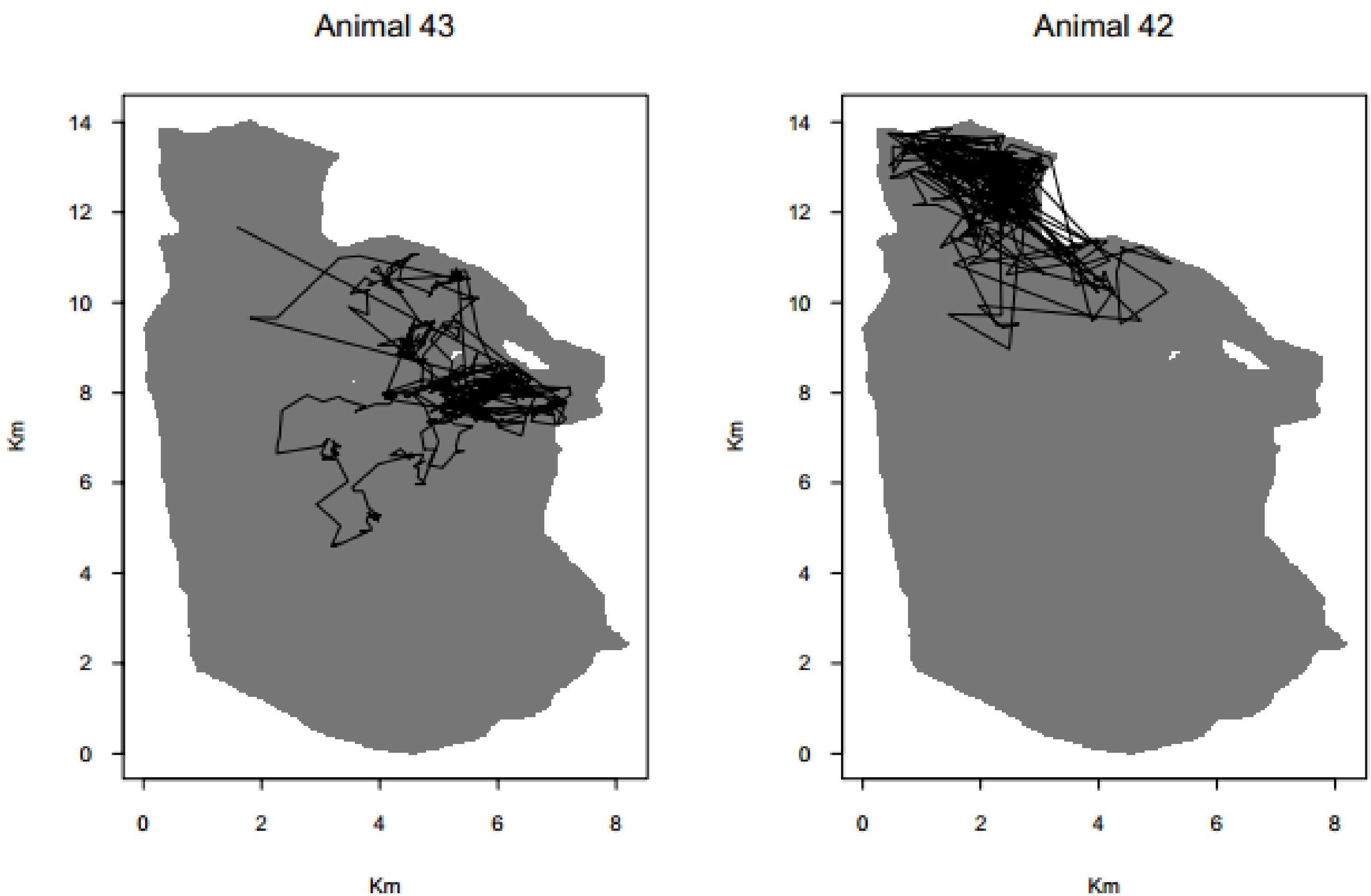
An SDE

A discrete approximation

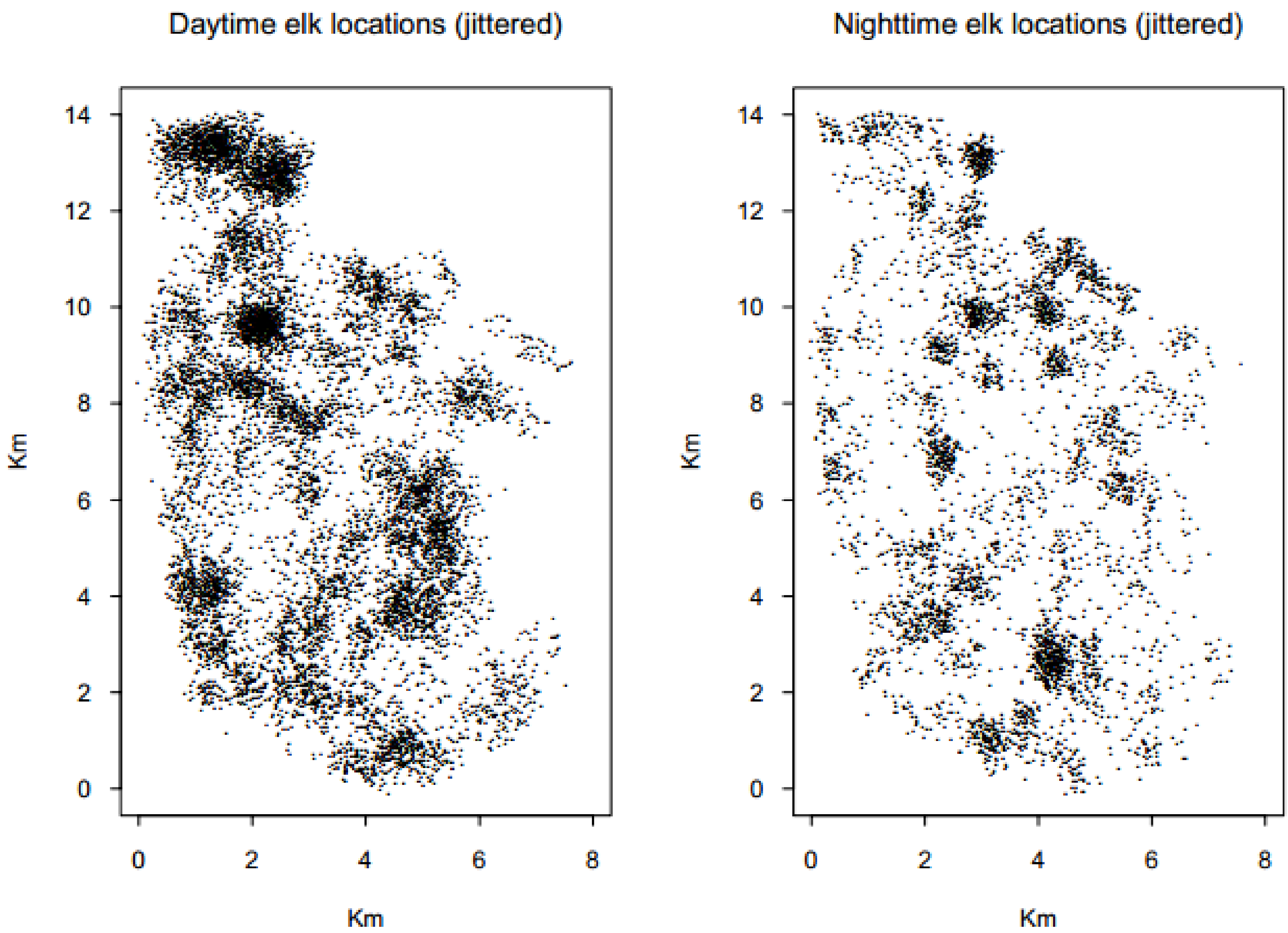
$$(\mathbf{r}(t_{i+1}) - \mathbf{r}(t_i))/(t_{i+1} - t_i) \approx \mu(\mathbf{r}(t_i), t_i) + \Sigma(\mathbf{r}(t_i), t_i)\mathbf{Z}/\sqrt{t_{i+1} - t_i}$$

LHS velocity **μ estimated by R-gam**

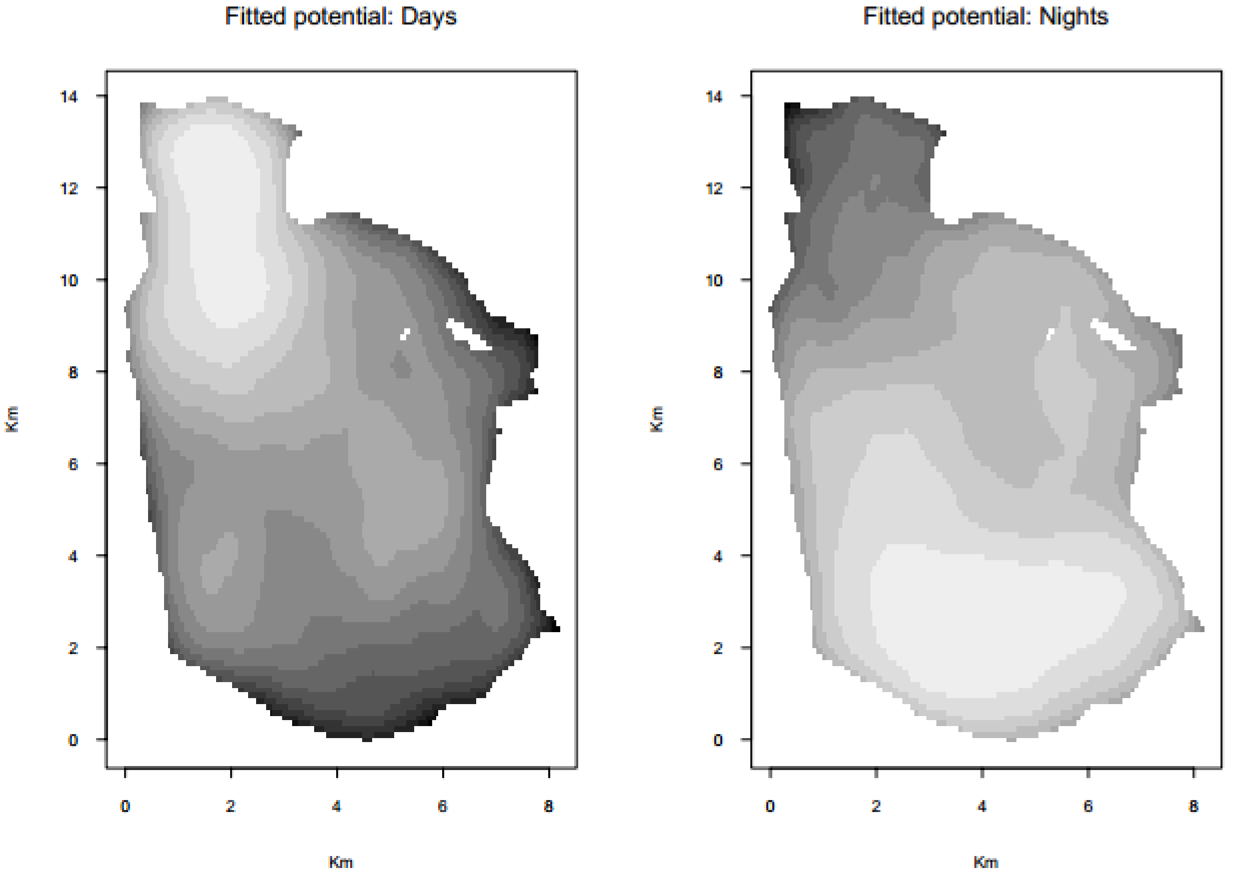
Data



Locations


Daytime elk locations (jittered)
Nighttime elk locations (jittered)
Km
Km
Km
Km
0
2
4
6
8
10
12
14
0
2
4
6
8

Potential estimate



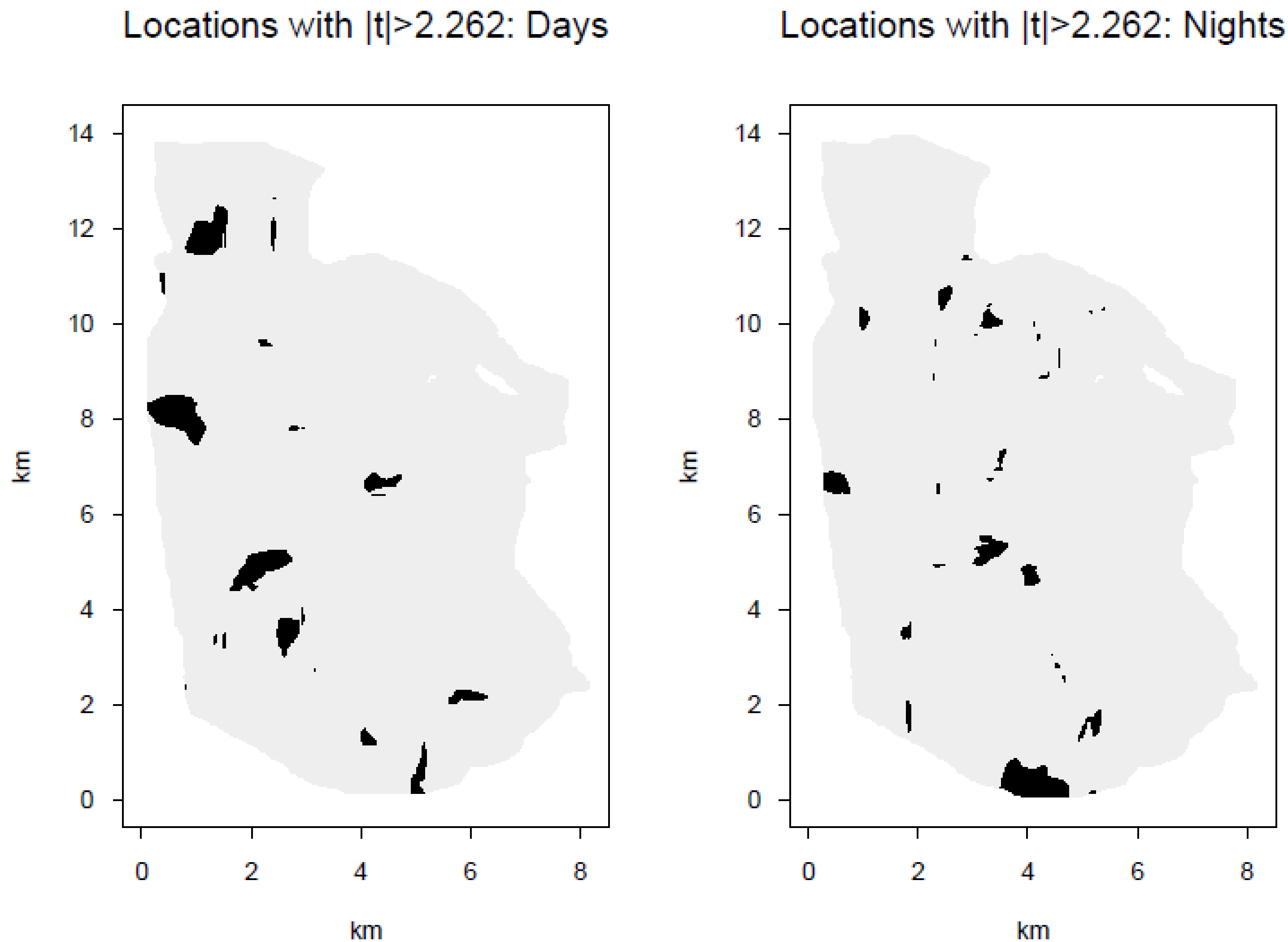
Simulation of trajectories used to estimate potential function, H.

But does a potential function exist?

$$\partial H_x / \partial y = \partial H_y / \partial x \quad ?$$

t-statistic

Jackknife employed, dropping 5 elk out of 50 ten times



Proportions exceeding 2.262

day 036

night .026

nominal ,050

Most recent paper

Brillinger, D. R., Preisler, H. K . and Wisdom, M. J. Modelling particles moving in a potential field with pairwise interactions and an application . Brazilian J. Prob. Statist. 25, 421-436 (2011).

radio collared $\Delta t \approx 5$ min

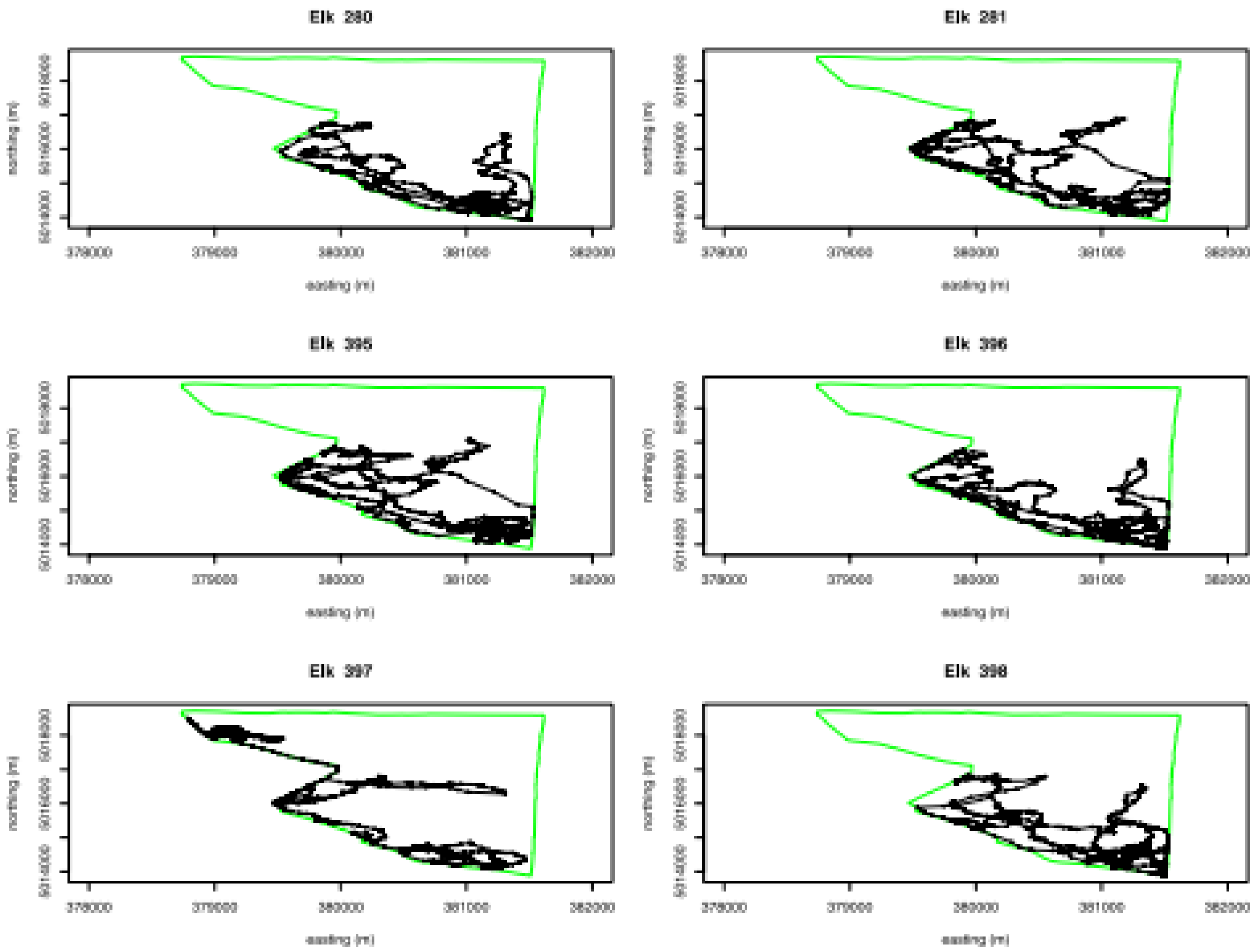
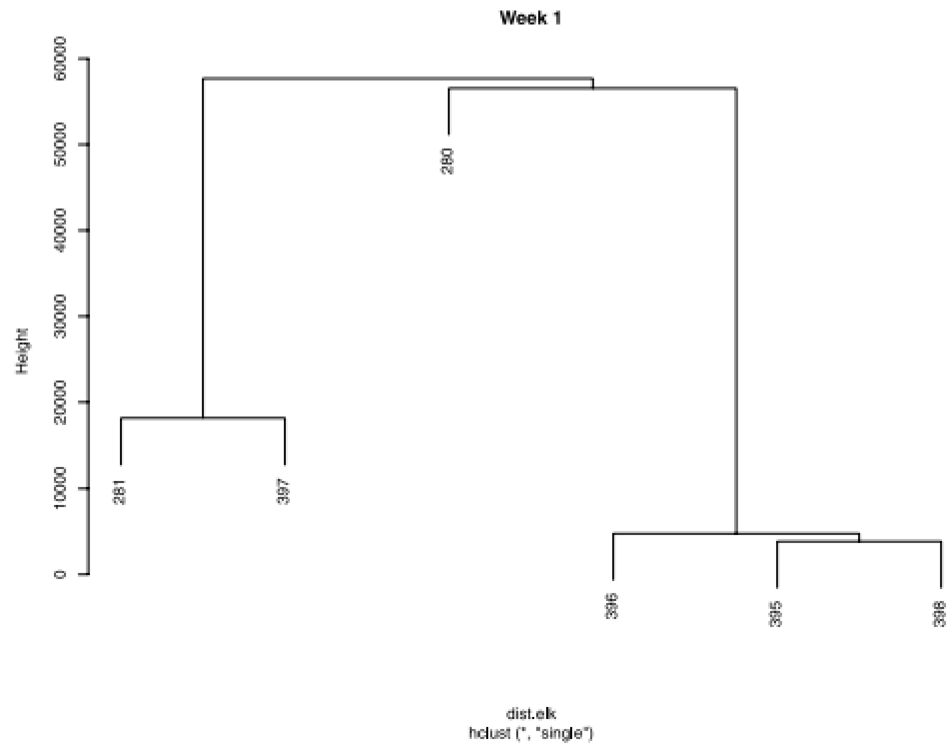


Figure 3 *Individual paths of six animals for the second week.*

Table 1 *Coherence estimates for first week of data*

elk	280	281	395	396	397	398
280	1	0.416	0.466	0.472	0.471	0.477
281	0.416	1	0.343	0.349	0.897	0.356
395	0.466	0.343	1	0.995	0.415	0.997
396	0.472	0.349	0.995	1	0.424	0.994
397	0.471	0.897	0.415	0.424	1	0.427
398	0.477	0.356	0.997	0.994	0.427	1



Hierarchical clustering

Figure 4 *Result of the hierarchical clustering analysis. First week's data for the six animals.*

Table 1 *Coherence estimates for first week of data*

$$d\mathbf{r}_i = -\sum_{j \neq i} \nabla V(\mathbf{r}_i - \mathbf{r}_j)\, dt + \sigma\, d\mathbf{B}_i .$$

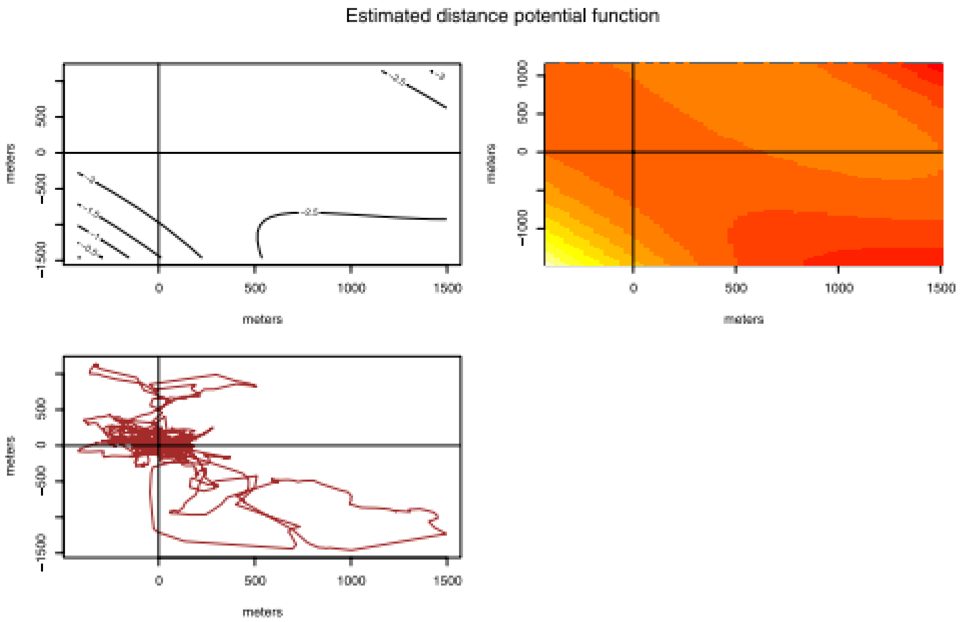


Figure 6 *Elk 398's movement with respect to elk 395.*

398 wrt 395 V(x,y) cubic in x, y

Hunters and elk, Suzette Puente's work

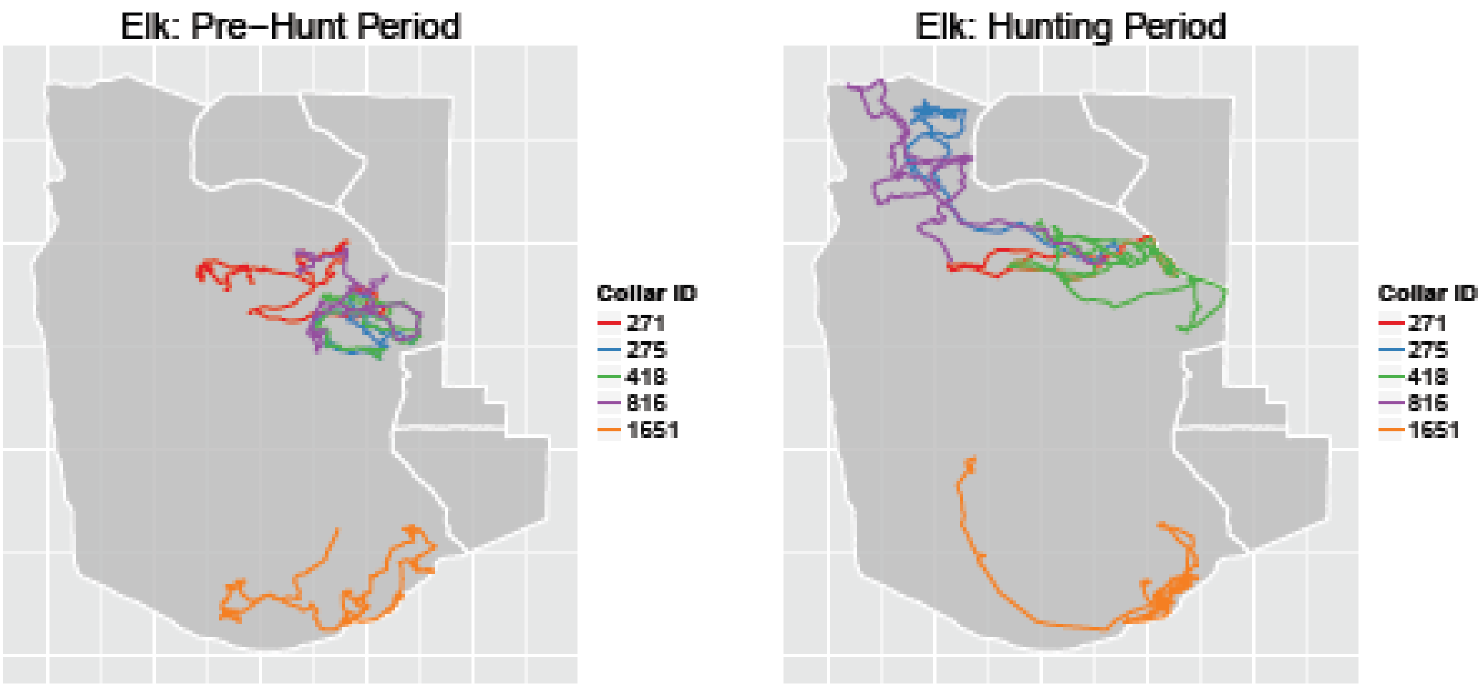


Figure 4: Trajectories of 5 randomly selected elk during the rifle-elk hunt (5 days). White lines represent the fences.

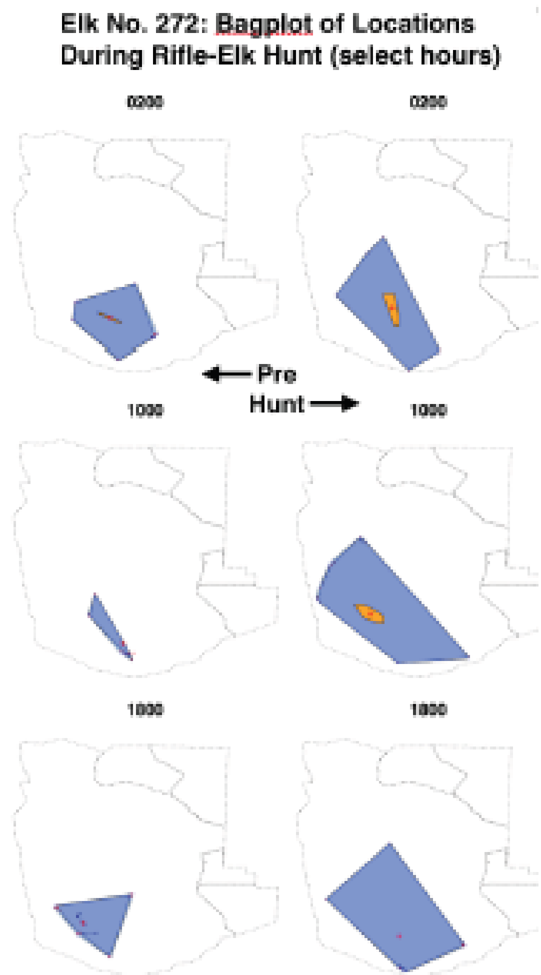


Figure 9: This elk covers more space during the hunt. The same pattern was consistent in all animals and all hours.

bagplots

Wildfire data.

Probabilistic risk assessment

Prob{fire in particular region and time period | explanatories}

insurance approach helpful

The data are fire locations (x_j, y_j, t_j) , $j = 1, \ldots, J$

voxel (x,y,t) $N(dx,dy,dt) = 1$ or 0 discrete $N_{x,y,t}$

$\eta_{x,y,t} = g_1(x,y) + g_2(d) + g_3(e)$ linear predictor

$\Pr\{N = 1 \mid \eta\} = \exp\{\eta\}/(1 + \exp\{\eta\})$ logistic regression R-gam

Earliest wildfire paper

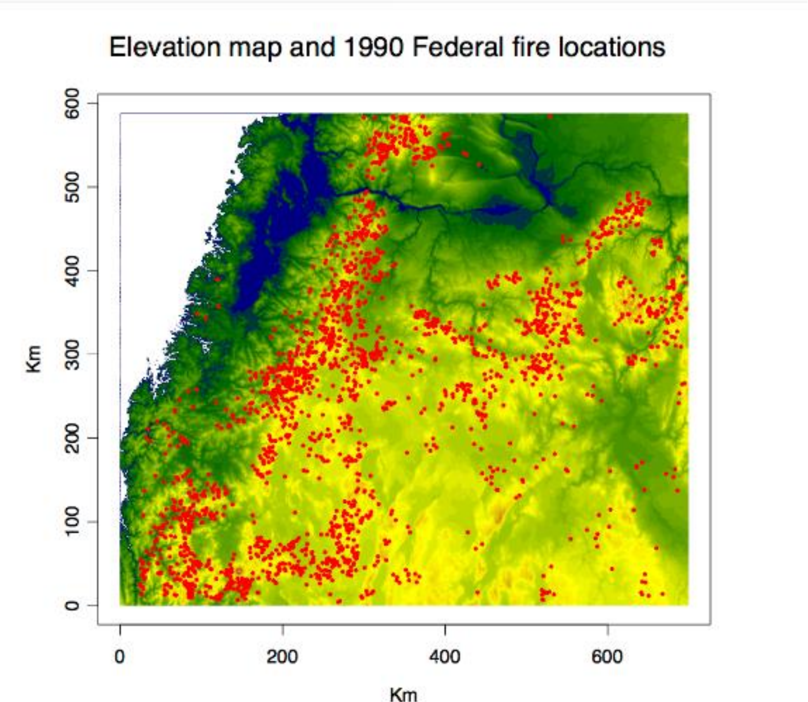
Brillinger, D.R., Preisler, H. K. and Benoit, J. W. "Risk assessment: a forest fire example”. Lecture Notes in Statistics 40, 177-196. IMS (2003)

Federal Lands in Oregon

1989 – 1996, 15786 fires > 0.1 acre

Fire locations

http://www.fs.usda.gov/rds/archive/Product/RDS-2013-0009.



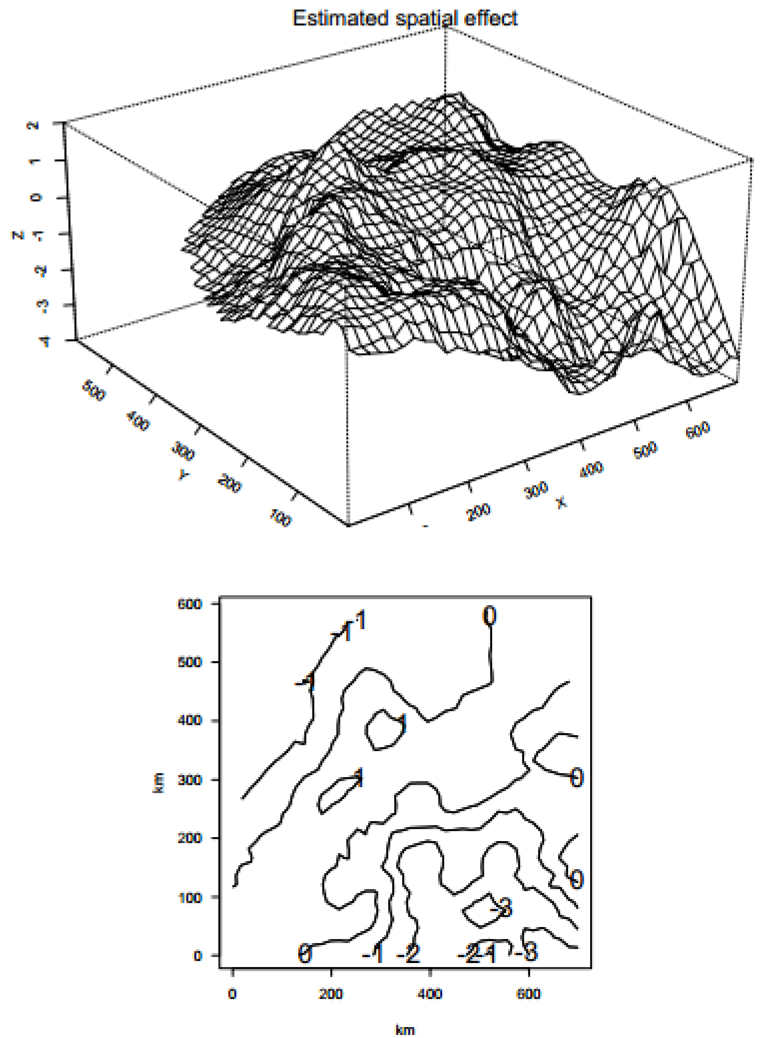


Figure 4: Estimated spatial effect, $\hat{g}_1$, for model (6a,b).

spatial effect g_3

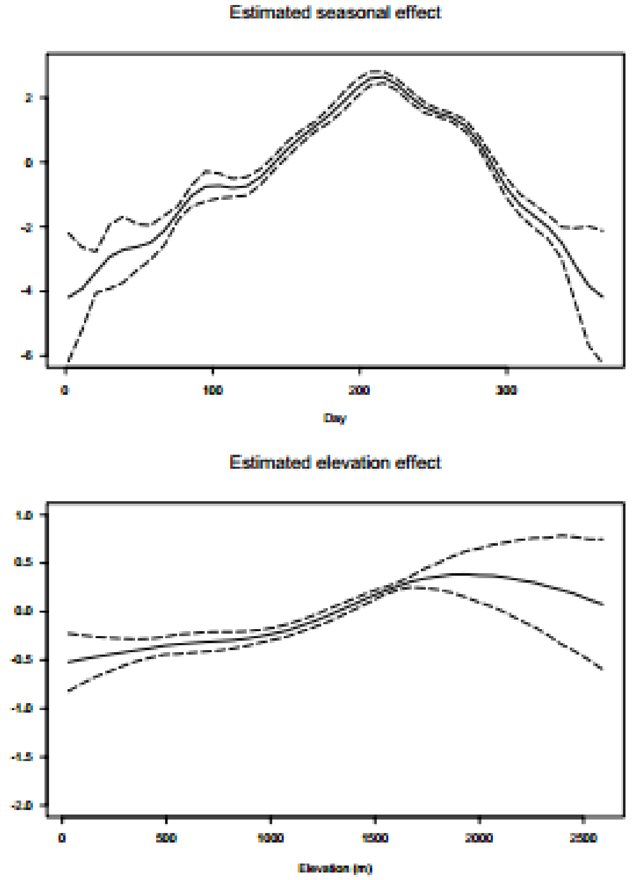


Figure 5: Estimated effects $\hat{g}_2$ of day in year and $\hat{g}_3$ of elevation for the model (6a,b). The dashed lines provide approximate marginal 95% bounds computed by a jackknife procedure.

seasonal and elevation effects g_2 , g_3

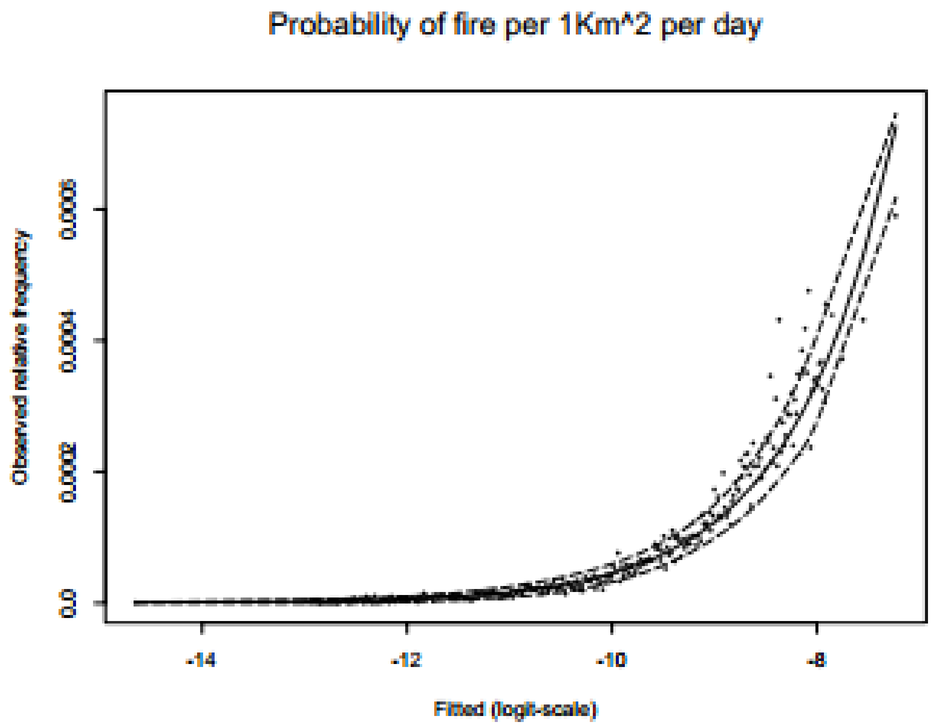


Figure 6: Observed relative frequencies of fire, after grouping the data into classes based on the fitted linear predictor, $\hat{\eta}$. The solid curve is the fitted logistic curve. The dashed lines are smoothed approximate 95% limits obtained via a binomial approximation.

assessment of logistic link

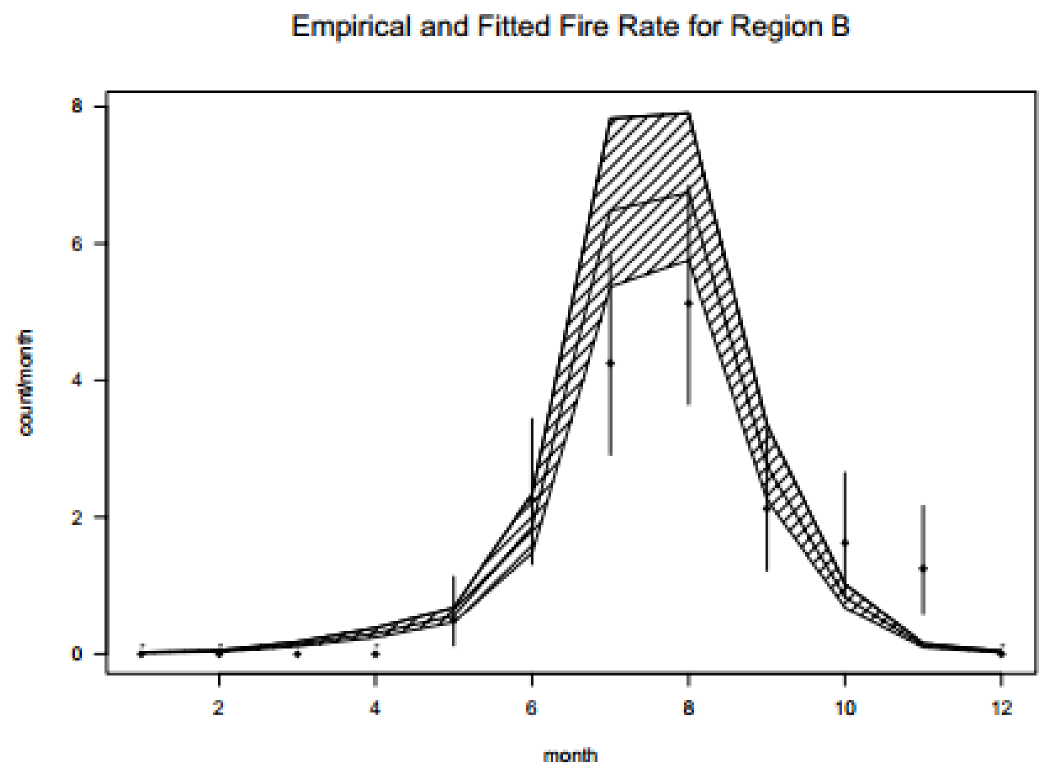


Figure 7: The solid central line gives the fitted rate of fires by month. The shaded region gives +- 2 s.e. limits. The points are the monthly empirical rates of fires. Vertical lines are +- 2 s.e. limits for the points.

Umatilla Forest result

i	*probability*	*confidence interval*
1	.989	(.967,.996)
2	.956	(.895,.983)
3	.887	(.771,.948)
4	.774	(.610,.883)
5	.628	(.441,.784)
6	.470	(.291,.658)
7	.324	(.176,.520)
8	.206	(.097,.386)
9	.121	(.049,.268)
10	.066	(.023,.174)
11	.033	(.010,.106)
12	.016	(.004,.060)
13	.007	(.001,.032)
14	.003	(.001,.016)
15	.003	(.000,.007)
16	.001	(.000,.003)
17	.000	(.000,.001)
18	.000	(.000,.001)
19	.000	(.000,.000)
20	.000	(.000,.000)

Table 1: Estimated probability of i or more fires and approximate 95% confidence limits for the month of July and region B.

Pr{ one or more fires }

Most recent paper

Brillinger, D. R., Preisler, H. K., and Benoit, J. "Probabilistic risk assessment for wildfires." Environmetrics 17, 623-633. (2006)

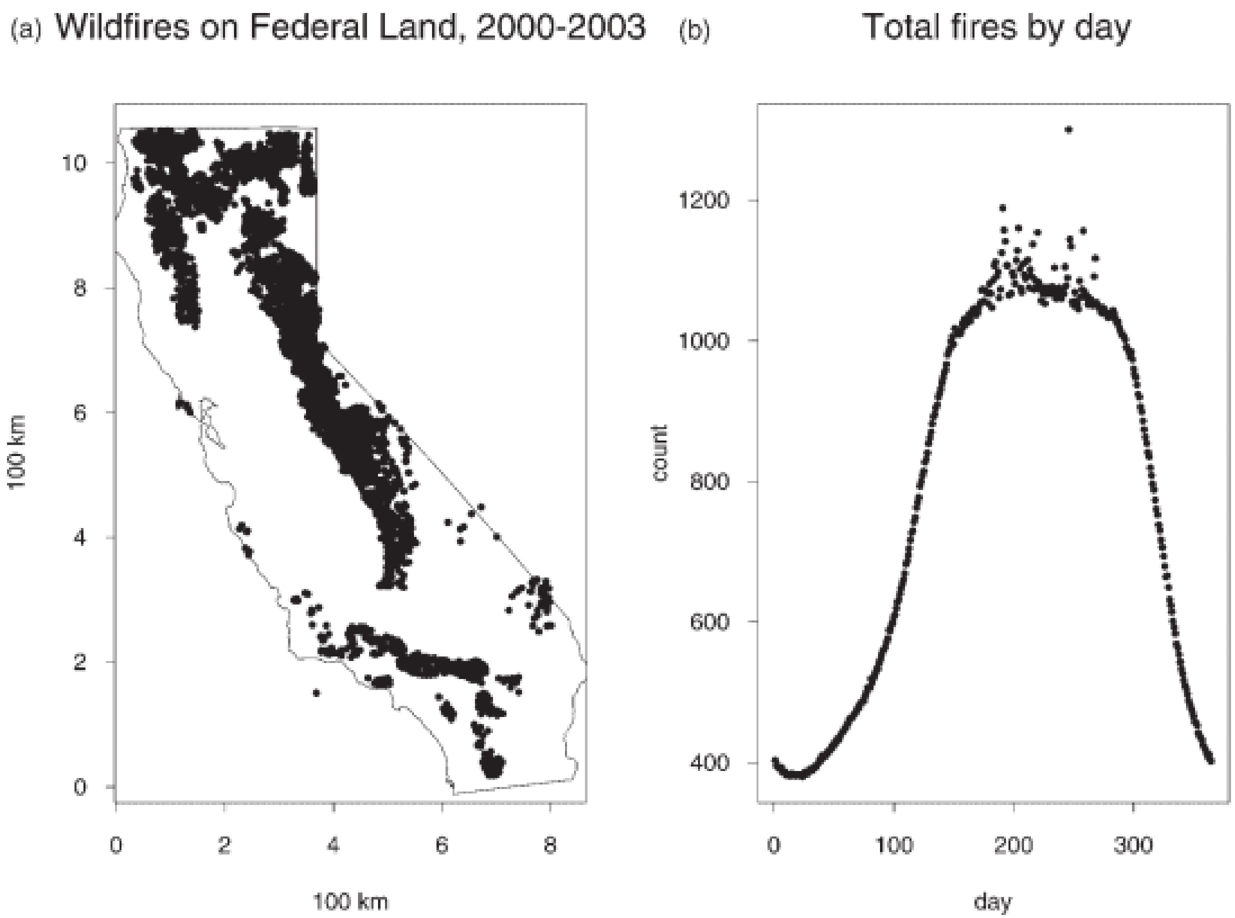


Figure 1. (a) Fire locations; the blocked off upper right area is the state of Nevada; (b) 4-year total for each day of the year

Model I: With Y, binary-valued and (x, y) and d fixed

$$\text{logit Prob}\{Y = 1 | x, y, d\} = g_1(x, y) + g_2(d)$$

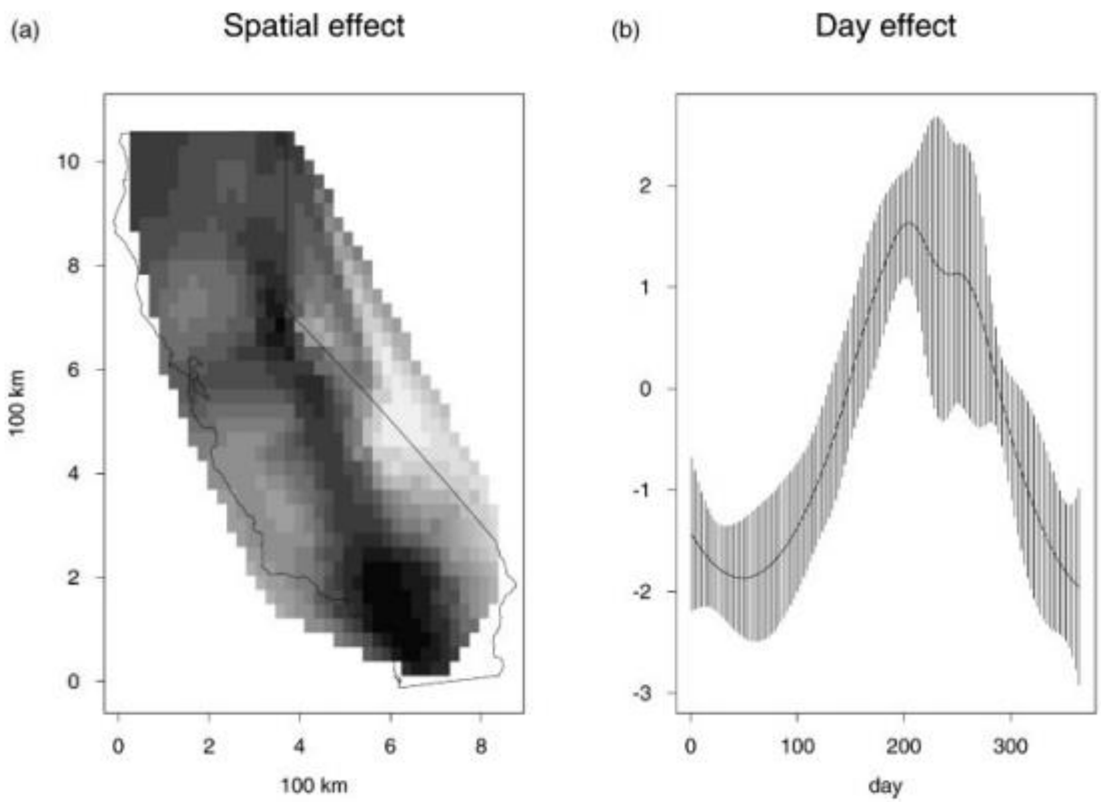


Figure 2. The estimated spatial and daily effects for Model I, (a) Provides an image plot of the estimated spatial effect. The darker values correspond to increased fire risk. In (b), the vertical lines provide approximate 95% confidence limits about a smoothed version of the solid line

Random effects (empirical Bayes)

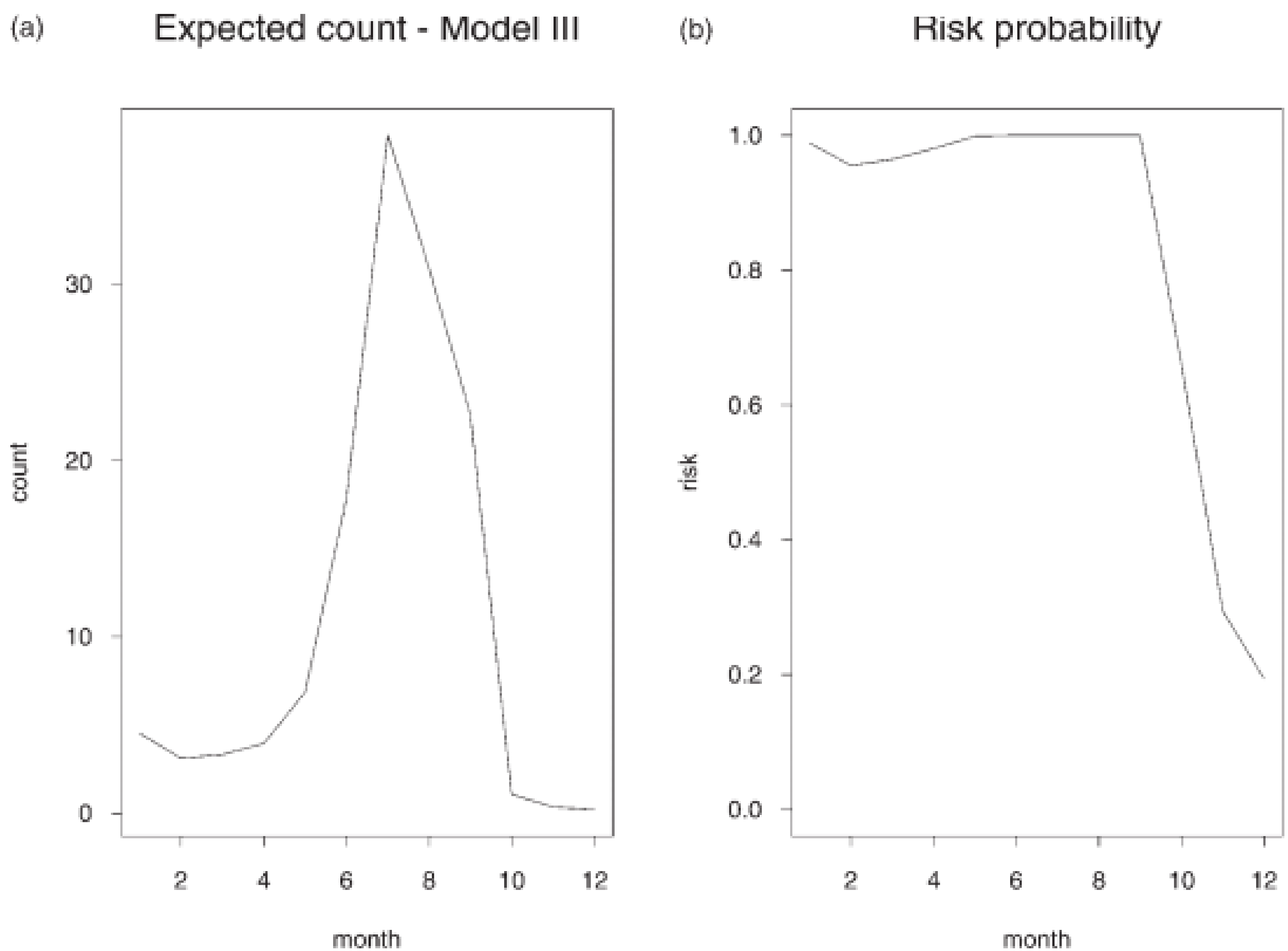
$$E_I\{\text{Prob}\{\text{fire in particular region and time period} \mid \text{explanatories}, I\}\}$$

Model III: With I, a factor whose effects are independent normals with mean 0 and variance τ^2

$$\text{logit Prob}\{Y = 1|x, y, d, I\} = g_1(x, y) + g_2(d) + I$$

$$\text{Prob}\{Y = 1|\text{explanatories}\} = \int \frac{\exp\{\eta + \tau z\}}{(1 + \exp\{\eta + \tau z\})}\phi(z)\text{d}z$$

Prob{At least one fire in a particular region and month}



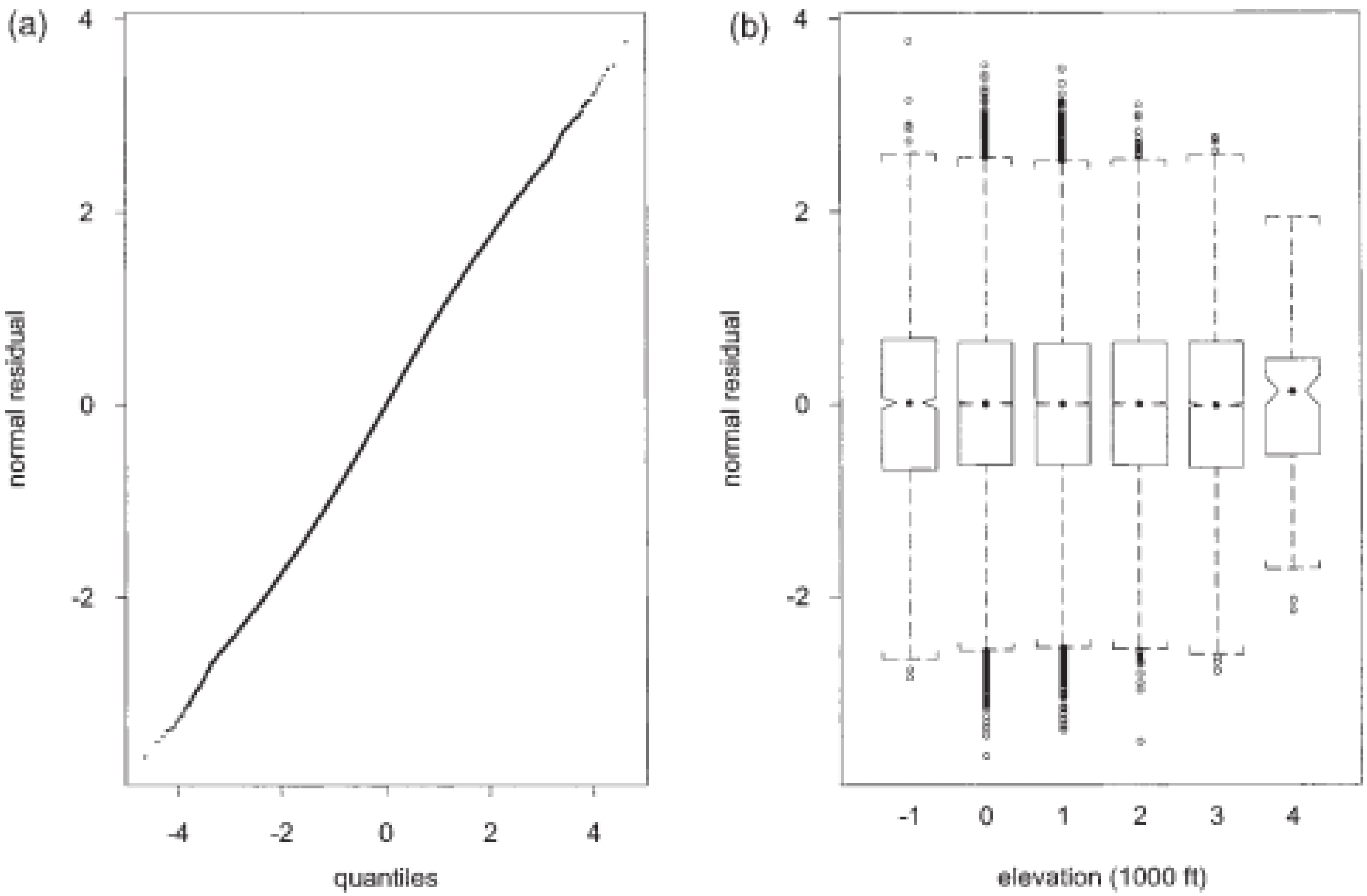


Figure 9. Model assessment results for Model III. The negative elevations in (b) come from Death Valley

Model assessment

Summary.

The USFS has advanced forest science by collecting data in novel ecological experiments and making it publicall yavailable and then encouraging its scientists collaborate with outside researchers.

There has been an emphasis on communication. I communicate as an academic:, consulting, papers, talks, teaching, thesis supervision

There has been an emphasis on assessing uncertainty.

I take this opportunity to thank the many foresters particularly USFS people who made this research possible and Abdel for publishing much of it. Thank you all.

DAN PIRARO 9-27-14
I heard that!
BIZARROCOMICS.COM Facebook.com/BizarroComics Dist by King Features

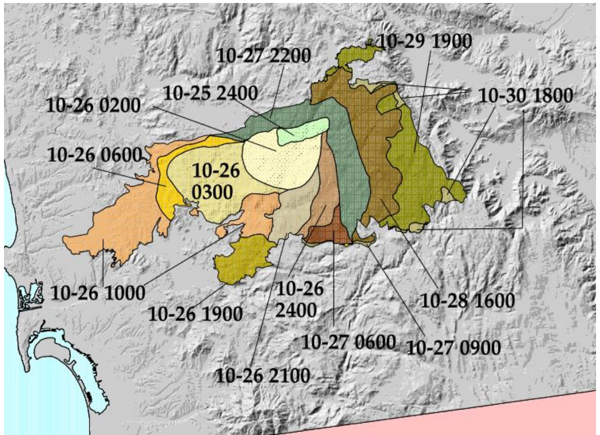
D. R. Brillinger, B. S. Autrey and M. D. Cattaneo, "Probabilistic Risk Modelling at the Wildland Urban Interface: the 2003 Cedar Fire". Environmetrics

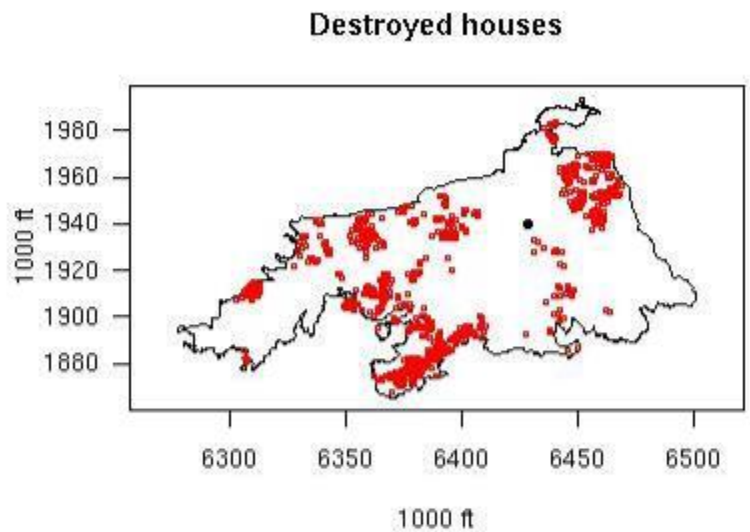
The 2003 Cedar Fire.

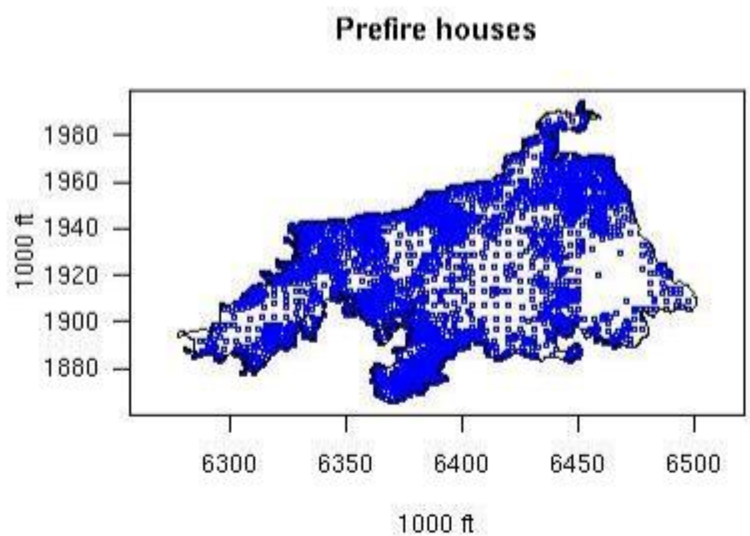
Southern California near San Diego.

Largest fire in California history.

Large data set created for analytic use by San Diego community.




Destroyed houses
1980
1960
1940
1920
1900
1880
1000 ft
6300
6350
6400
6450
6500
1000 ft


Prefire houses
1980
1960
1940
1920
1900
1880
1000 ft
6300
6350
6400
6450
6500
1000 ft

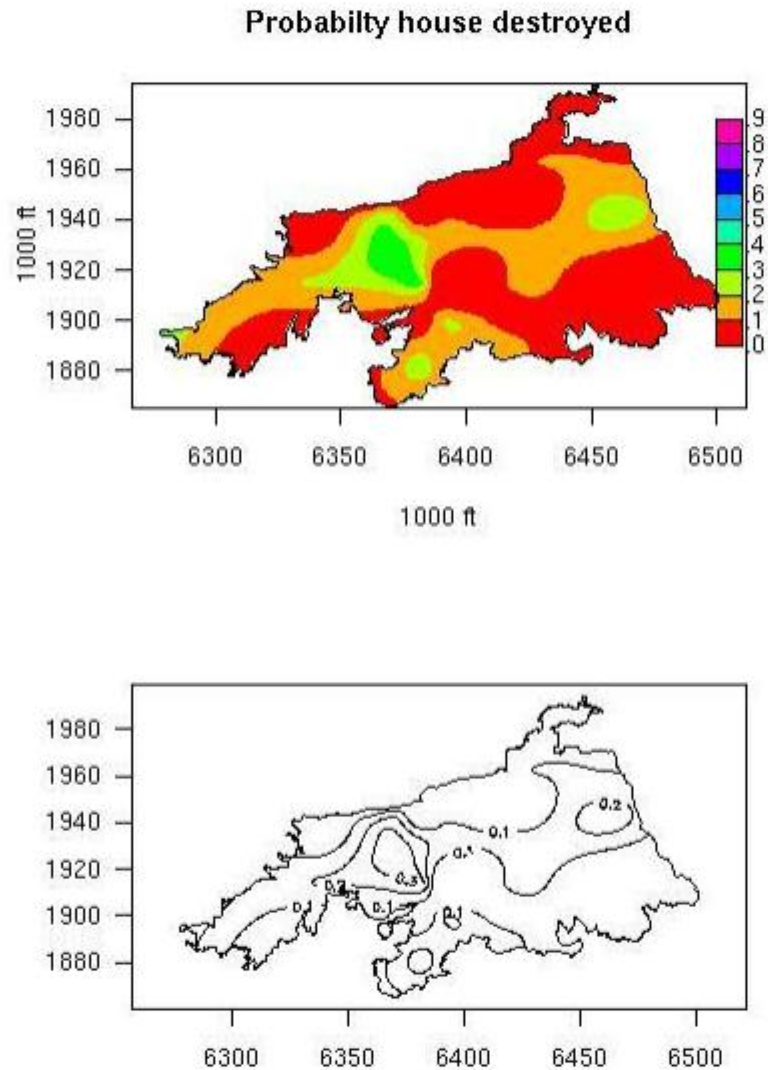


Figure 6. Estimated destruction probability as a function of location. The estimate is displayed in both perspective and contour form.

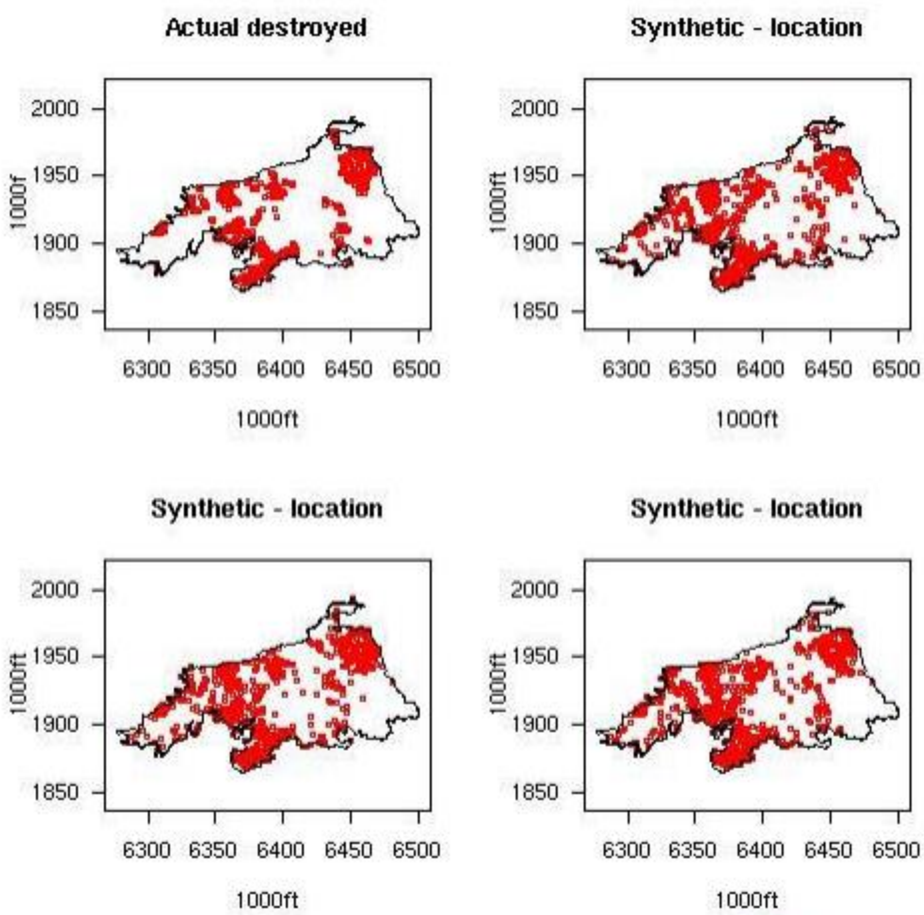
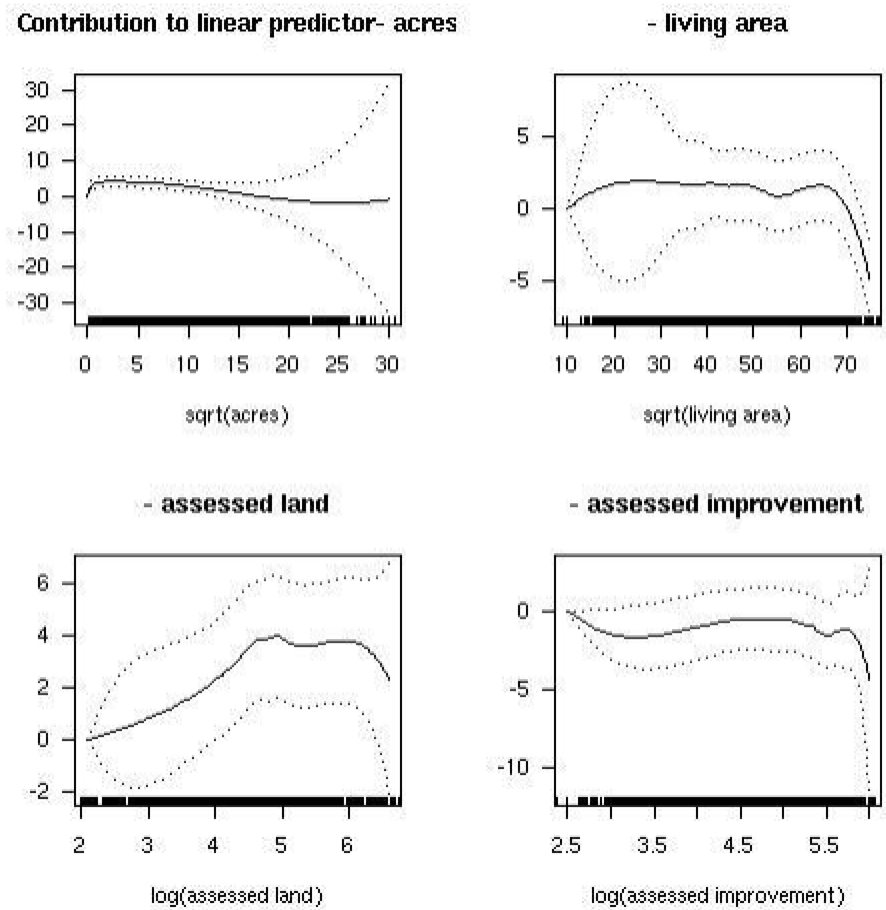


Figure 7. The locations of the original houses destroyed and three synthetic plots involving random thinnings employing the estimated probability function of Figure 6.



Contribution to linear predictor- acres
30
20
10
0
-10
-20
-30
0 5 10 15 20 25 30
sqrt(acres)
- living area
5
0
-5
10 20 30 40 50 60 70
sqrt(living area)
- assessed land
6
4
2
0
-2
2 3 4 5 6
log(assessed land)
- assessed improvement
0
-5
-10
2.5 3.5 4.5 5.5
log(assessed improvement)

Brillinger, D. R. and Finney, M. An exploratory data analysis of the temperature fluctuations in a spreading fire Environmetrics DOI:10.1002/env.2279 (2014)

Figure 1. The combs of tines are laid out regularly along a straight testbed

Figure 2. This photo was taken from the side of the testbed. The wind is blowing from the left, and the flames are moving to the right. Some tines are still standin

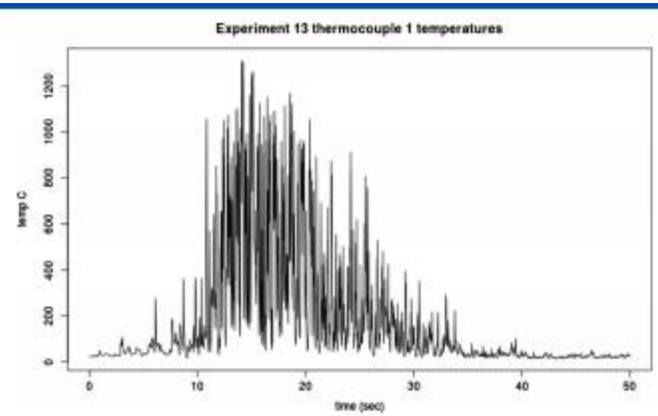


Figure 3. The temperatures (°C) recorded at thermocouple 1. There are 25,001 measurements in all

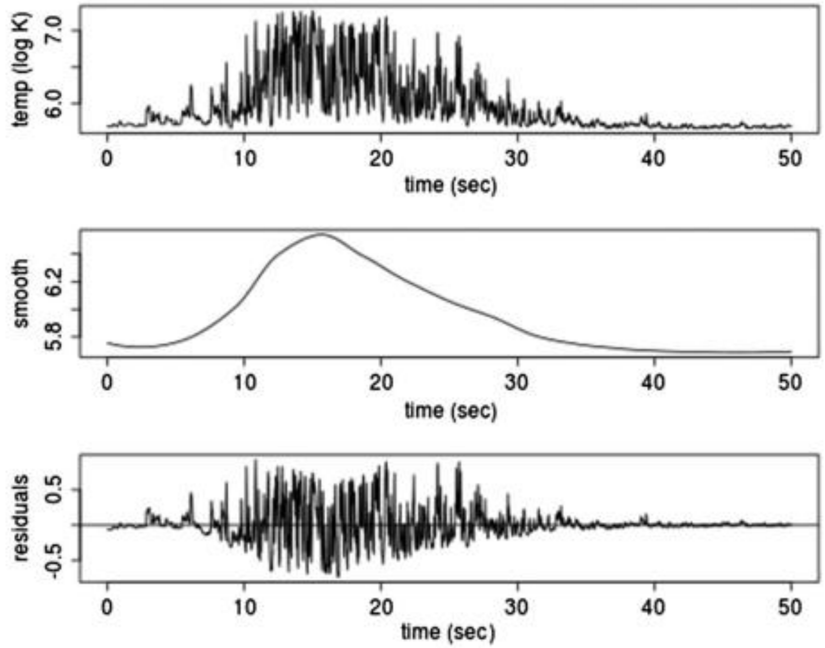


Figure 5. The series shown are, respectively, the data, the smoothed values obtained by loess, and the difference between these two, the residuals. The data

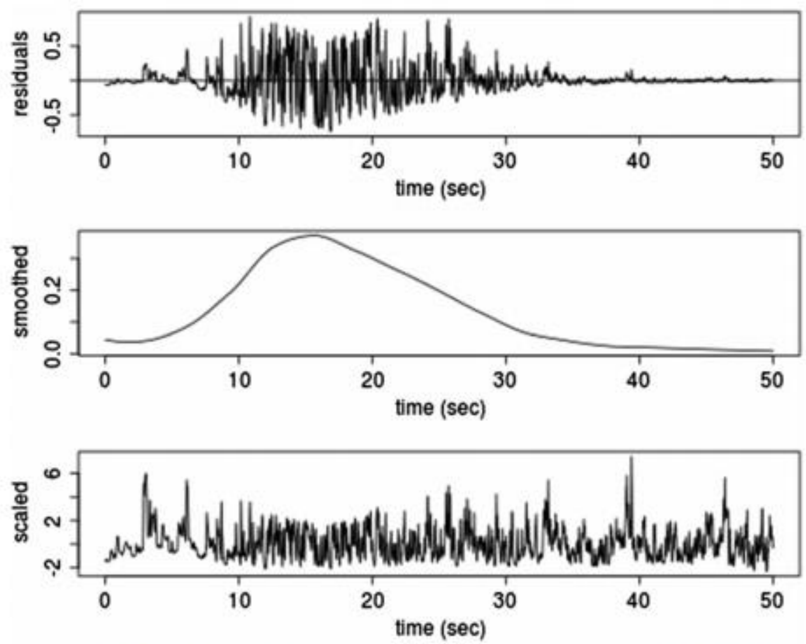


Figure 6. The panels are, respectively, the residuals of the last panel of Figure 5, the smooth of the absolute residuals, and the result of dividing these latter into the corresponding values of the top plot