Probabilistic Risk Modelling at the Wildland Urban Interface: the 2003 Cedar Fire, II

David R. Brillinger, Statistics UC Berkeley
(brill@stat.berkeley.edu, www.stat.berkeley.edu/~brill)

Benjamin Scott Autrey, PopularMedia Inc. (Sequoia Ventures) (benjamin.scott@gmail.com)

Matias D. Cattaneo, Economics UC Berkeley
(cattaneo@econ.berkeley.edu)

Data preparation, data preparation, data preparation

Booker's Law

An ounce of application is worth a ton of abstraction.

Overview.

A story of wildfires at the urban-wildland interface

"… where humans and their development meet or intermix with wildland fuel." Federal Register (2004)

Getting/preparing data – yiiiih

Focus on the San Diego County Cedar Fire of 2003

Looking for: associations with explanatories, understanding of movement, …

Trying to understand costs - losses of life, property, animals, social cost (veg), private cost (home), fire suppression, …

Insurance premium?

The Cedar Fire.

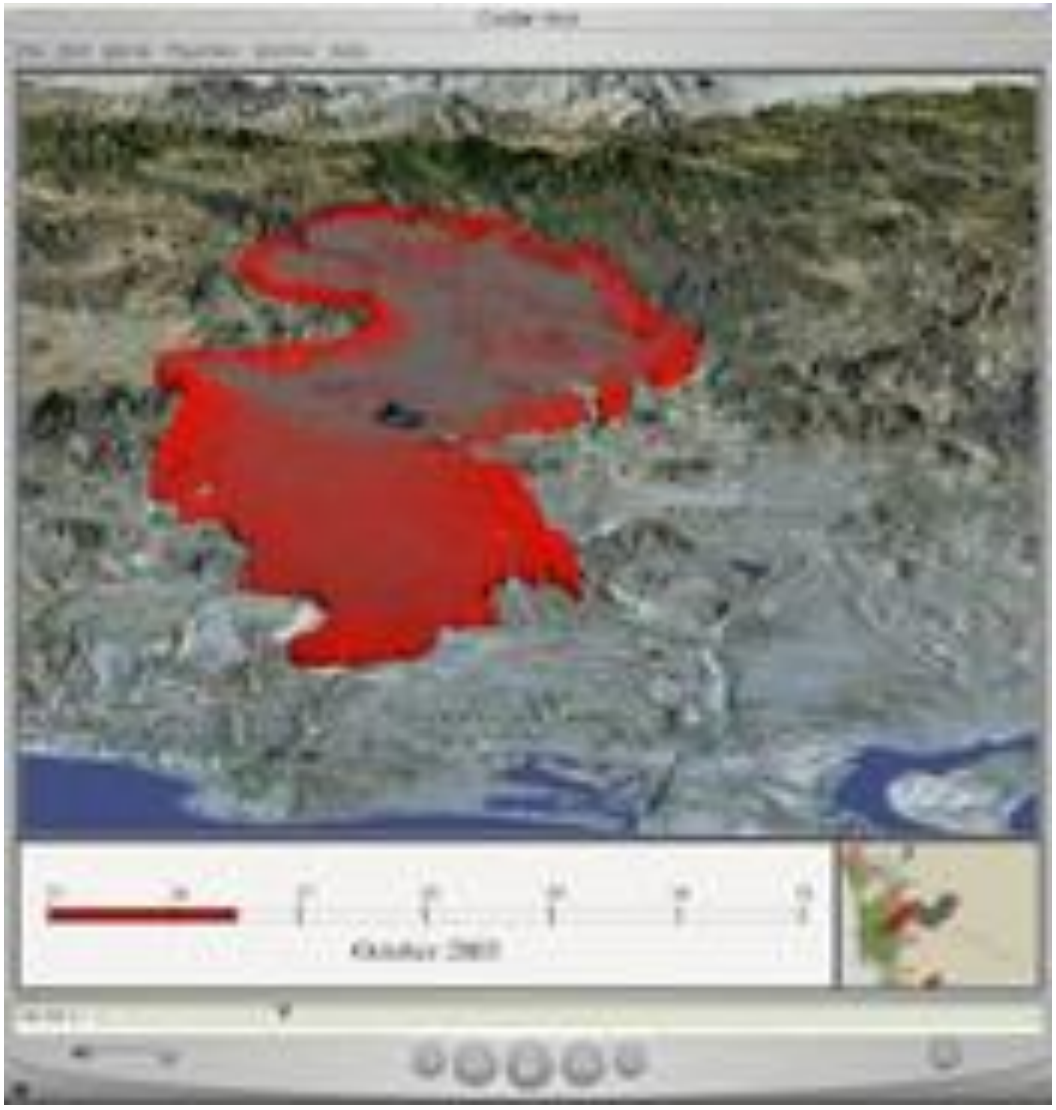
25 October – 4 November, 2003

15 deaths, 6000 firefighters, 2232 homes, 273246 acres, many evacuations, … (All ±)

Santa Anna conditions

A disaster

Large amounts of data, but …



J. Gibbons

Some formalism.

Spatial marked point process

Data (x_i , y_i , M_i)

(x_i,y_i): location, M_i: mark

How to describe a point process X?

$dX(x,y)/dxdy = \Sigma\ \delta(x-x_i,y-y_i)$

Dirac delta

Rate/intensity

$\mu_X(x,y) = E\{\Sigma\ \delta(x-x_i,y-y_i)\}$

Perhaps Y a subset of X (e.g. destroyed)

Ratio of rates

$p(X,Y) = \mu_Y(x,y) / \mu_X(x,y)$

Useful for comparison, …

How to describe a m.p.p.

$dU(x,y)/dxdy = \Sigma M_i \delta(x-x_i,y-y_i)$

Average

$v_U(x,y) = E\{\Sigma M_i \delta(x-x_i,y-y_i)\}$

Thinning with $M_i = 0$ or 1 randomly yields p.p.

Y subset of X

Ratio of averages

$v_V(x,y)/v_U(x,y)$

Logit-gam model

Logit{Prob[destroyed|explanatories]}

= α_j with j vegetation class

= $\beta(x,y)$ with (x,y) location

= $\gamma(s)$ with s slope

= $\delta(a)$ with a assessed improvement value

= $\alpha + \beta + \gamma + \delta + (\alpha\beta) + \ldots$

After first case, function is assumed smooth

Developing “the” data set.

Many people, organizations, file formats, coordinate-systems, decisions, definitions, authorities, issues, skills, tricks, uncertainties,Nas, errors, checks,…

Publically available data

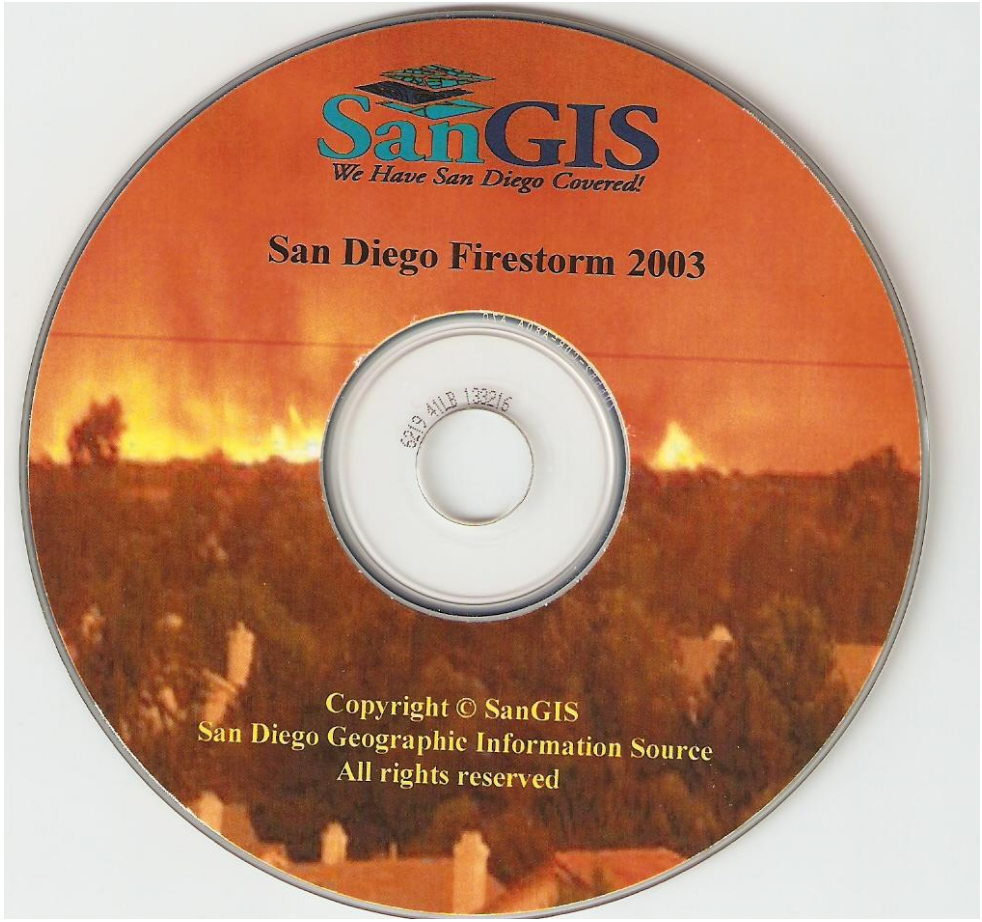
Tax records, assesors, satellites

GIS files – didn’t need package

Difficulty merging – APN, (X,Y), address,…

Response: 0-1 (destroyed) or continuous (sq ft)

Explanatories: topography, vegetation, roofing, brush,…


SanGIS
We Have San Diego Covered!
San Diego Firestorm 2003
Copyright © SanGIS
San Diego Geographic Information Source
All rights reserved

TAX RECORD DATA – all houses

AREA PERIMETER PARCEL_ PARCEL_ID PARCELID OVERLAY_JU POSTID
POSTDATE SUBDIVID GRAPHSRC CONFACTR APNID APN_POSTID
APN_POSTDA PENDING APN APN_8 MULTI OWN_NAME1 OWN_NAME2
OWN_NAME3 FRACTINT OWN_ADDR1 OWN_ADDR2 OWN_ADDR3 OWN_ADDR4
OWN_ZIP ASR_SITENA LEGLDESC ASR_LAND ASR_IMPR ASR_TOTAL ACREAGE
TAXSTAT OWNEROCC TRANUM ASR_ZONE ASR_LANDUS SUBMAP SUBNAME
UNITQTY ADDRNO ADDRFRAC ADDRUNIT ROADPDIR ROADNAME ROADSFX
JURIS ZIP X_COORD Y_COORD SITUS_ADDR SITUS_FRAC SITUS_SUIT
SITUS_PRE_ SITUS_NAME SITUS_SUFF SITUS_POST YEAR_EFFEC
TOTAL_LVG_ BEDROOMS BATHS ADDITION_A GARAGE_CON GARAGE_STA
CARPORT_ST POOL PAR_VIEW USABLE_SQ_ OBJECTID

SUMMARY							RESIDENTI	
Report Number	Community	Street Number	Street Name	GPS Location	Photos	Assessor Parcel No.	Sq ft	COUNTY Assess Valuation
5012	Alpine	502	S. Glen Oaks		disk 1 #12	40307501	1600	$ 127,500
5015	Alpine	2198	Larkspur		disk 2/photo :	#########	1584	$ 114,444

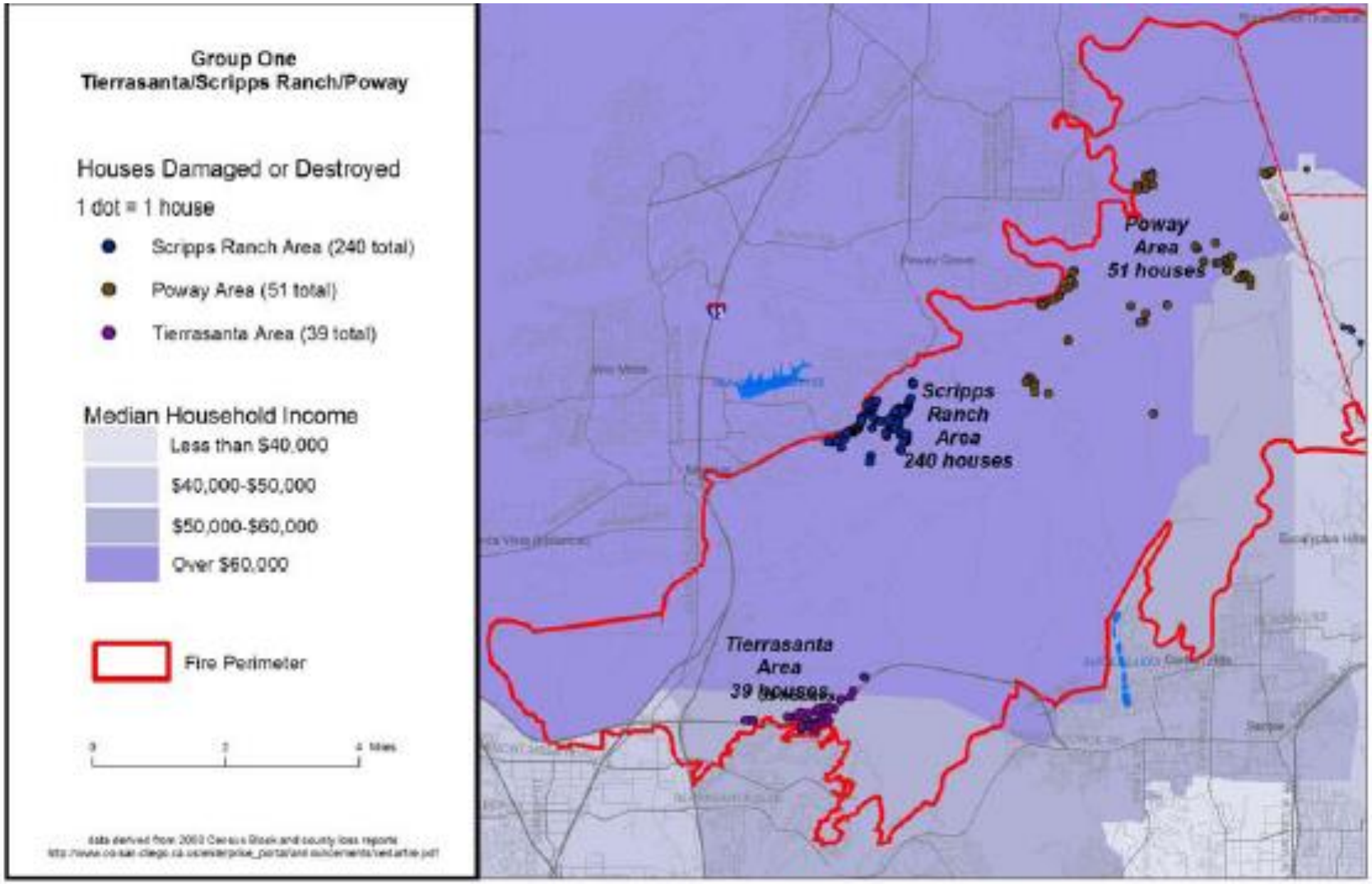
AL & COMMERCIAL STRUCTURES				OTHER LOSS				
Replacement Cost Per Sq Ft	Structure Damage			Out Building Damage, Other Improvements			Vehicles, Travel Traile Tractors	
$150	DS	DM	Pct.	DS	DM	Loss $$ ($20/sq ft)	DS	DM
$ 240,000	1			1				
$ 237,600	1			2				

ers, Boats,	Structure Contents % of Rep. Cost	
$$ Loss	50%	COMMENTS
$ 120,000		utility shed
$ 118,800		patio cover 2

Damage Assessment

R. Martin

CN Fire Name Photo Log Foundation Number Address Number Street Name Add'l Location Info Latitude Longitude Township Range Occupant Name Owner Name Insurance Carrier Structure Type Construction Type Fire Rated? Occupancy Type Type of Business Property Use #Dwellings Damaged #Dwellings Destroyed #Dwellings Saved #outbuildings damaged #Outbuildings Destroyed #Outbuildings Saved #Vehicles Damaged #Vehicles Destroyed #Vehicles Saved Structure Condition Structure Status Defensibe Space Defensive Actions Taken? By Whom? Roof Covering Ground Floor Length Ground Floor Width SF Number of Stories Construction Quality Year Built Property Management Civilian Injuries Civilian Deaths FF Injuries FF Deaths Area of Fire Origin Area: Level of Certainty Area: INFO SOURCE Form of Heat of Ignition Form: Level of Certainty Form: INFO SOURCE Structural Factors Vegetation Facors Logistical Factors Environmental Factors Operational Factors HYDRANT? Location Slope Property Line Setback Adjacent Structure Setback Prevailing Vegetation Type Veg Specific Veg Distance Veg Condition Access Grade Access Width Access One Way? Access Dead End? Access Turnaround? Driveway Grade Driveway Width Driveway Vertical Driveway Passing Lane? Driveway Turnaround? Wall Const. Deck/Porch Window Glass Type Window Frame Type Attic/ Subfloor Vents Skylight Present? Skylight Surface Area Skylight Type Door: Sliding Glass Type Door: French Type Door: Other Type Eave Const. Overhang Width Rain Gutter Construction Address Present? Visible from Road? Contrasting? Letter Height Letter Width Stroke Width Greenbelt or Fuelbreak Present? Fuelbreak Width Fuelbreak Length Fuelbreak Observed Effect Type of Water Supply Fire Sprinklers Present? Interior or Exterior Sprinkler Type Remarks Observations with Address and Damage info Observations with GPS and Damage info



Map 2: Effects/Income Map

Analyses.

spatial, spatial-temporal, binary, continuous

R functions: str(), read.shapefile(), inout(), match(), read.xls(), read.dbf(), image(), as.numeric(as.character()), library(),…

Example of shapefile contents

```
str(L)
List of 3
 $ shp:List of 2
  ..$ shp   : num [1, 1:4]       1  309373 -549829       1
  .. ..- attr(*, "dimnames")=List of 2
  .. .. ..$ : NULL
  .. .. ..$ : chr [1:4] "record" "x" "y" "shape.type"
  ..$ header:List of 12
  .. ..$ file.code   : int 9994
  .. ..$ file.length : int 64
  .. ..$ file.version: int 1000
  .. ..$ shape.type  : int 1
  .. ..$ xmin        : num 309373
  .. ..$ ymin        : num -549829
  .. ..$ xmax        : num 309373
  .. ..$ ymax        : num -549829
  .. ..$ zmin        : num 0
  .. ..$ zmax        : num 0
  .. ..$ mmin        : num 0
  .. ..$ mmax        : num 0
 $ shx:List of 2
  ..$ index : num [1, 1:2] 50 10
  .. ..- attr(*, "dimnames")=List of 2
  .. .. ..$ : NULL
  .. .. ..$ : chr [1:2] "Offset" "Length"
  ..$ header:List of 12
  .. ..$ file.code   : int 9994
```

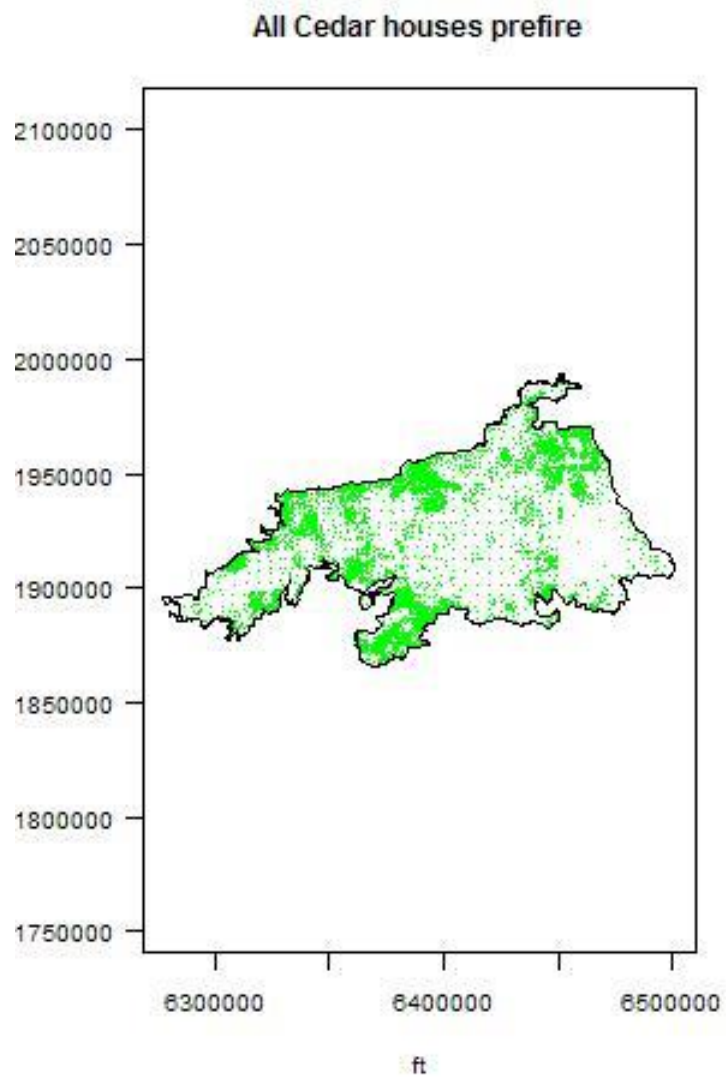
Looking at the point process data.

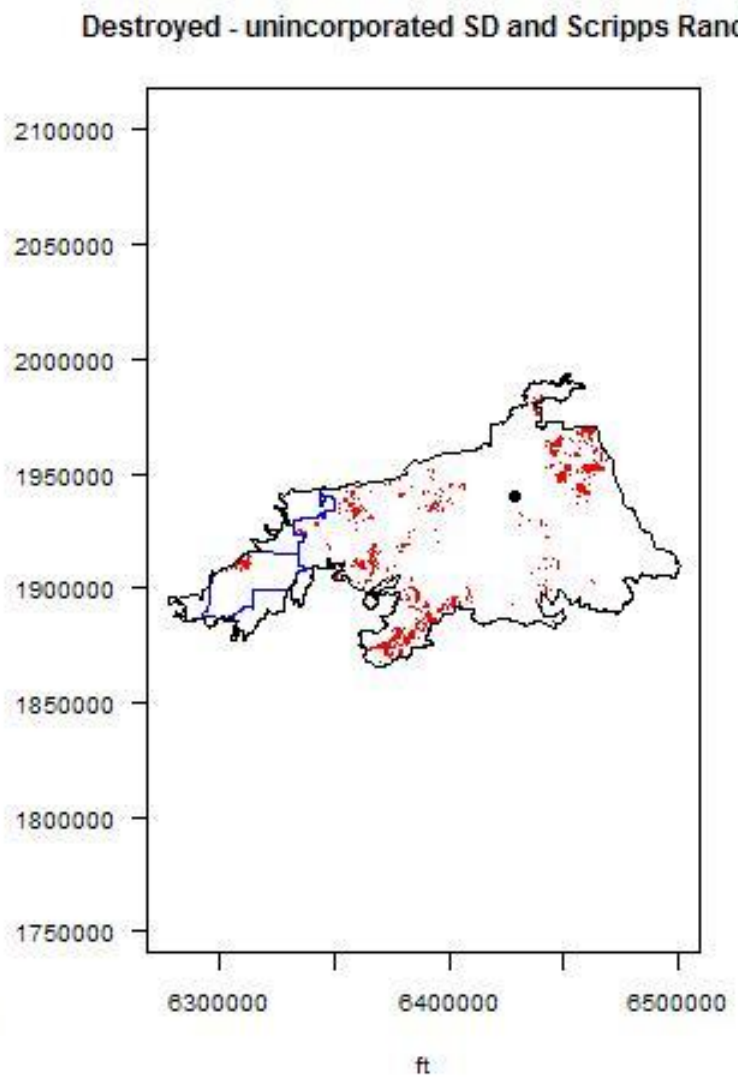
Unincorporated SD County + Scripps Ranch

Fire boundary

Locations (destroyed and not)

Rates/intensities and ratio


All Cedar houses prefire
2100000
2050000
2000000
1950000
1900000
1850000
1800000
1750000
6300000
6400000
6500000
ft


Destroyed - unincorporated SD and Scripps Ranch
2100000
2050000
2000000
1950000
1900000
1850000
1800000
1750000
6300000
6400000
6500000
ft

Looking at the continous data.

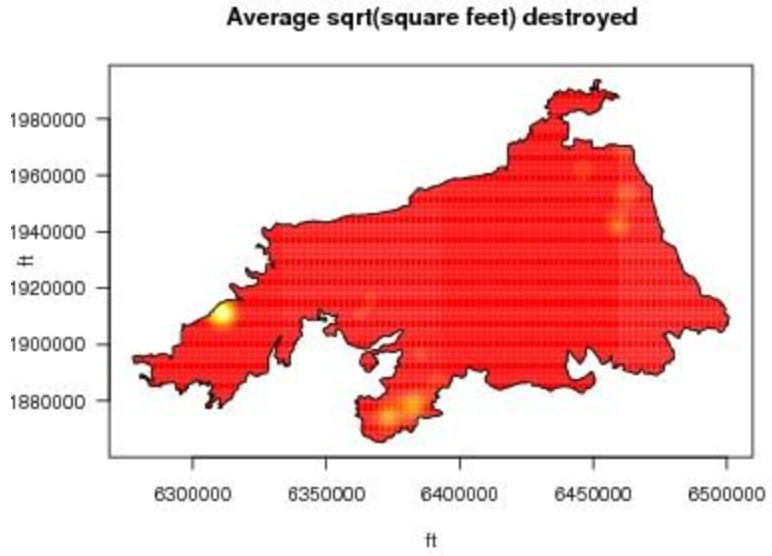
m.p.p. : area of house (square feet)

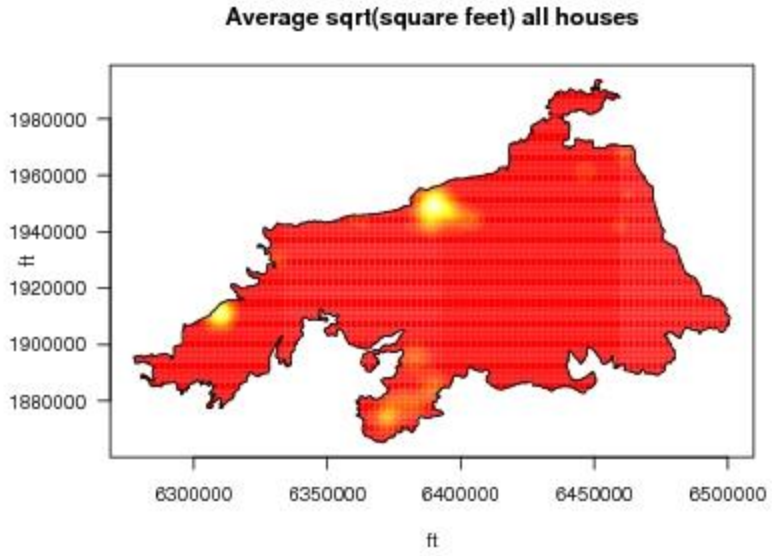
a cost proxy ($150/sqft)

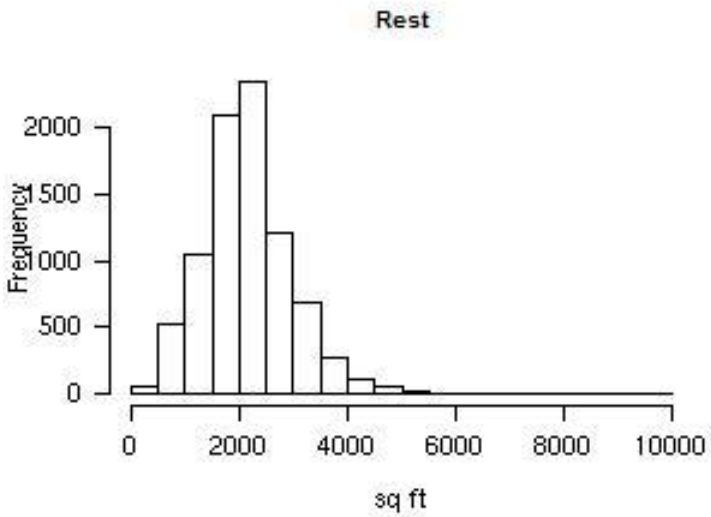
Smoothed sqrt(squared feet)

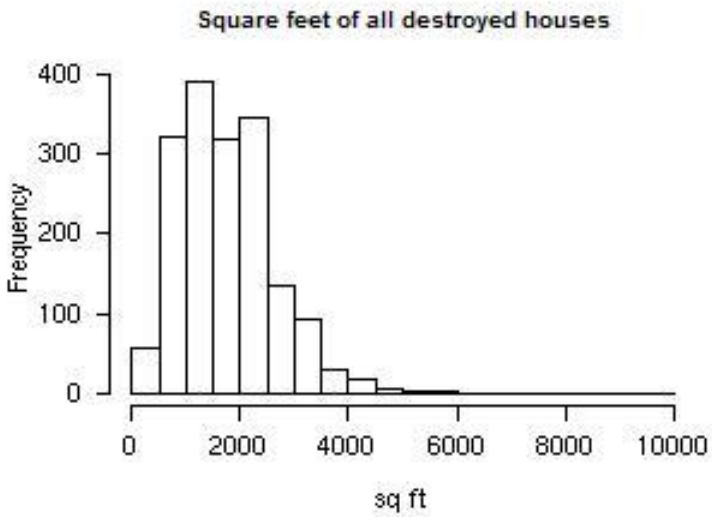
$\Sigma Z_l K(x-x_i,y-y_i)$

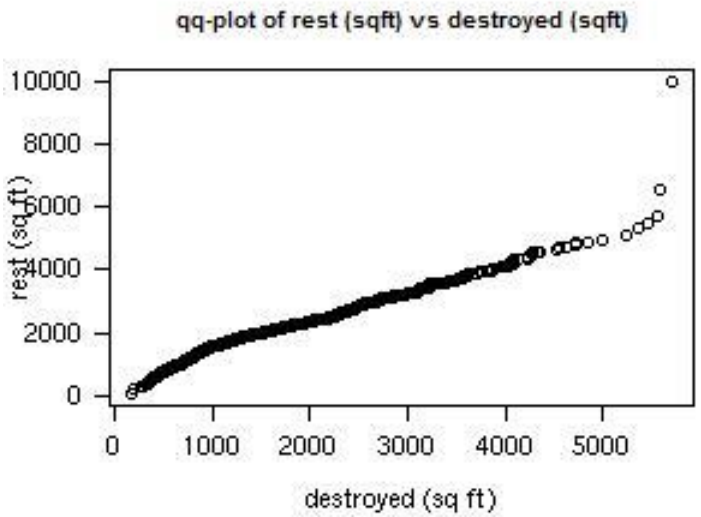
Some descriptive statistics

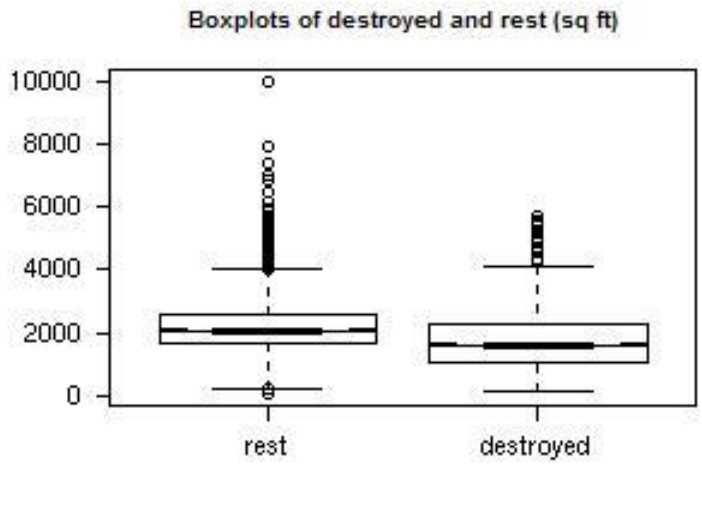

Average sqrt(square feet) destroyed
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000
6350000
6400000
6450000
6500000
ft


Average sqrt(square feet) all houses
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000
6350000
6400000
6450000
6500000
ft


Rest
Frequency
0
500
1000
1500
2000
0
2000
4000
6000
8000
10000
sq ft


Square feet of all destroyed houses
Frequency
0
100
200
300
400
0
2000
4000
6000
8000
10000
sq ft


qq-plot of rest (sqft) vs destroyed (sqft)
rest (sq ft)
0
2000
4000
6000
8000
10000
0
1000
2000
3000
4000
5000
destroyed (sq ft)


Boxplots of destroyed and rest (sq ft)
0
2000
4000
6000
8000
10000
rest
destroyed

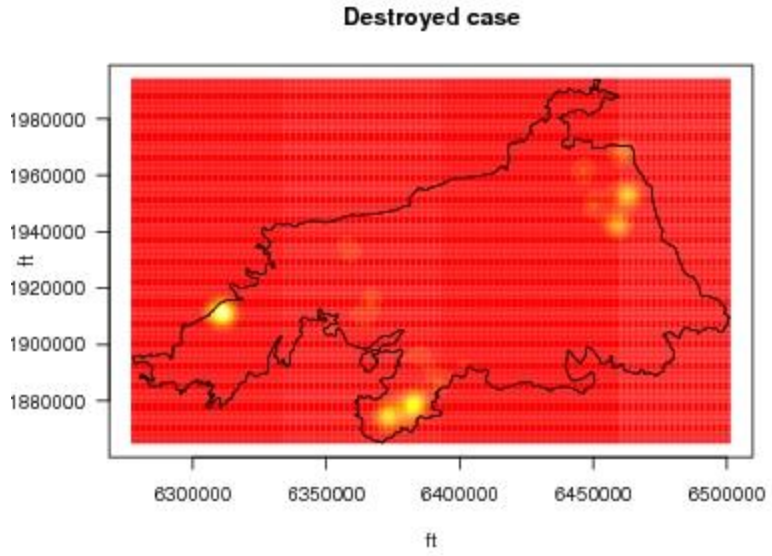
Inference results. *Point process case*.

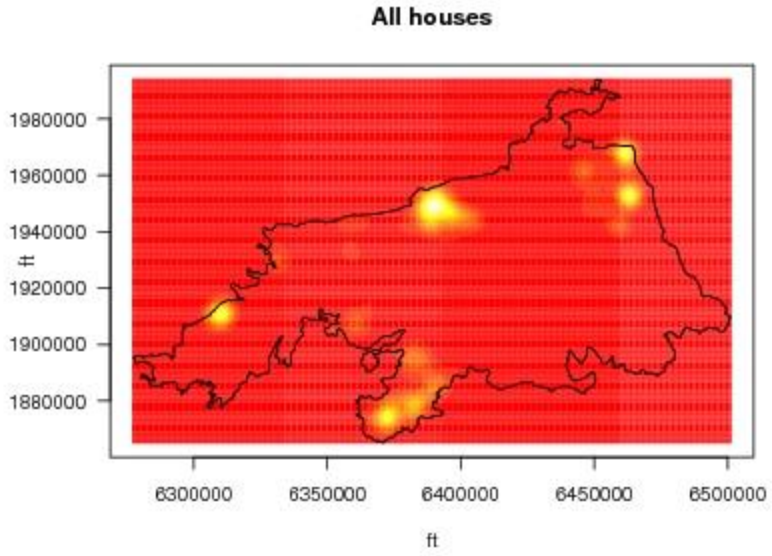
Intensity of houses at (x,y) initially $\mu_X(x,y)$

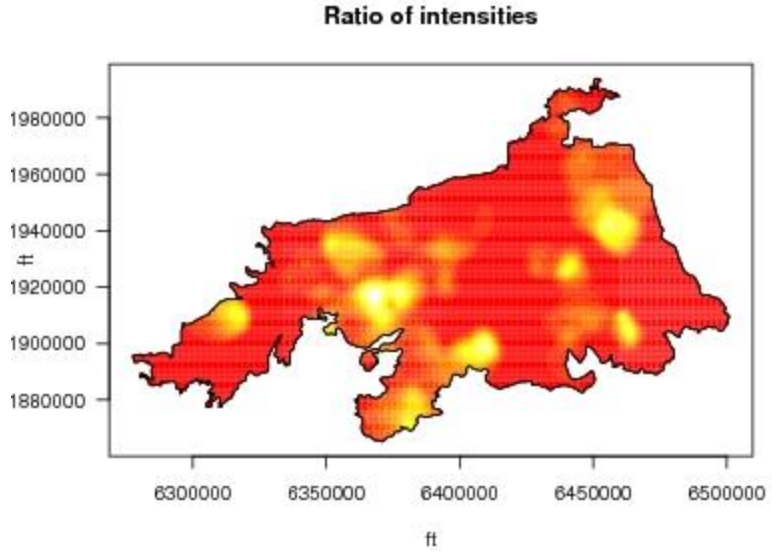
Intensity of destroyed $\mu_Y(x,y)$

$p(x,y) = \mu_Y(x,y)/ \mu_X(x,y)$

"probability" of a house's destruction


Destroyed case
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000
6350000
6400000
6450000
6500000
ft


All houses
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000
6350000
6400000
6450000
6500000
ft

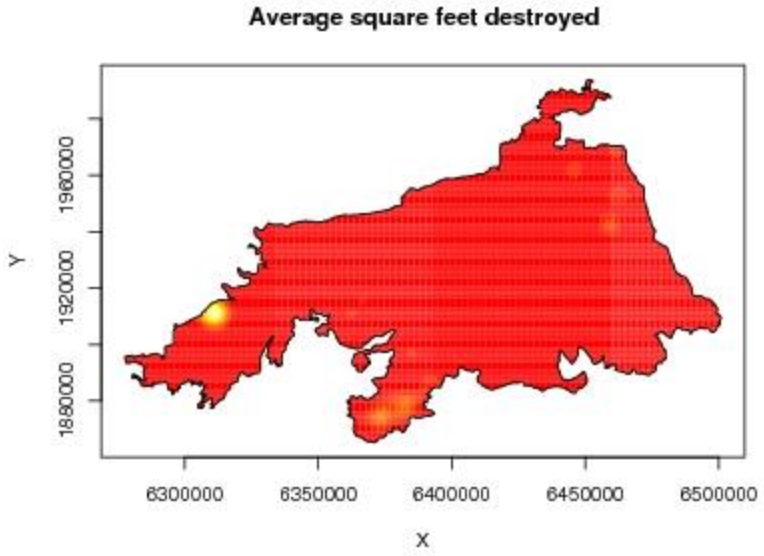

Ratio of intensities
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000
6350000
6400000
6450000
6500000
ft

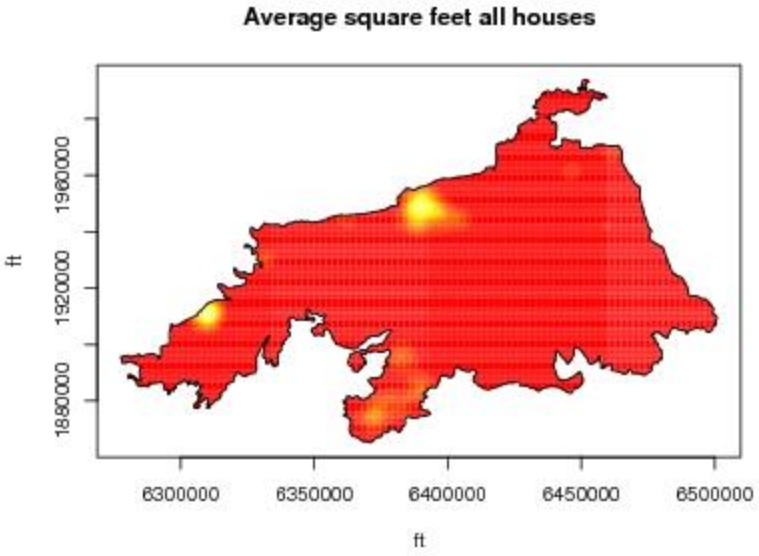
Inference results. *Continuous case.*

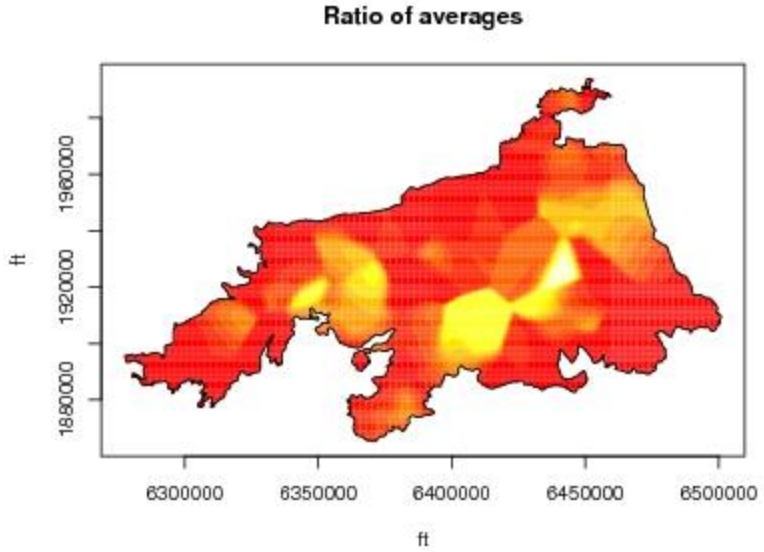
Square feet (from tax records)

Is there a difference wrt squared feet between destroyed and rest?

Estimate $v_V(x,y)/v_U(x,y)$


Average square feet destroyed
Y
1880000
1920000
1960000
6300000
6350000
6400000
6450000
6500000
X


Average square feet all houses
ft
1880000
1920000
1960000
6300000
6350000
6400000
6450000
6500000
ft


Ratio of averages
ft
1880000
1920000
1960000
6300000
6350000
6400000
6450000
6500000
ft

Does size depend on location?

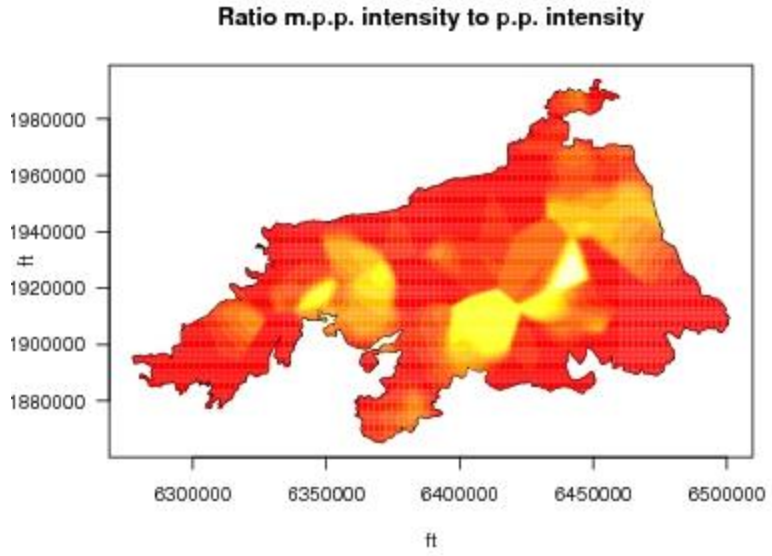
$$dN(x,y,z)/dxdydz = \Sigma\, \delta(x-x_i, y-y_i, z-z_i)$$

$$\Sigma\, z_i\, \delta(x-x_i, y-y_i, z-z_i)$$

If Z independent of p.p. {X(x,y)}, average satisfies

$$\gamma(x,y,z) = \gamma_1(x,y)\, \gamma_2(z)$$

Consider $\gamma(x,y,z)/\gamma_1(x,y)$


Ratio m.p.p. intensity to p.p. intensity
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000
6350000
6400000
6450000
6500000
ft

Explanatories.

Vegetation type (15 categories)

Slope

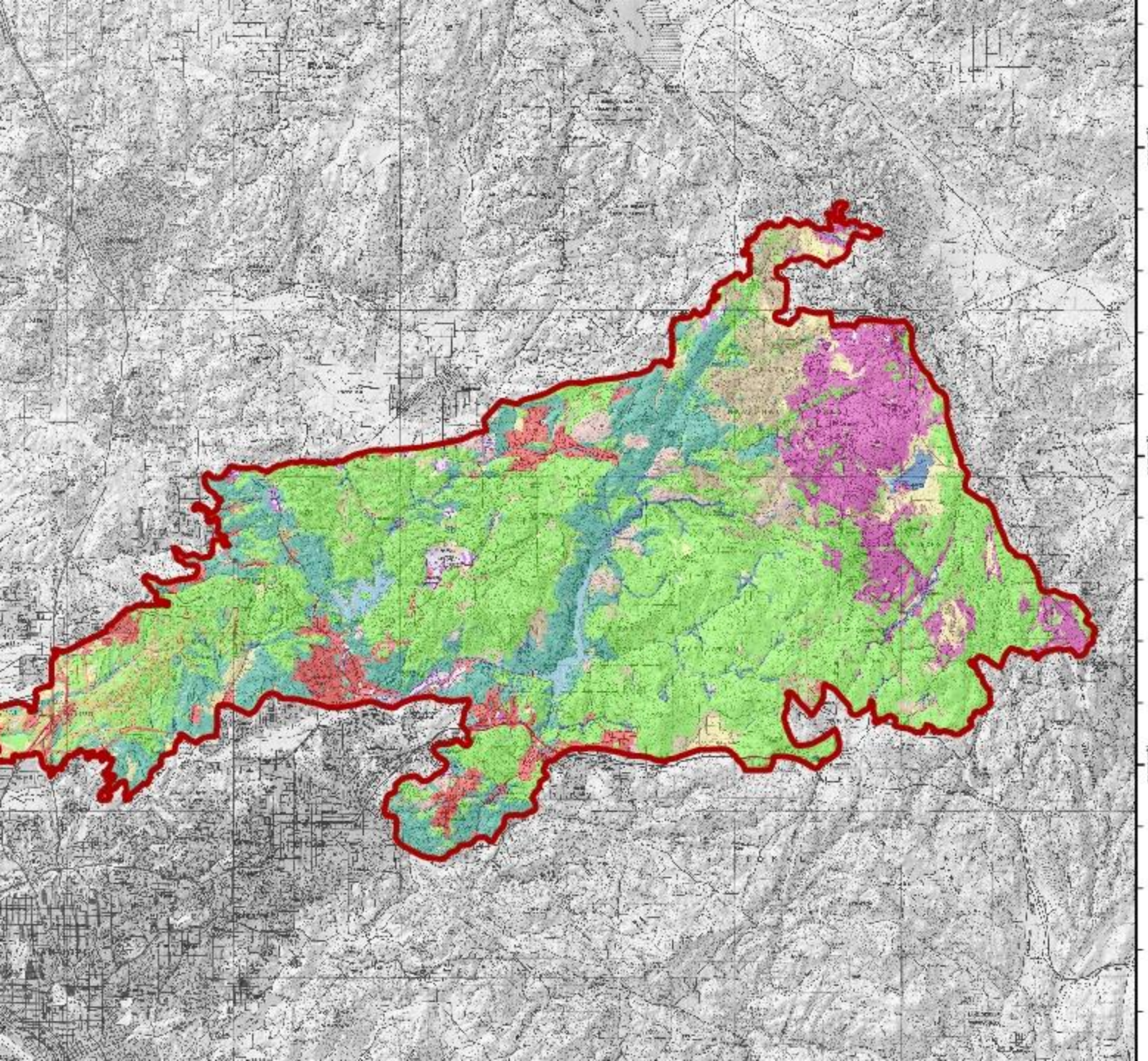
Assessed improvement value

Destroyed

Structure location

Square feet

…



2003
So Cal Fires

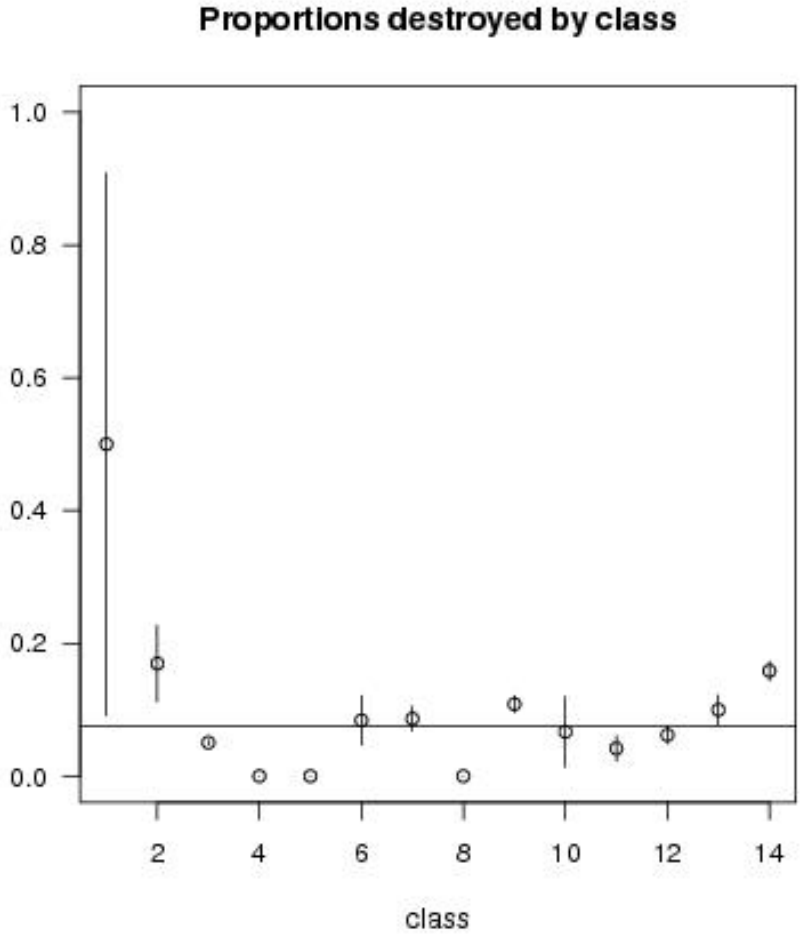
Cedar

October 2003

Pre-Fire Vegeta

- Agriculture
- Chaparral
- Coastal Sage Scrub
- Coastal Sage/Chaparral S
- Developed
- Disturbed
- Eucalyptus Woodland
- Forest
- Grasslands/Meadows/Ver
- Open Water
- Riparian
- Unvegetated
- Scrub
- Woodland
- Fire Perimeter

The data represented in this map were gathered from m
which may vary in accuracy, scale and date. This map
only and not an offical declaration of boundaries or own



14. Forest (2637)

13. Woodland (732)

12. Riperian (1421)

11. Grassland/Meadows/Vernal pools (456)

10. Coastal sage/Chaparral scrub (90)

9. Chaparral (2554)

8. Scrub (2)

7. Coastal sage scrub (959)

6. Agriculture (226)

5. Unvegetated (1)

4. Open water (13)

3. Developed (13353)

2. Disturbed (171)

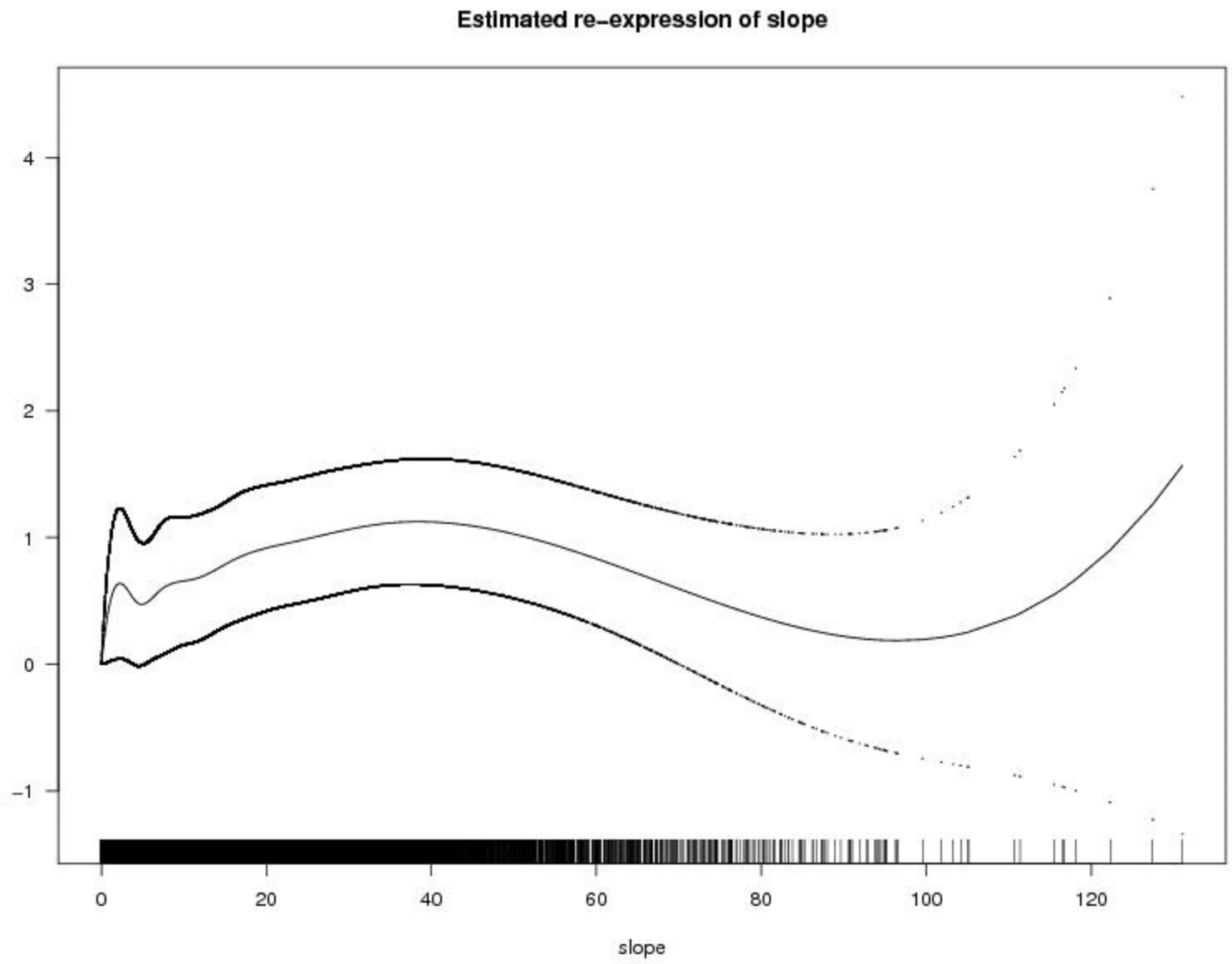
1. Eucalyptus woodland (6)

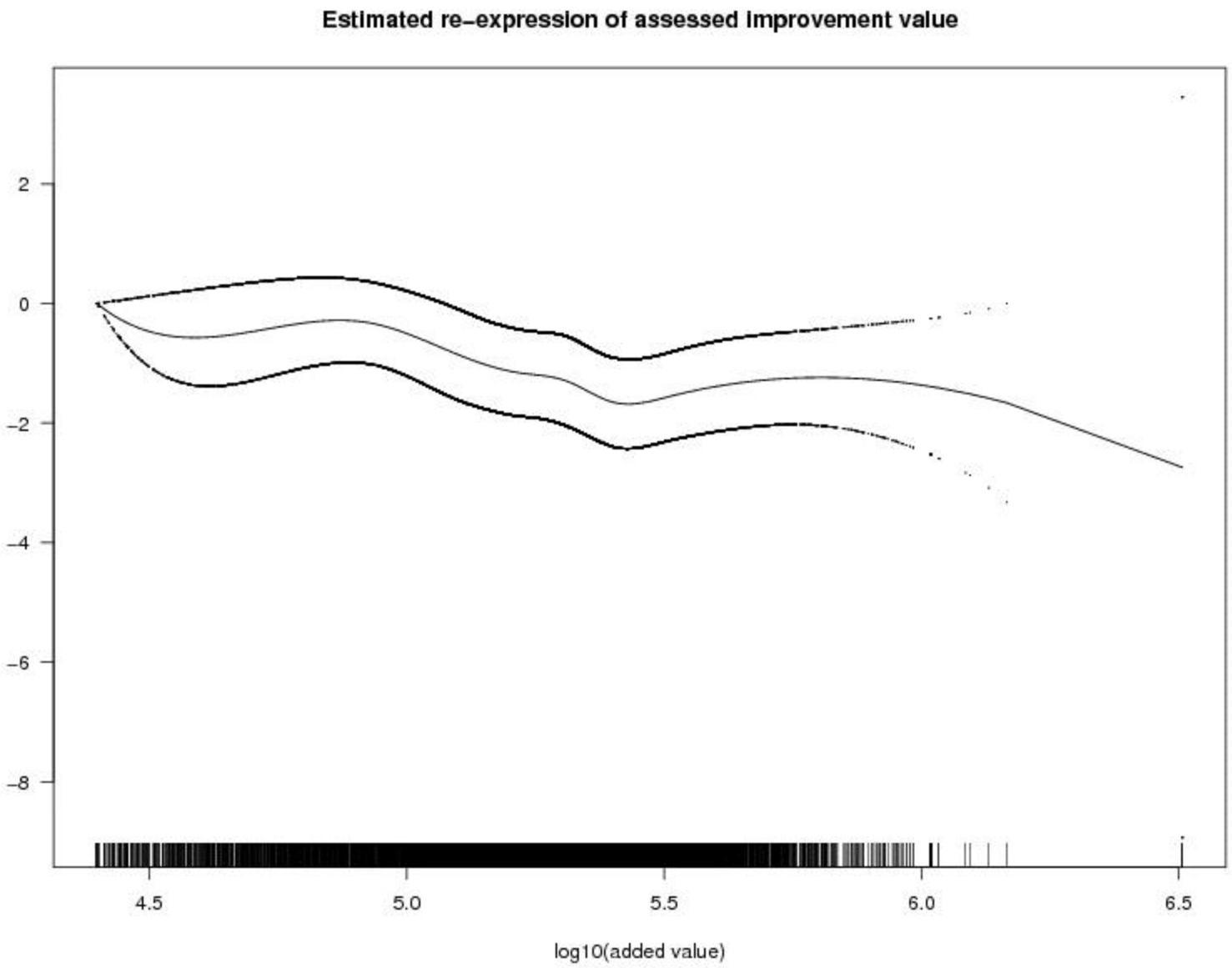
Logit-gam model results.

Logit{Prob[destroyed|explanatories]}

= γ(s) with s slope

γ smooth


Estimated re-expression of slope
4
3
2
1
0
−1
0
20
40
60
80
100
120
slope


Estimated re−expression of assessed improvement value
2
0
−2
−4
−6
−8
4.5
5.0
5.5
6.0
6.5
log10(added value)

Spatial-temporal results.

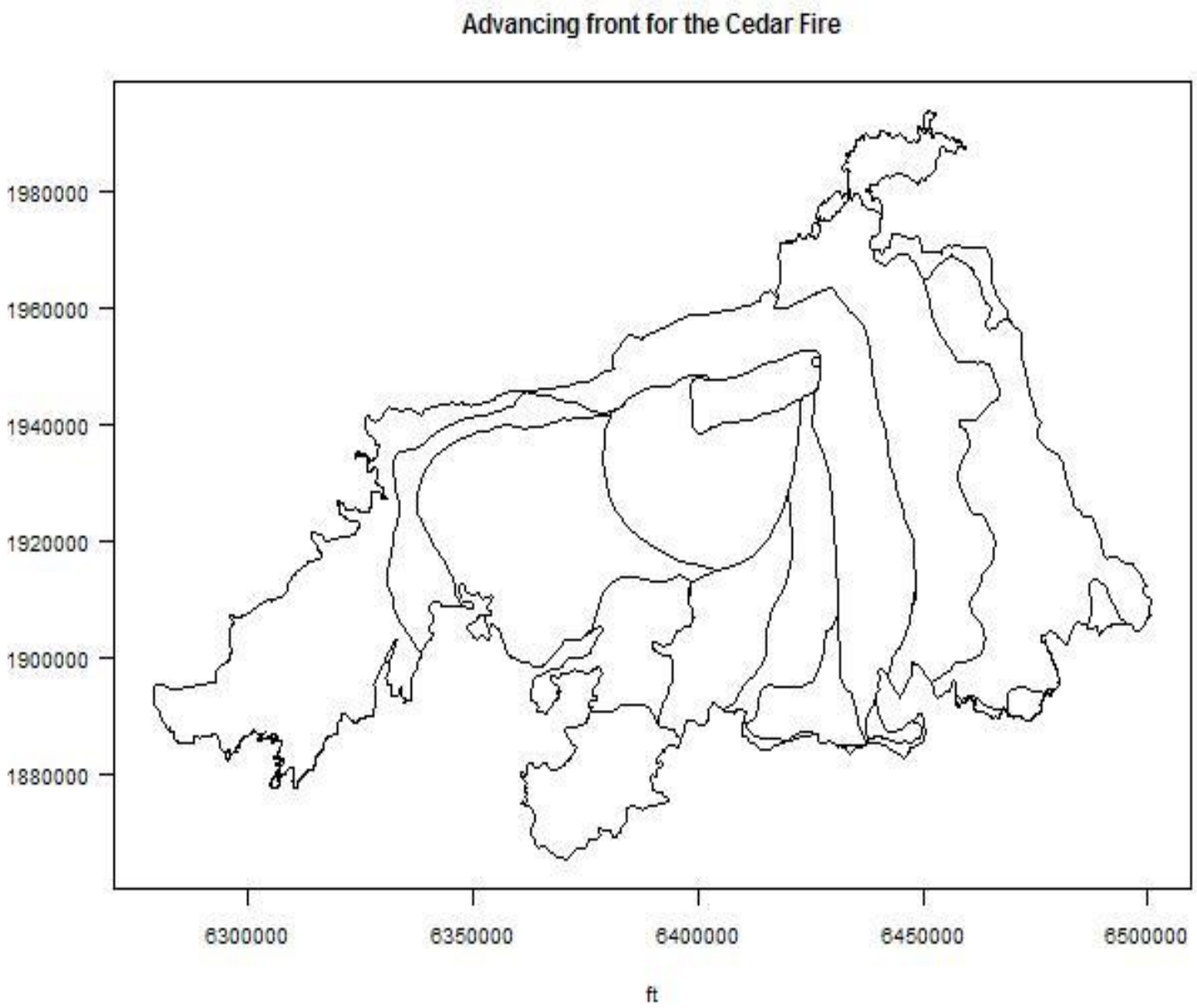
polygons

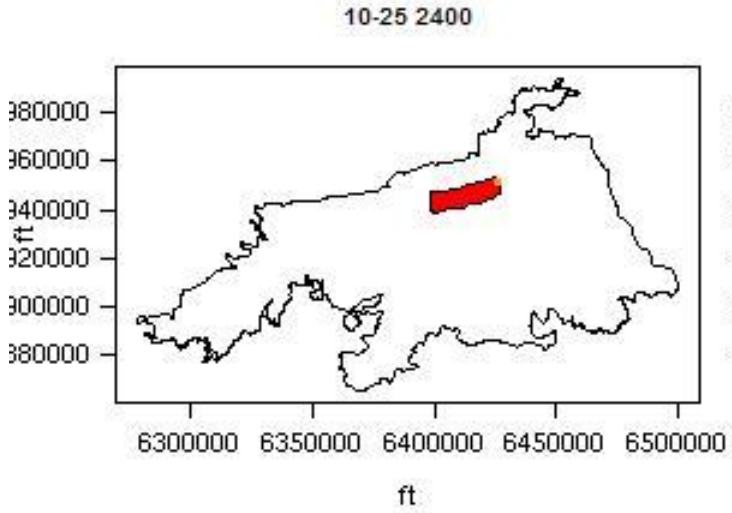
wavefront

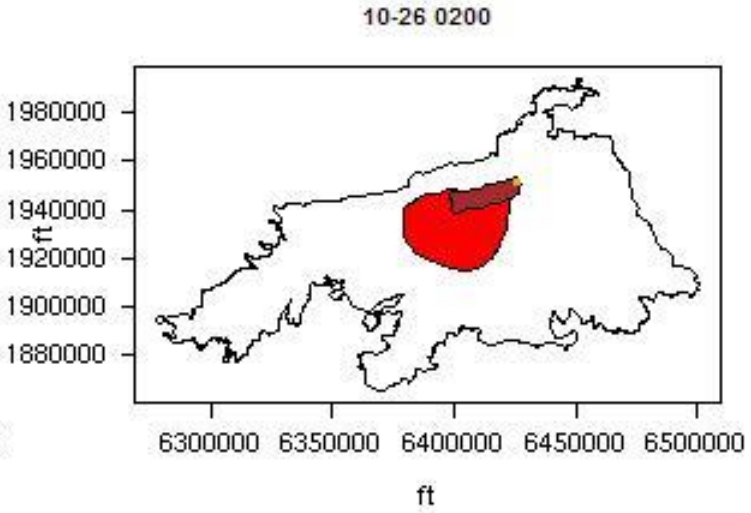
How quantities in polygons depend on time

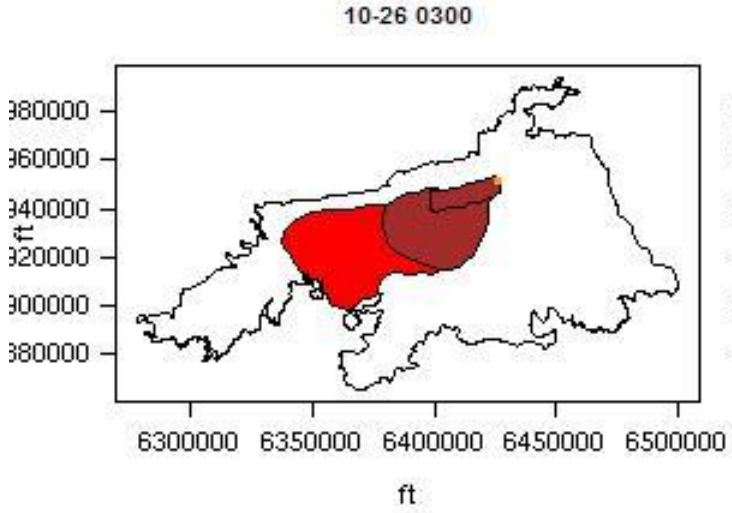
Time defined as interval from midnight 25 October to last fire boundary

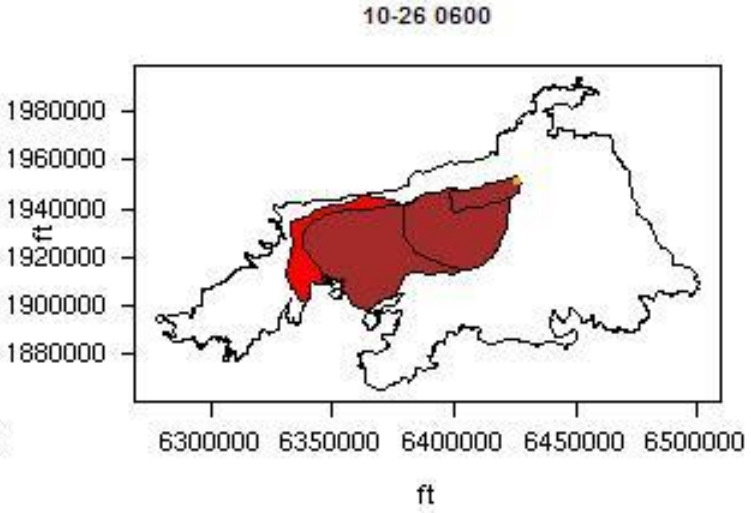
Observed fire boundaries

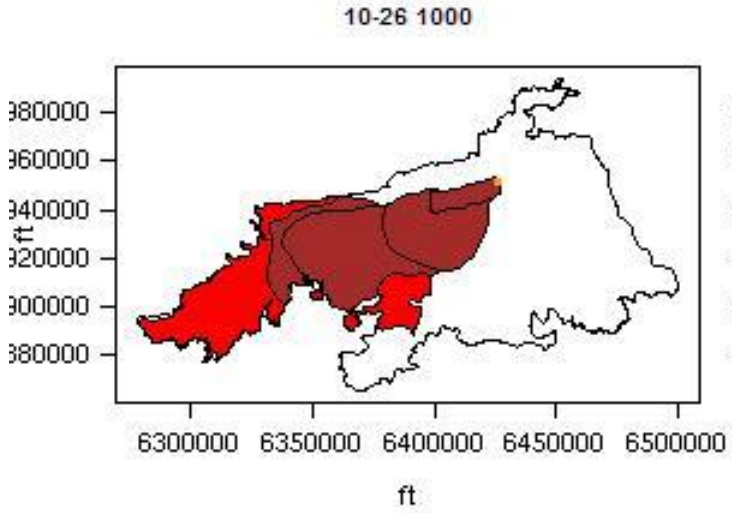


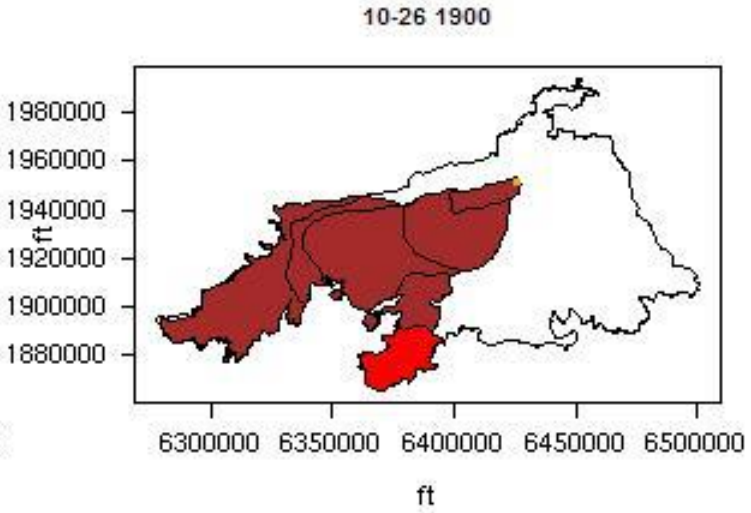

10-25 2400
ft
6300000 6350000 6400000 6450000 6500000
ft

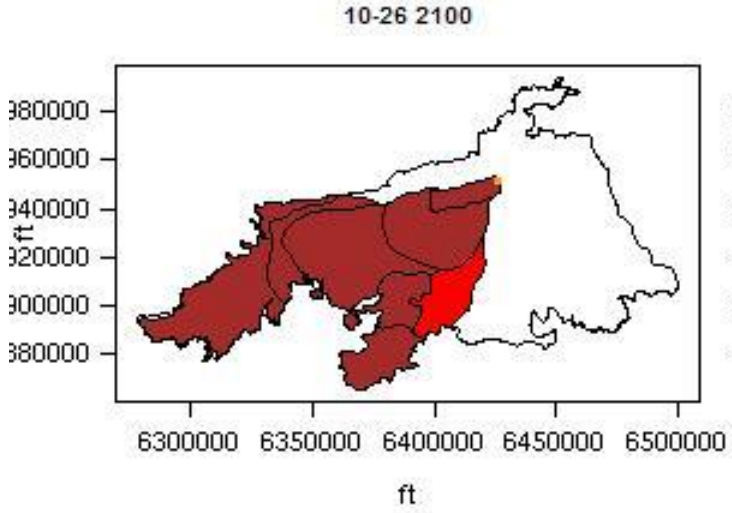

10-26 0200
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft

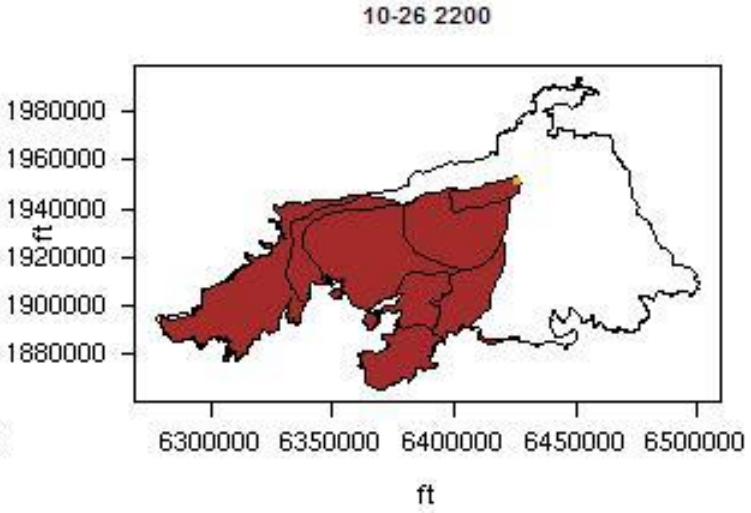

10-26 0300
ft
6300000 6350000 6400000 6450000 6500000
ft

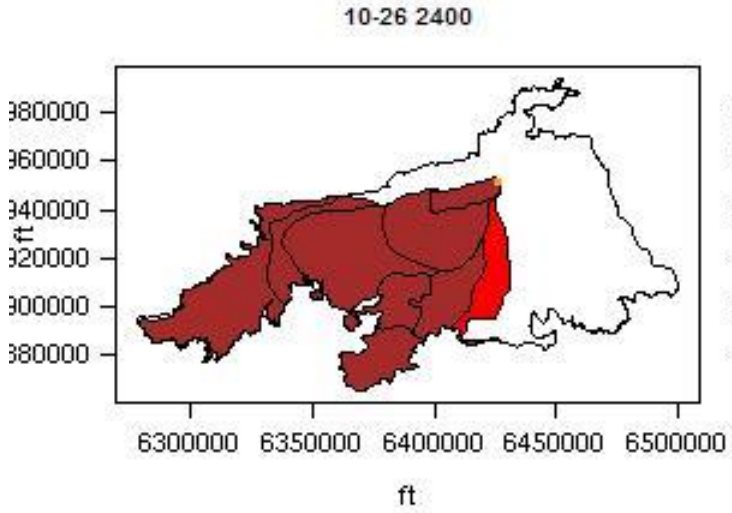

10-26 0600
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft

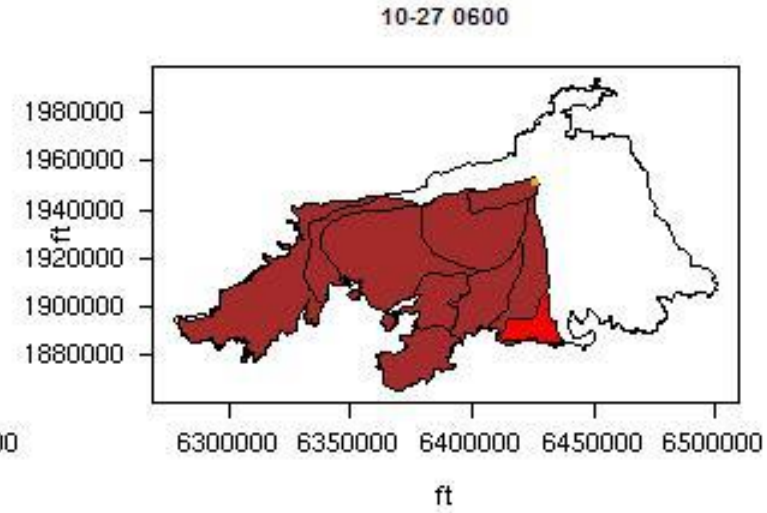

10-26 1000
ft
6300000 6350000 6400000 6450000 6500000
ft

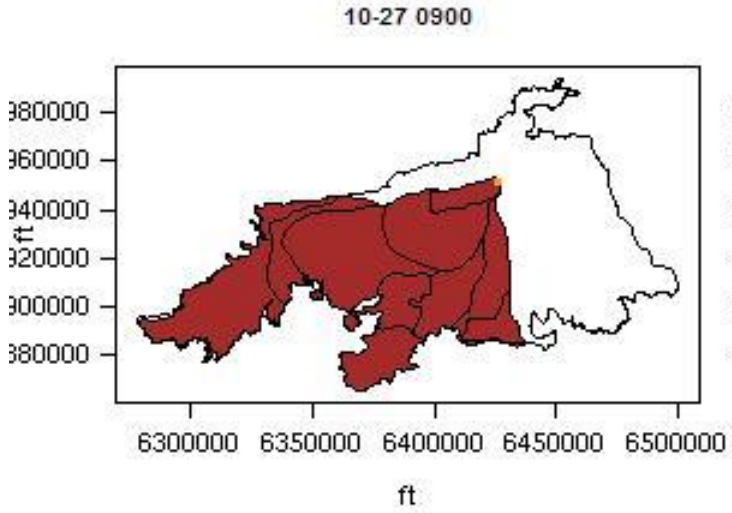

10-26 1900
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft

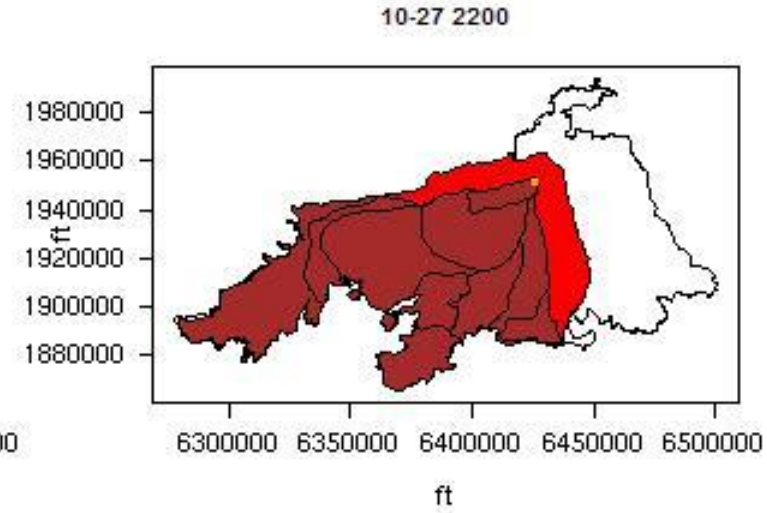

10-26 2100
ft
6300000 6350000 6400000 6450000 6500000
ft

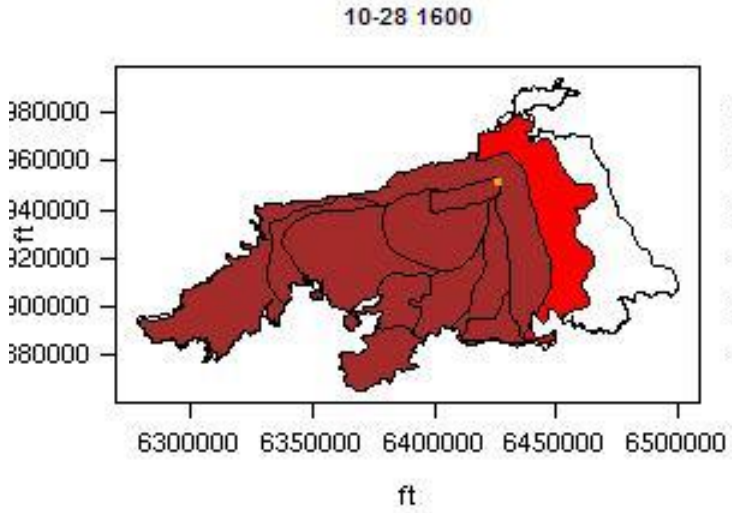

10-26 2200
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft

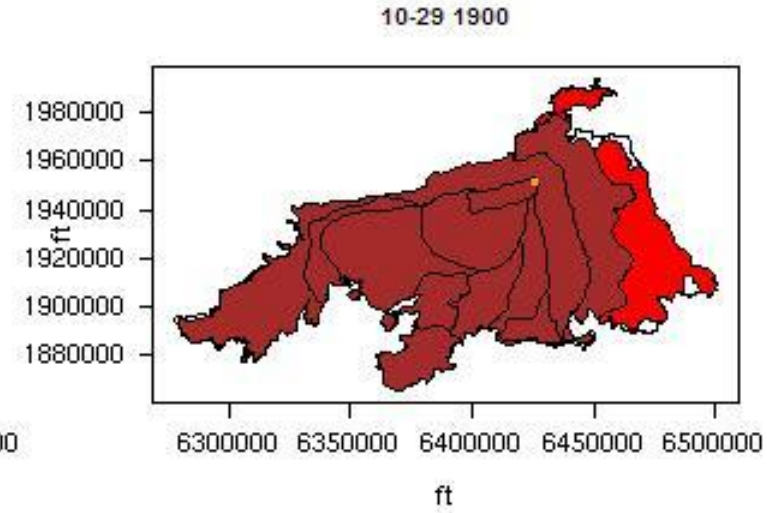

10-26 2400
980000
960000
940000
920000
900000
880000
ft
6300000 6350000 6400000 6450000 6500000
ft

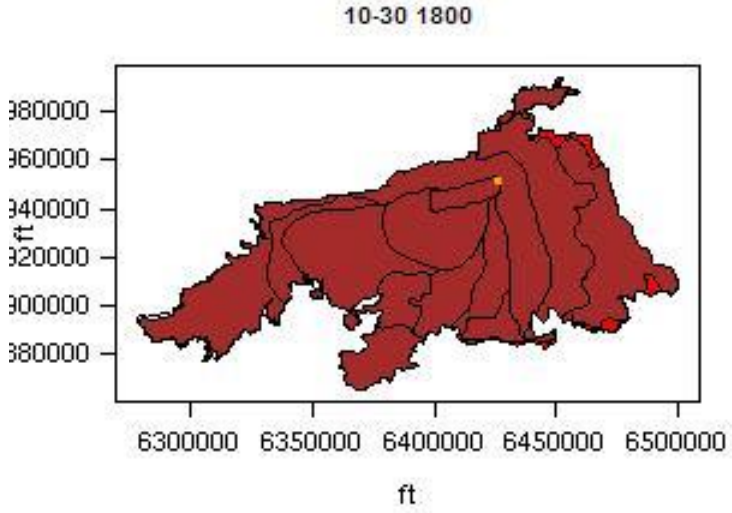

10-27 0600
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft

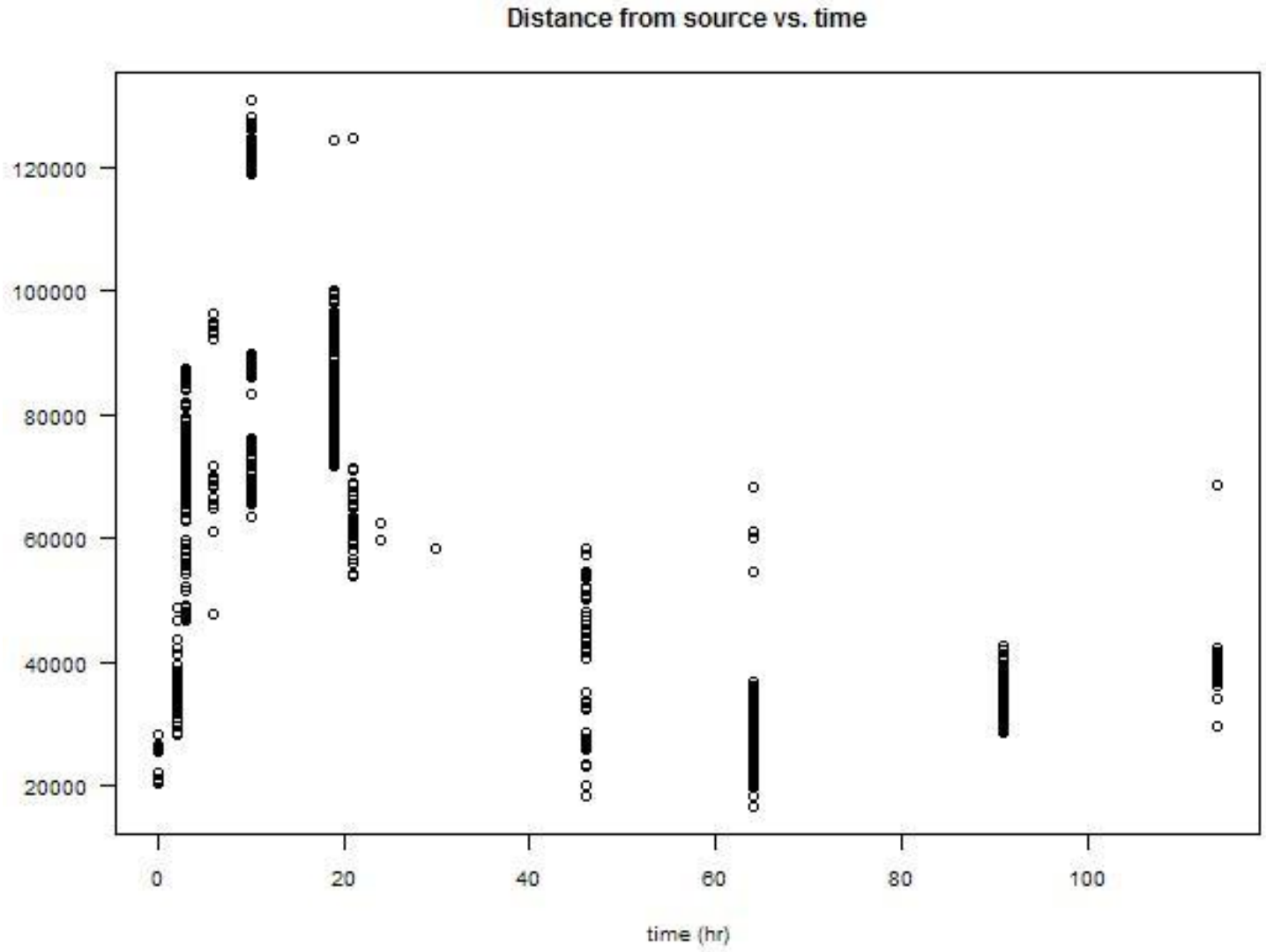

10-27 0900
980000
960000
940000
920000
900000
880000
ft
6300000 6350000 6400000 6450000 6500000
ft

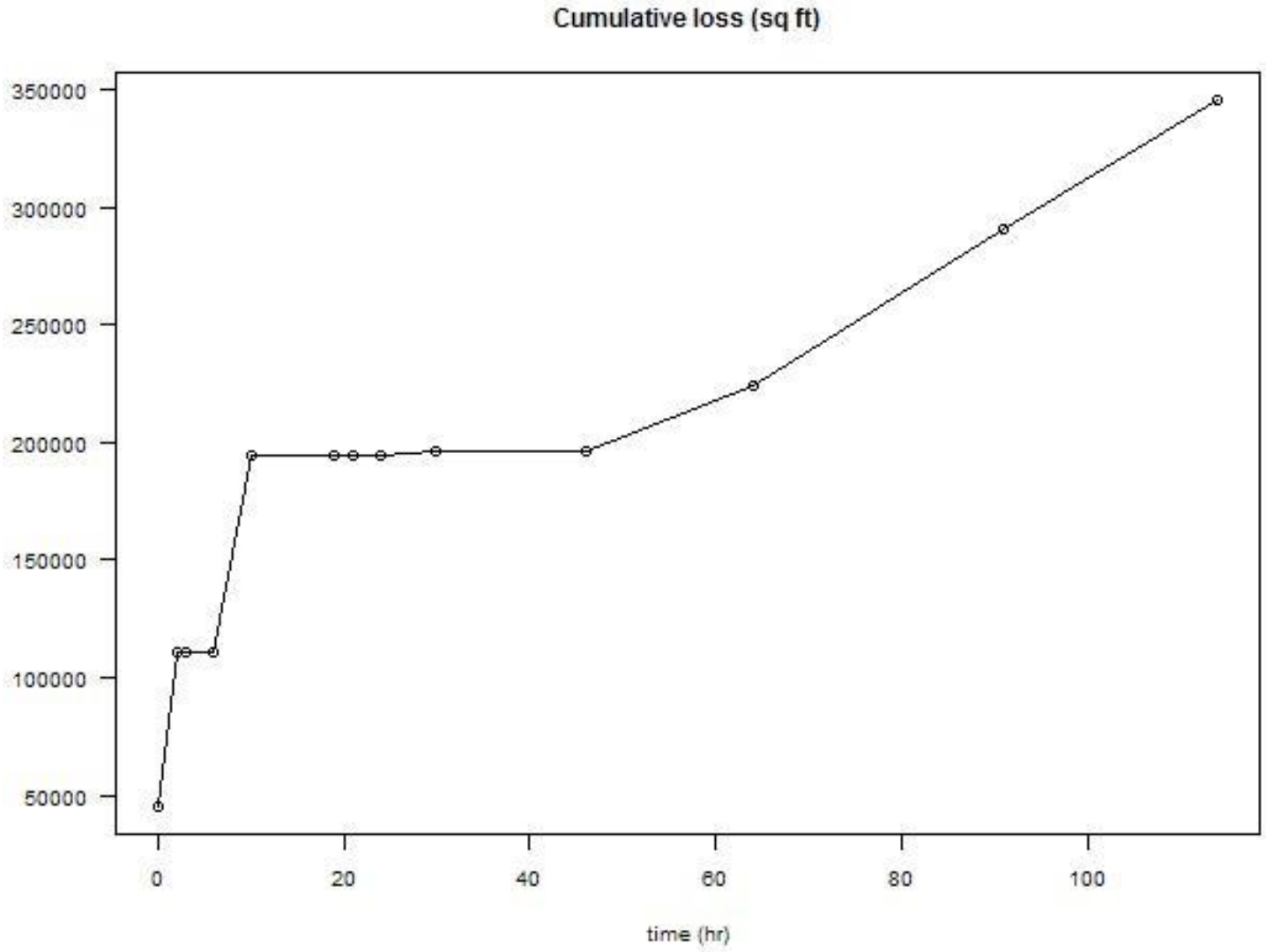

10-27 2200
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft


10-28 1600
980000
960000
940000
920000
900000
880000
ft
6300000 6350000 6400000 6450000 6500000
ft


10-29 1900
1980000
1960000
1940000
1920000
1900000
1880000
ft
6300000 6350000 6400000 6450000 6500000
ft


10-30 1800
980000
960000
940000
920000
900000
880000
ft
6300000 6350000 6400000 6450000 6500000
ft


Distance from source vs. time
120000
100000
80000
60000
40000
20000
0
20
40
60
80
100
time (hr)


Cumulative loss (sq ft)
350000
300000
250000
200000
150000
100000
50000
0
20
40
60
80
100
time (hr)

Economic Valuation: $$$

- Key distinction:
  - Social Cost (public goods)

    (e.g., vegetation lost or air pollution)
  - Private Cost (private goods)

    (e.g., properties or assets destroyed)
- Short-run vs. Long-run Effects

Social cost - loss of Chaparral

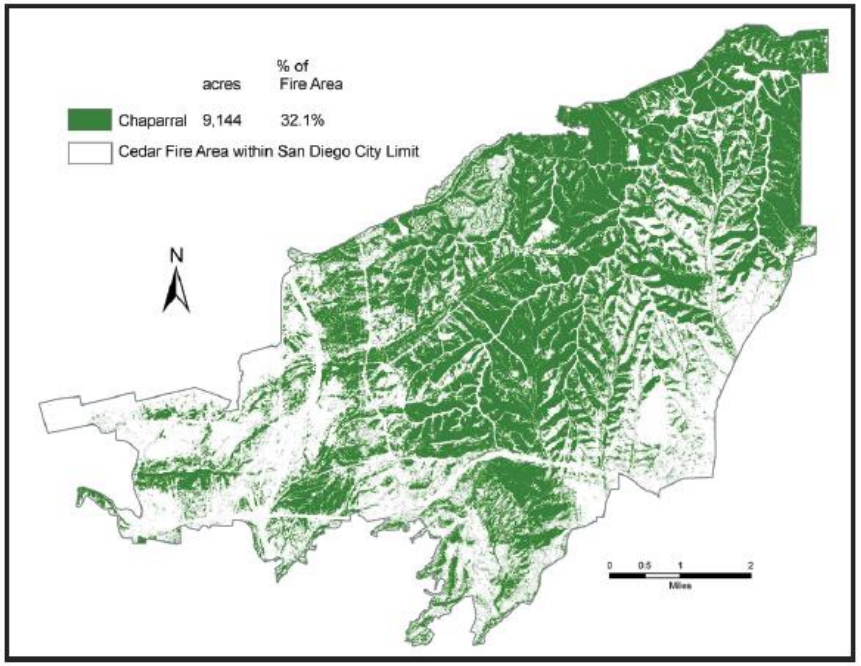


Figure 2: Chaparral coverage in the Cedar Fire - pre fire 2002

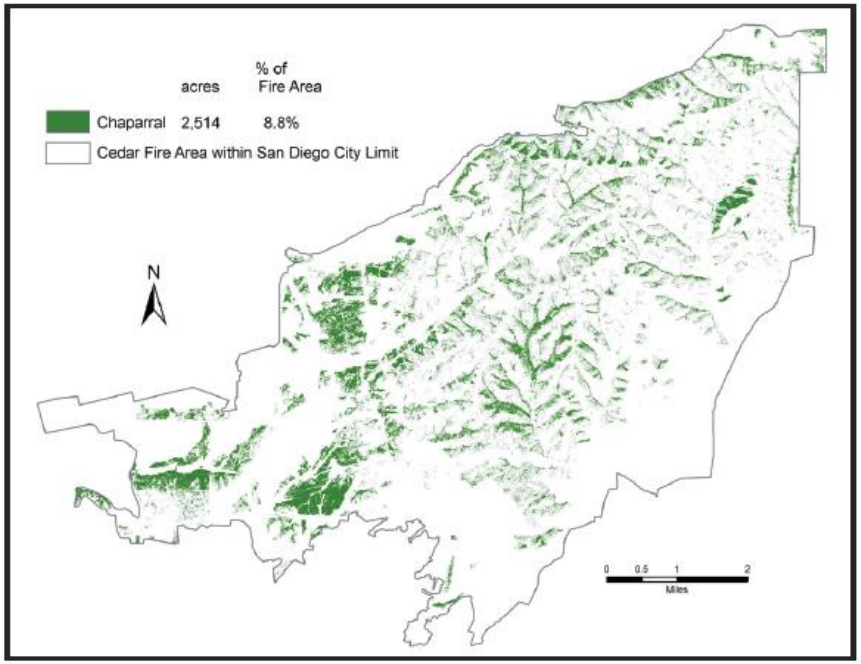
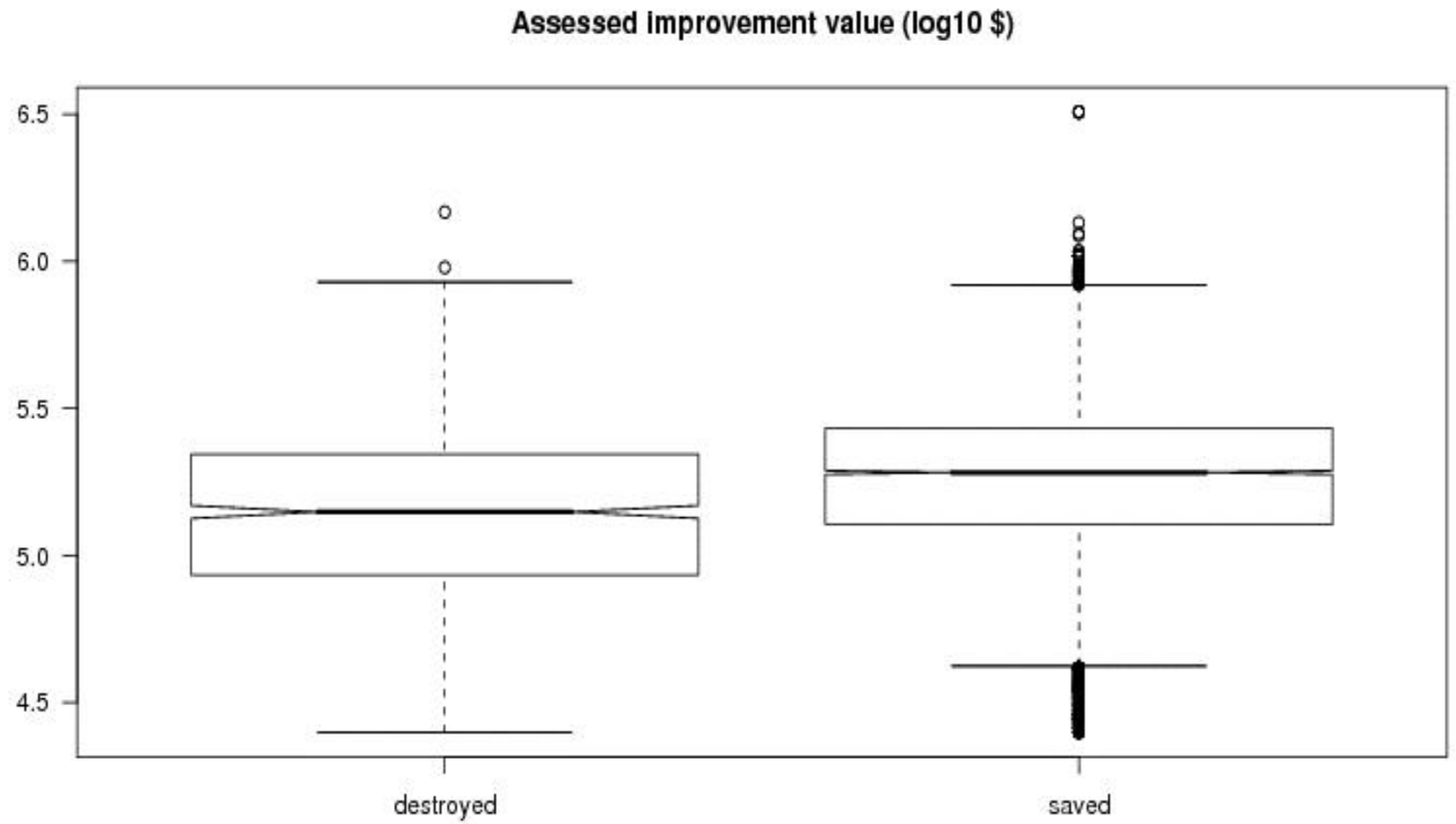


Figure 3: Chaparral coverage in the Cedar Fire - post fire 2004

Example of non-market valuation:

- Stormwater runoff increased by 12 million cubic feet.
- Cost of retaining is estimated at $25 million dollars.
- Underestimation: This reflects only one dimension of value.



- Downward trend in chance of destruction as assessedvalue increases.

Other thoughts.

Damaged houses

Other explanatories

Other models

Other fires

Spatial correlation

Uncertainties

…

Discussion.

Limitations

"they are 'messy' datasets and do require a bit of massaging to make sense" … "the damage assesment we performed … was a rapid assessment. There were 18 … teams. … we used a variety of GPSs of varying accuracy. The individual team members also had varying degrees of competency." J. Batchelor (SD County)

Just one fire, lurking variables/proxies

GISs – Cedar fire areal time success for the GIS industry

Can grab shapefile data for R analyses

Would robust/resistant methods have helped?

Summary.

A work in progress, a story

Difficulties of getting, cleaning and employing data

Used statistical package, R, with Sangis data layers

Acknowledgements.

SanGIS, SDSU, SDCounty, SDFoundation, SDCity, …

A. Ager (USFS), J. Benoit (USFS), C. Hunter (Rancho Santa Fe), D. Sapsis (State of Ca), P. Spector (UCB), C. Westling (SD County), K. Wright (USFS), J. Batchelor (SD County), R. Martin (SD County), M-H. Tsou (SDSU), D. Gilmore (SD County),…