Pacific Southwest Research Station

Some Statistical Issues in Predicting Wildland Fire Risk

by

Haiganoush K Preisler

May 23-28 2005

Toronto, Canada

THE FIELDS INSTITUTE

FOR RESEARCH IN MATHEMATICAL SCIENCES

Outline

Fire risk predictions at

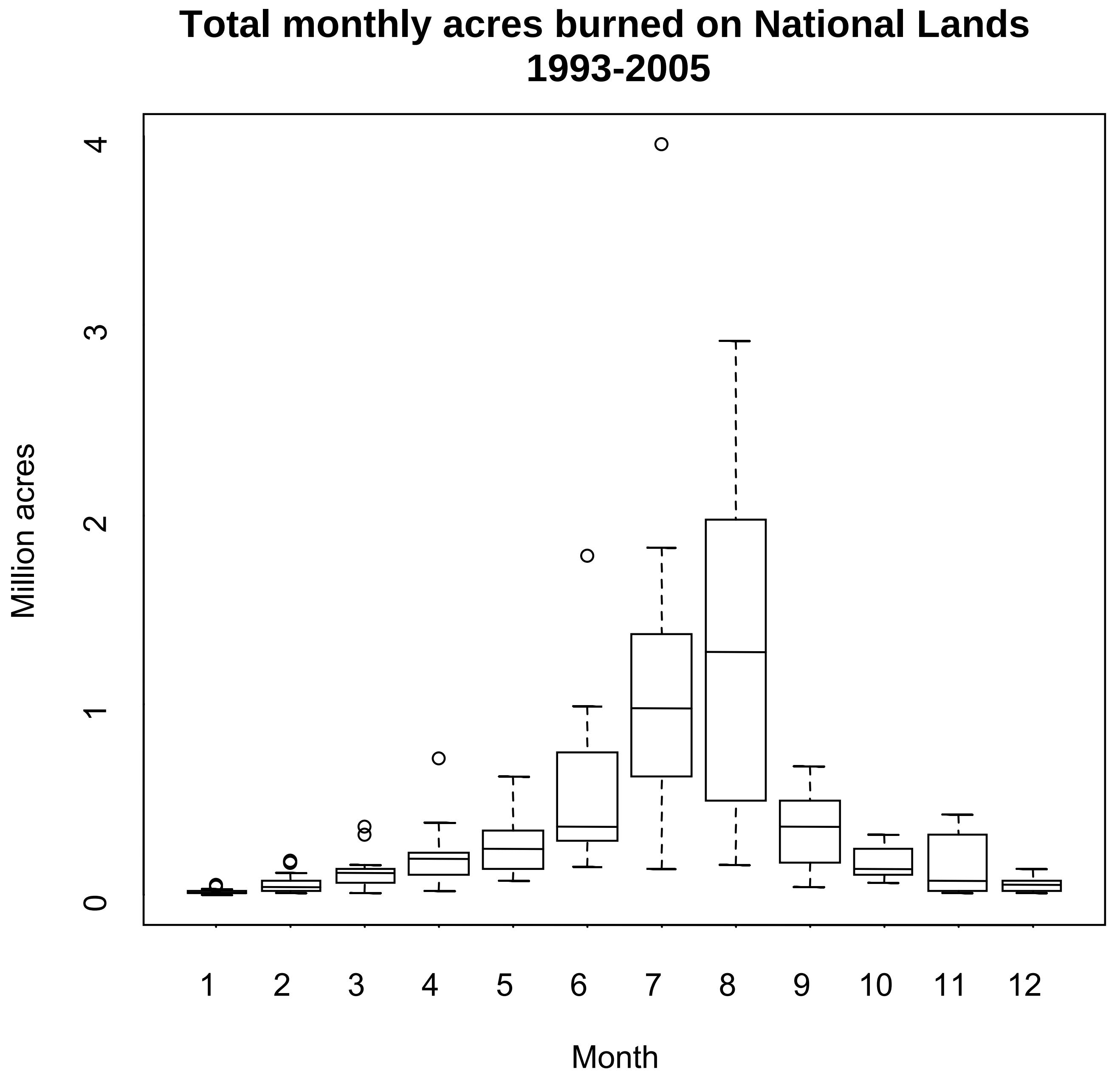
1) National level
2) climate divisions level
3) 1 km square grid level
4) individual fire level

1) National Level

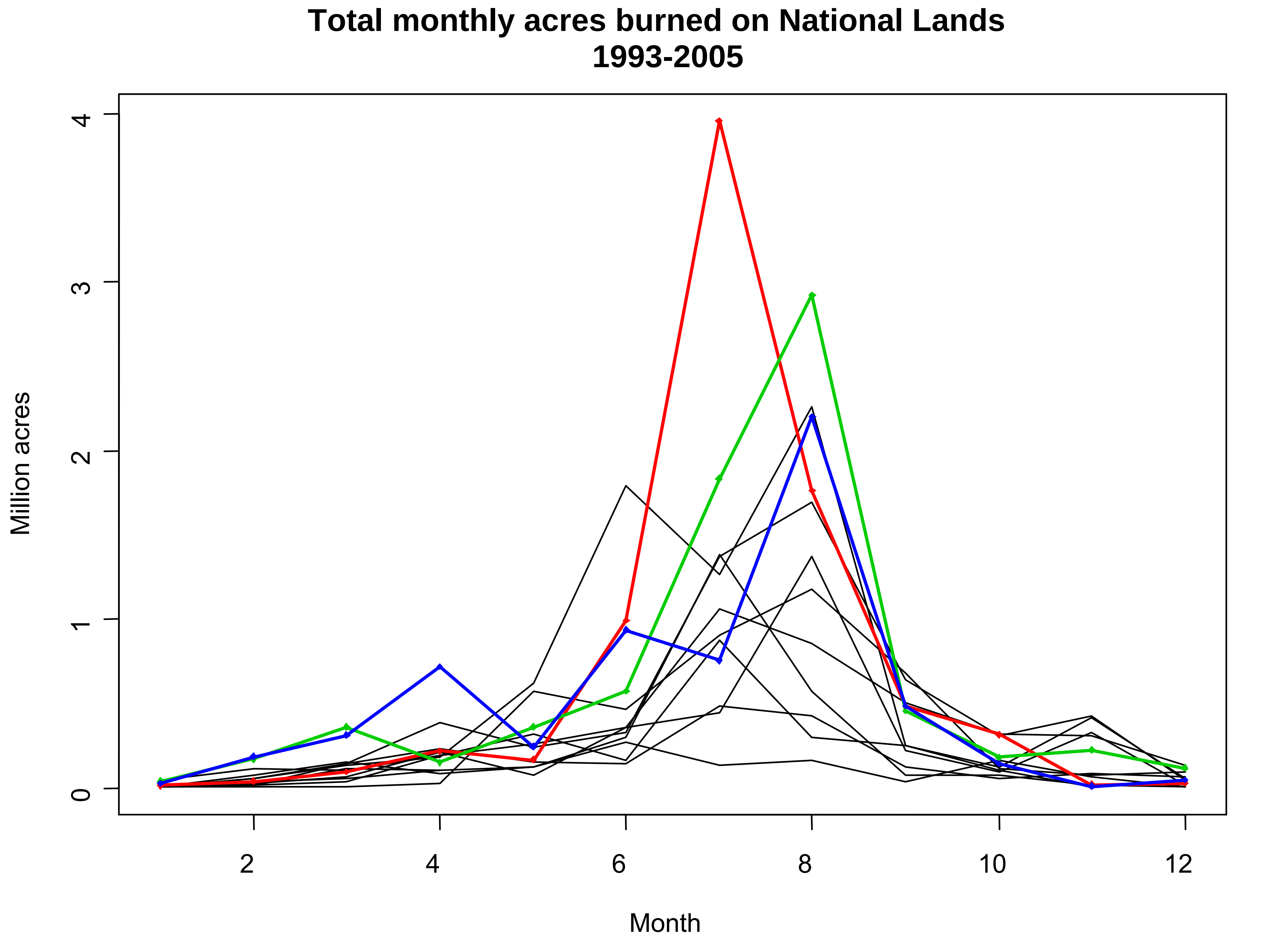
Total monthly acres burned on National Lands

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
2	1	2192	4955	7985	27754	10448	4159	6586	40257	49252	10079	18777	15386	9735
3	2	36566	11331	19622	184276	75925	17803	59030	168710	113954	55487	28893	39787	18414
4	3	55939	11662	141227	312205	154350	35030	139263	360600	103341	135186	67971	95457	115749
5	4	105538	24881	231376	719952	80469	192480	384814	148948	194693	180158	213040	219470	123791
6	5	126168	573817	150097	245198	122365	317754	240733	361570	259890	622441	76238	166299	
7	6	274174	462372	148486	936702	295471	167644	333807	573856	359202	1791865	360104	993709	
8	7	133467	904041	484900	756031	1378817	879843	1377181	1834991	450789	1262447	1057950	3954566	
9	8	160463	1178107	429339	2198280	569121	301459	1697388	2922752	1373747	2265935	860156	1762367	
10	9	39890	678772	121912	484817	71878	255653	639006	455226	225048	255527	507110	481982	
11	10	166767	112352	59948	143435	71075	102449	309467	181296	98886	128544	321513	315811	
12	11	60846	71542	84633	8513	19104	7331	426407	226565	326336	417180	314034	19087	
13	12	11774	92756	64573	45298	7936	48099	48294	118722	15087	58130	133437	30610	
14														

Thomas A. Wordell
Wildland Fire Analyst
NICC - Predictive Services Group



Total monthly acres burned on National Lands
1993-2005
Million acres
0
1
2
3
4
1
2
3
4
5
6
7
8
9
10
11
12
Month



Total monthly acres burned on National Lands
1993-2005
Million acres
0
1
2
3
4
2
4
6
8
10
12
Month

Two-way array

Robust techniques

1) Least absolute residuals (L1 solution)

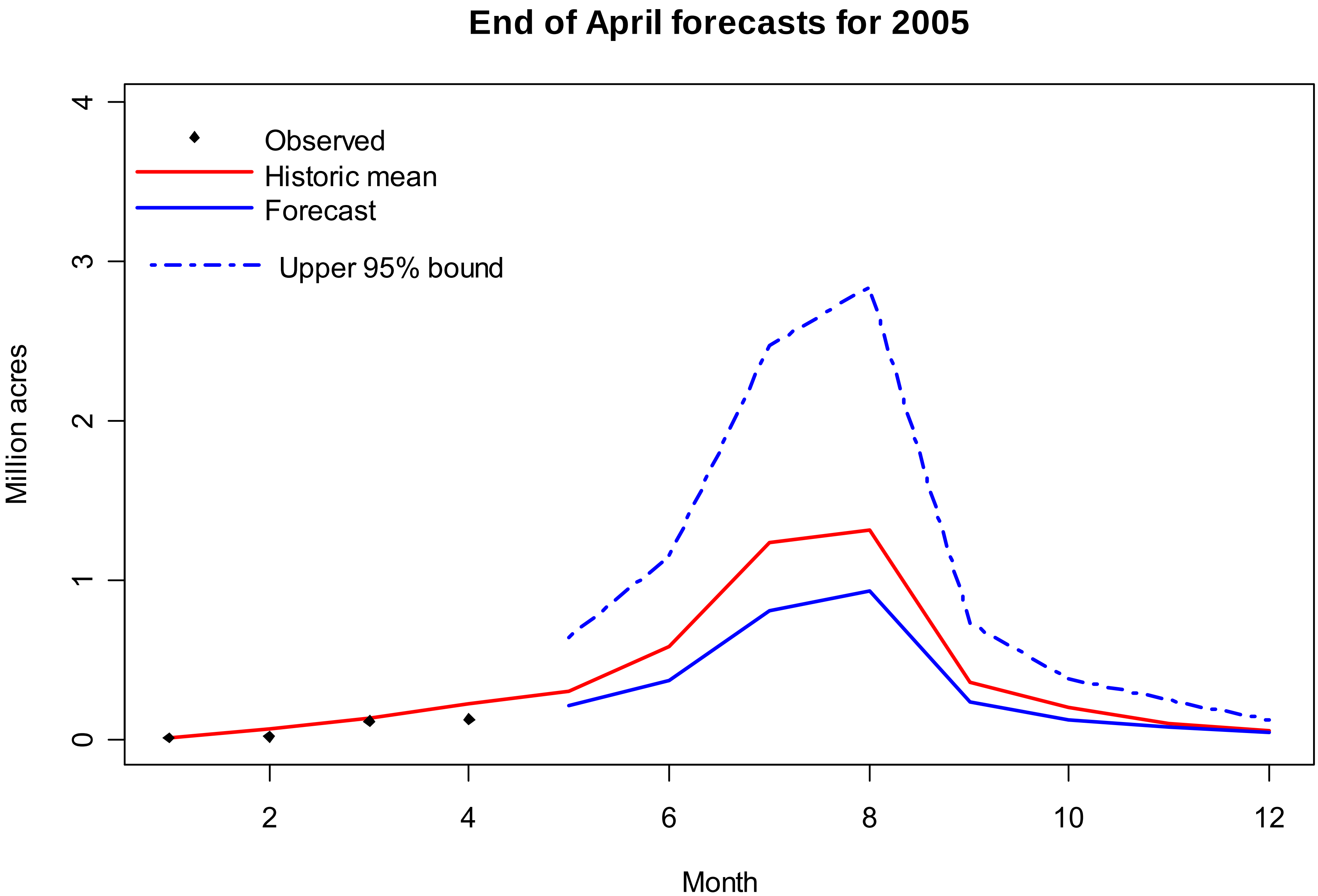
$$\min_{\mu,\alpha,\beta} \sum_{i,j} \left| y_{ij} - \mu - \alpha - \beta \right|$$

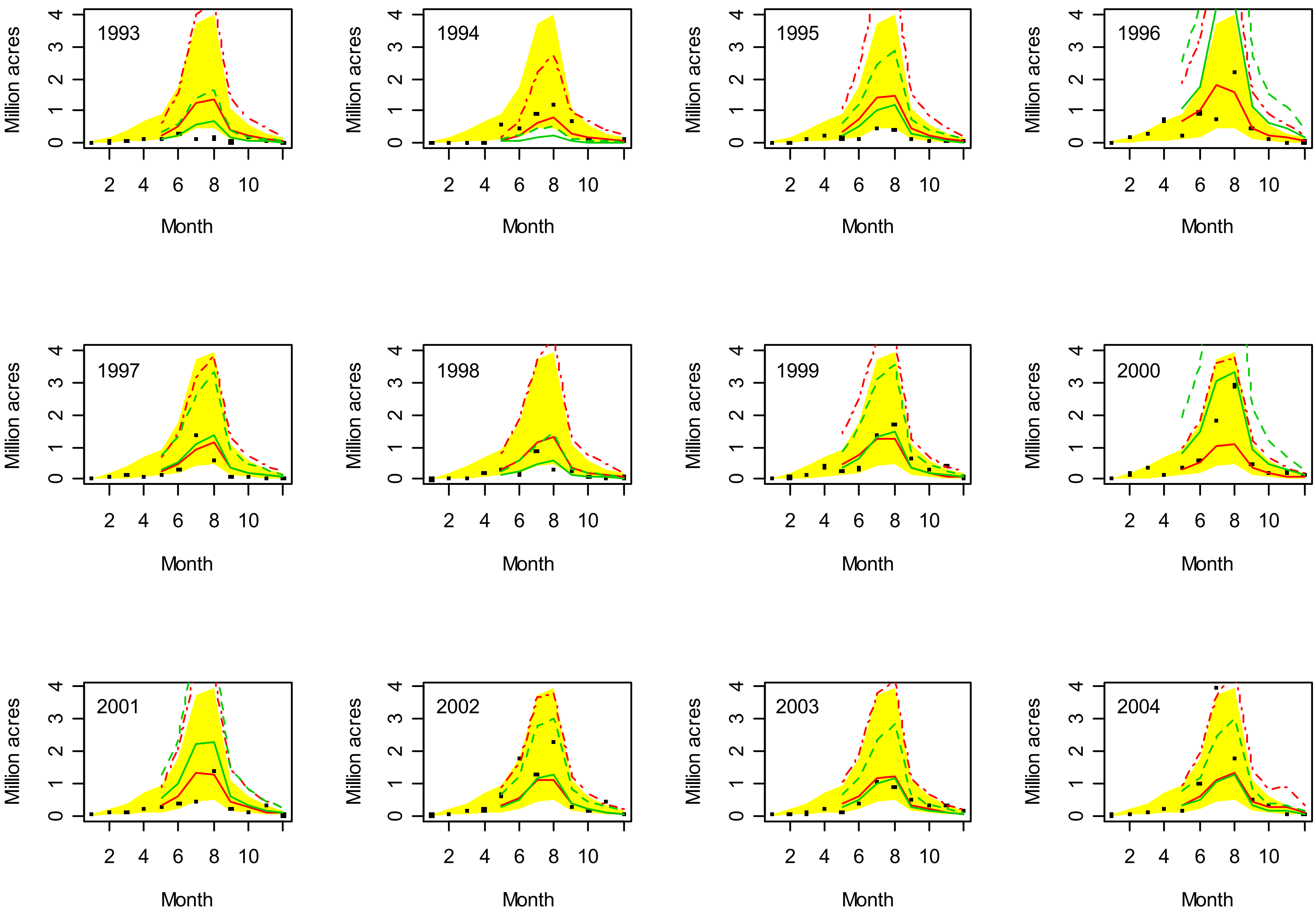
2) Median polish (approximate L1 solution)

iteratively remove row and column medians until row/column has median zero

$$y_{ij} = m + a_i + b_j + r_{ij}$$

3) Trimmed mean. In Splus use twoway(x,y, trim=.1)

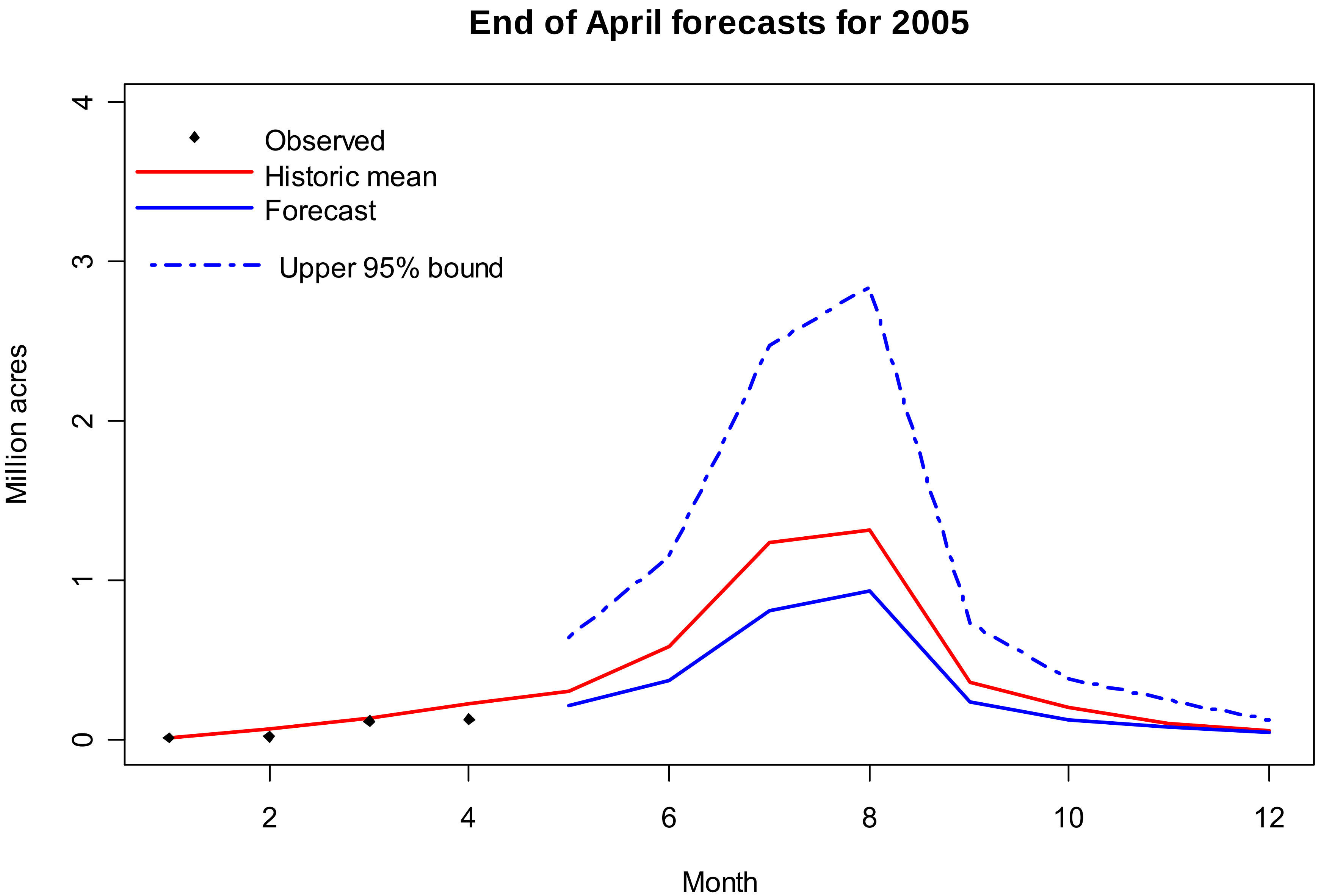

End of April forecasts for 2005
Observed
Historic mean
Forecast
Upper 95% bound
Million acres
0
1
2
3
4
2
4
6
8
10
12
Month



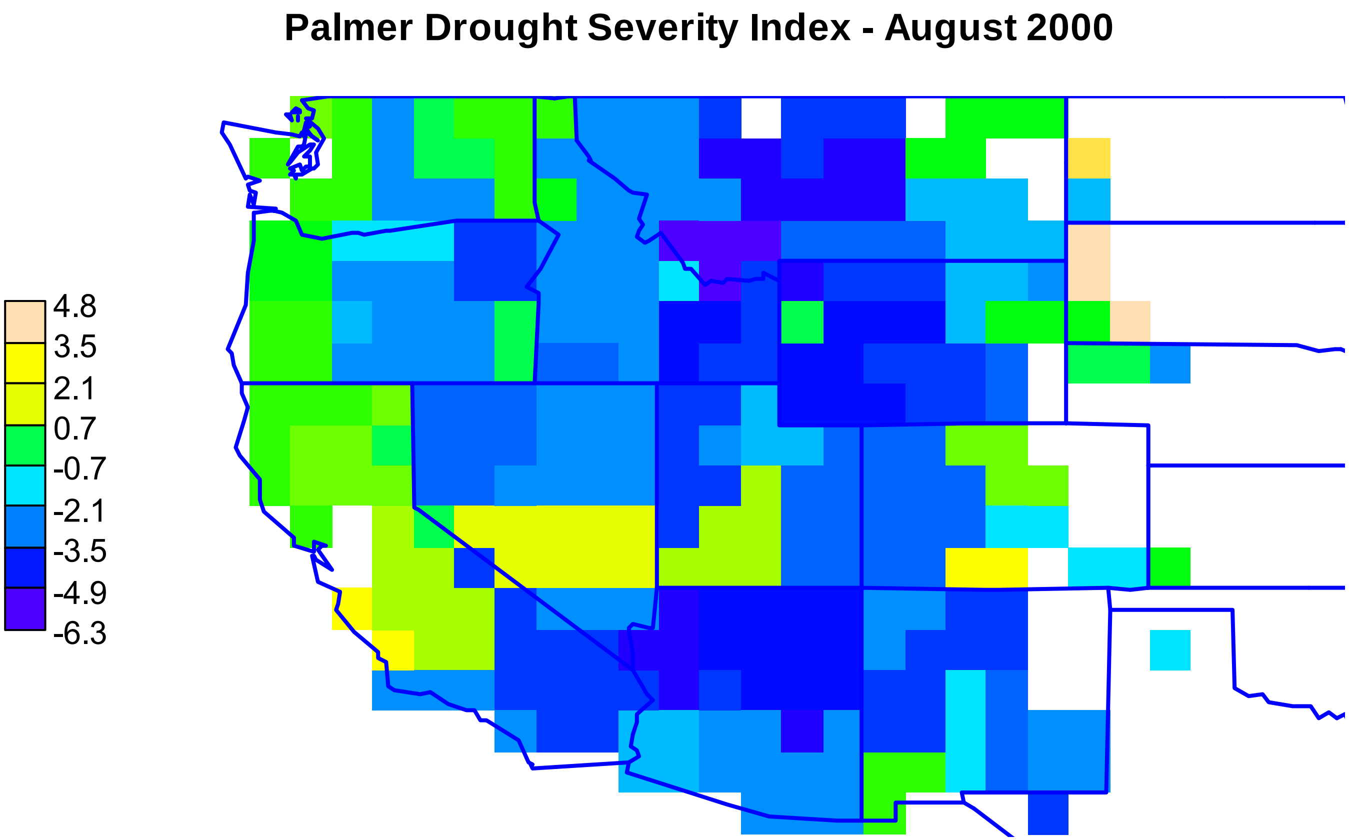
Black = Observed Yellow =historic 10-90% Green = Twoway polish Red = Arima

Table 1: Observed percentiles

	Percentile		
	80	90	95
Two-way April	79	84	89
Two-way June	81	90	94
Arima April	88	95	99
Arima June	94	97	100


End of April forecasts for 2005
Observed
Historic mean
Forecast
Upper 95% bound
Million acres
Month
0
1
2
3
4
2
4
6
8
10
12

2) 1x1 degree grids (~ 60-100 km²)



Anthony Westerling
Scripps Institution of Oceanography
La Jolla, CA

Data – monthly totals in one degree grid cells 1980 - 2000

Y_i = total number of acres burned in voxel i

N_i = number of fires in voxel i

Voxel = degree x degree x month

$\mathbf{X_i}$ = { PDSI, Temp, Precip, mPDSI, PDO, ENSO, Location, Month)

Fire Danger rating based on

$\pi_i =$ Pr[> 1000 acres burn in voxel i]

From the joint distribution of Y and N we have

$\pi_i = \Pr[\, N_i > 0 \mid \mathbf{X_i} \,] \Pr[\, Y_i > 1000 \mid N_i > 0, \mathbf{Xi} \,]$

mPDSI= maximum PDSI in last 12 months
PDO= Pacific Decadal Oscillation
ENSO = El Niño-Southern Oscillation

Odds Ratio

$$\theta_i = \frac{\pi_i}{1-\pi_i} \div \frac{\pi_h}{1-\pi_h} \qquad h = \text{historic}$$

$$\text{Rating}\begin{cases} \text{below normal} & \hat{\theta}_i + 2se < 1 \\ \text{normal} & otherwise \\ \text{above normal} & \hat{\theta}_i - 2se > 1 \end{cases}$$

$$\text{Extreme} \qquad \hat{\theta}_i - 2se > 5$$

Estimation – Local regression techniques

$$\text{logit } p_i = \beta_o + \nu_{yr} : \delta_{yr} : g_1(lon_i, lat_i) + g_2(month_i) + \sum_k g_{k-2}(X_{ik})$$

g = nonparametric smoothing function (spline, loess, tps, …)

X_{ik} = value of k^{th} covariate in voxel i

ν_{yr} = PDO value for yr^{th} year

δ_{yr} = ENSO value for yr^{th} year

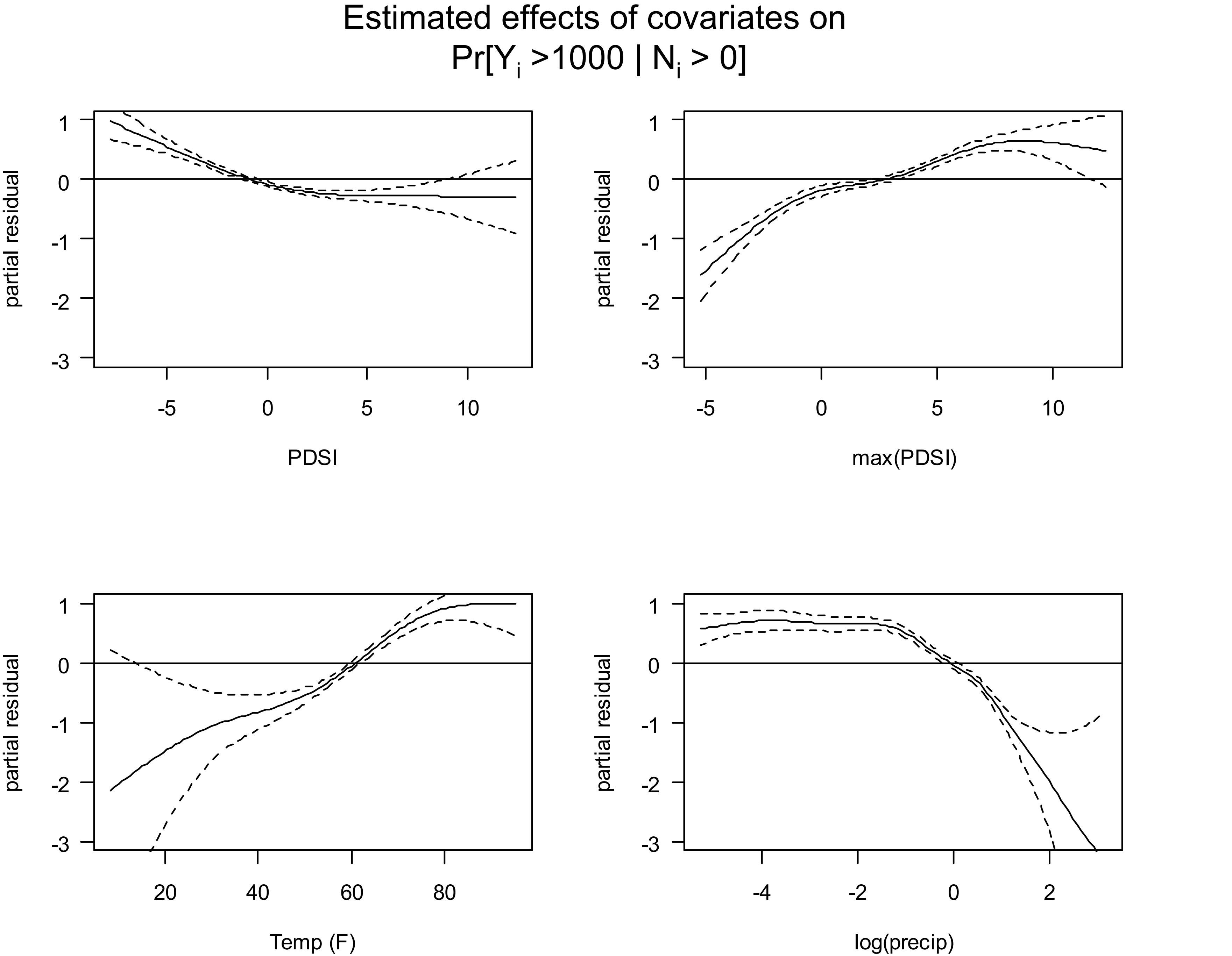
Piecewise Polynomials and Splines

1) Replace each vector ***X*** with a matrix whose columns are transformations of ***X***.
For example,

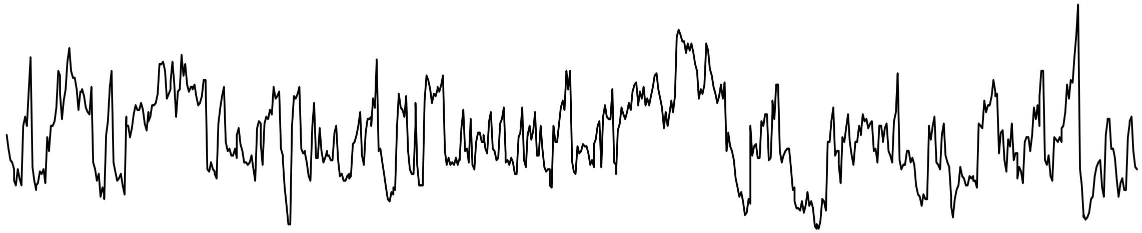
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \rightarrow \begin{pmatrix} x_1 & x_1^2 \\ \vdots & \vdots \\ x_n & x_n^2 \end{pmatrix} \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \xrightarrow{bs()} \begin{pmatrix} x_1 & (x_1 - k_1)_+ & (x_1 - k_2)_+ \\ \vdots & \vdots & \vdots \\ x_n & (x_n - k_1)_+ & (x_n - k_2)_+ \end{pmatrix}$$

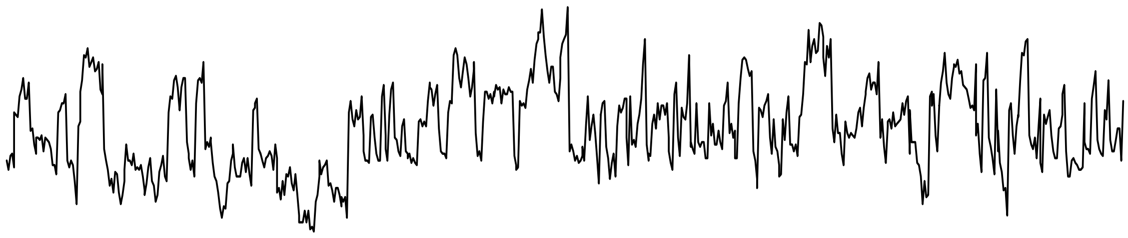
2) Use linear regression with the transformed variables

$$\text{logit } p_i = \beta_o + \beta_1 \text{bs}(X_{1i}) + \beta_2 \text{bs}(X_{2i}) + X_{3i} : \text{bs}(X_{4i})$$


Estimated effects of covariates on
Pr[Y_i >1000 | N_i > 0]
partial residual
1
0
-1
-2
-3
-5
0
5
10
PDSI
max(PDSI)
20
40
60
80
Temp (F)
-4
-2
0
2
log(precip)

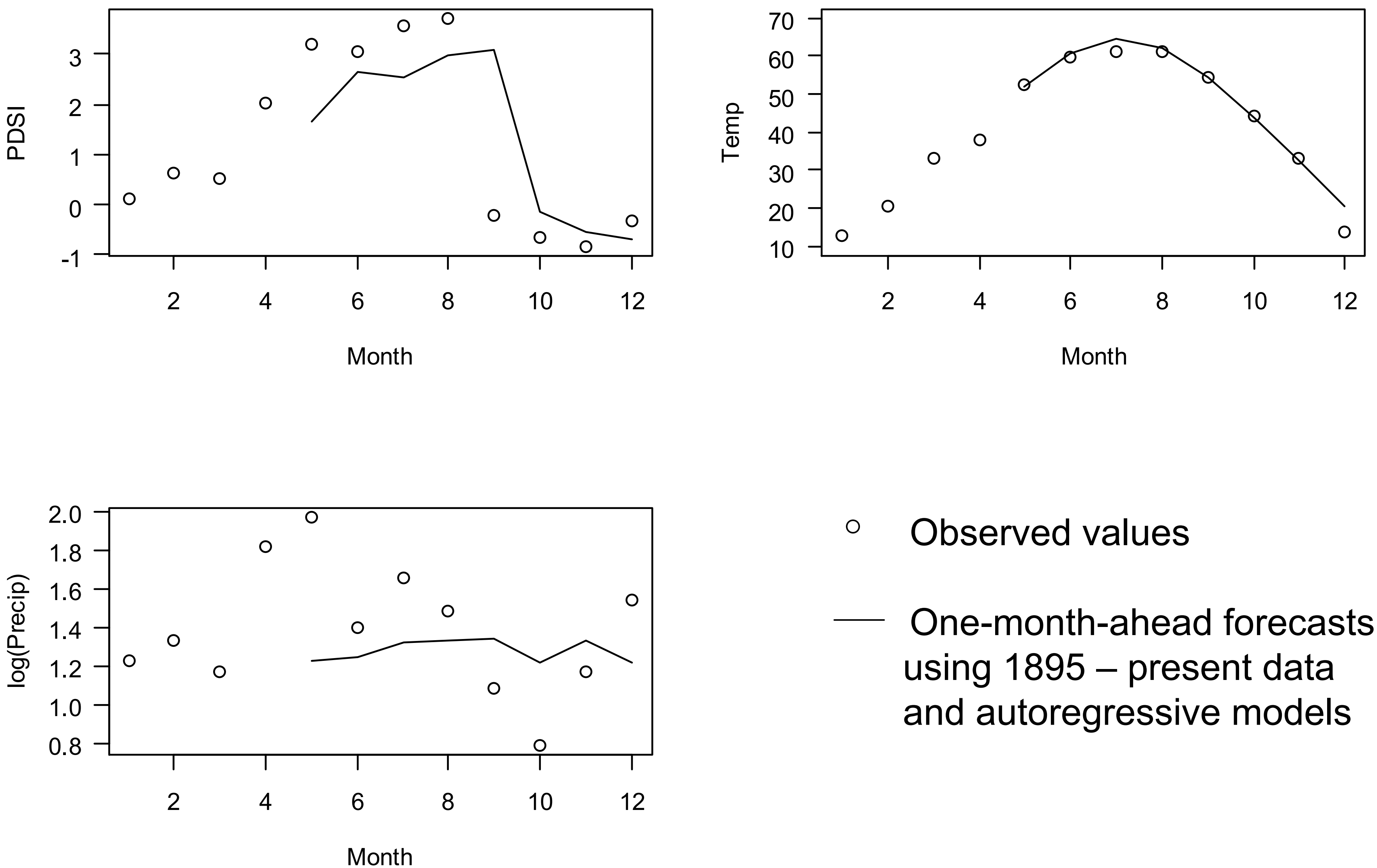
PDSI values 1895-2005





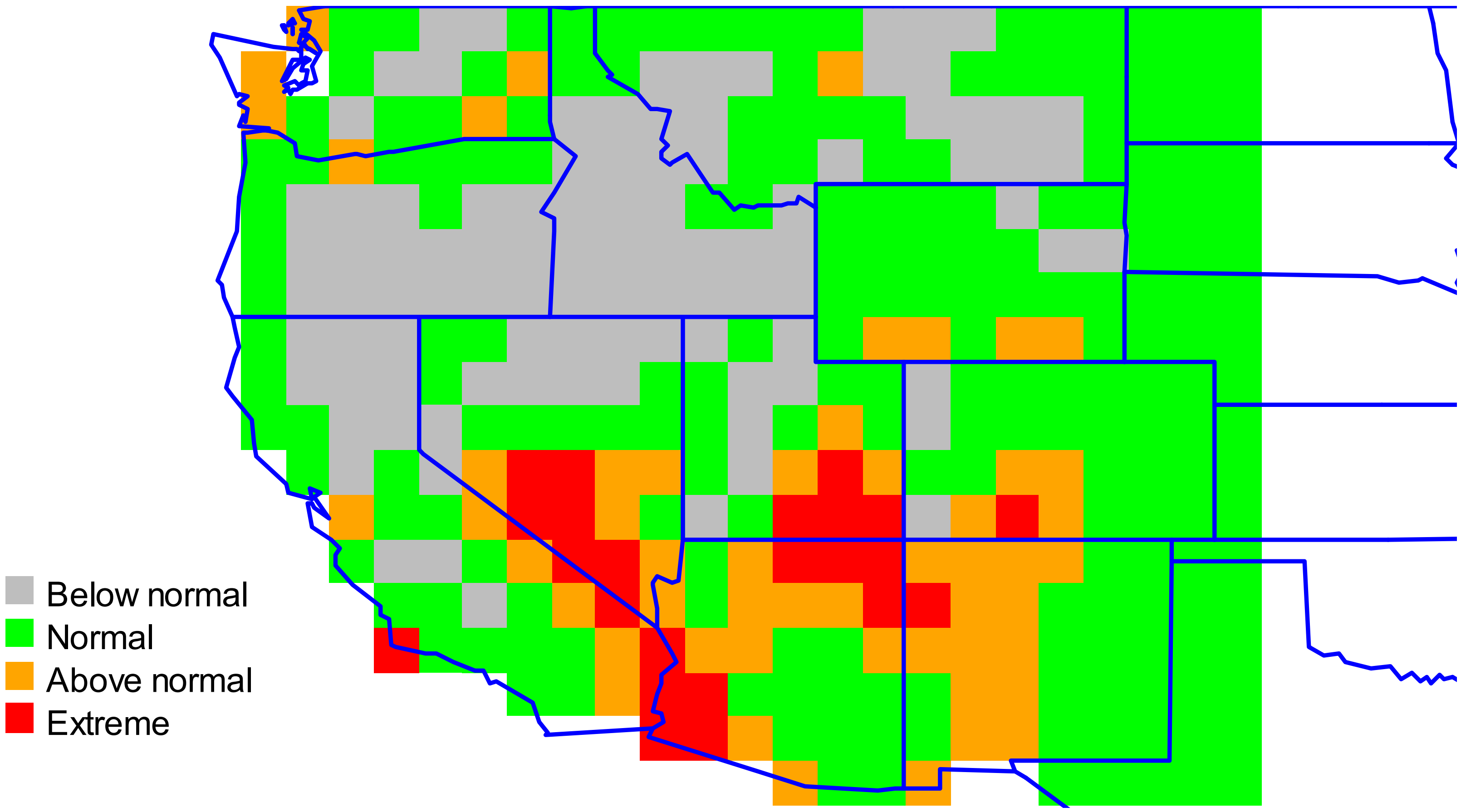
http://www1.ncdc.noaa.gov/pub/data/cirs

Monthly averages for each climate division

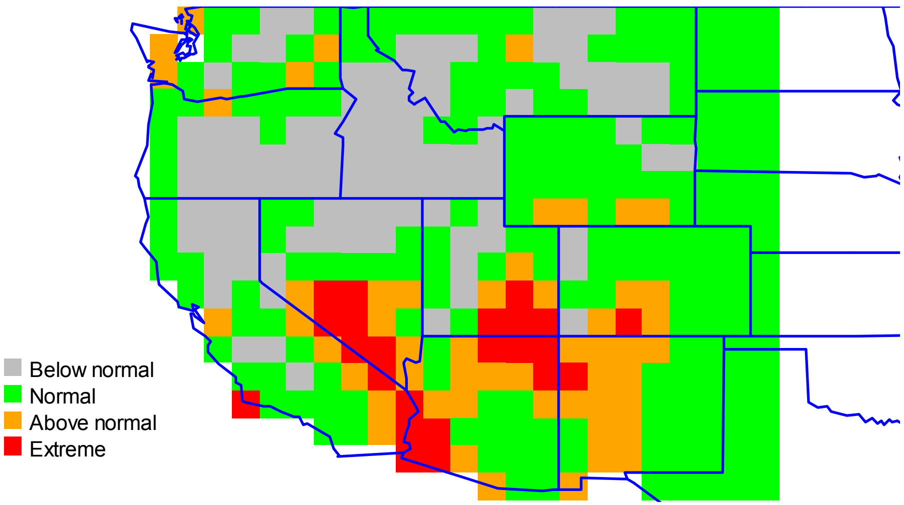


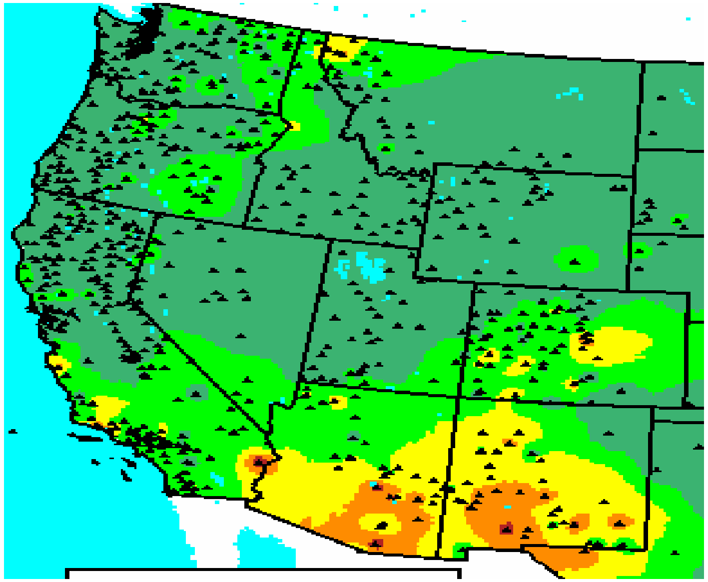
PDSI
3
2
1
0
-1
2
4
6
8
10
12
Month
Temp
70
60
50
40
30
20
10
2
4
6
8
10
12
Month
log(Precip)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
2
4
6
8
10
12
Month
Observed values
One-month-ahead forecasts
using 1895 – present data
and autoregressive models

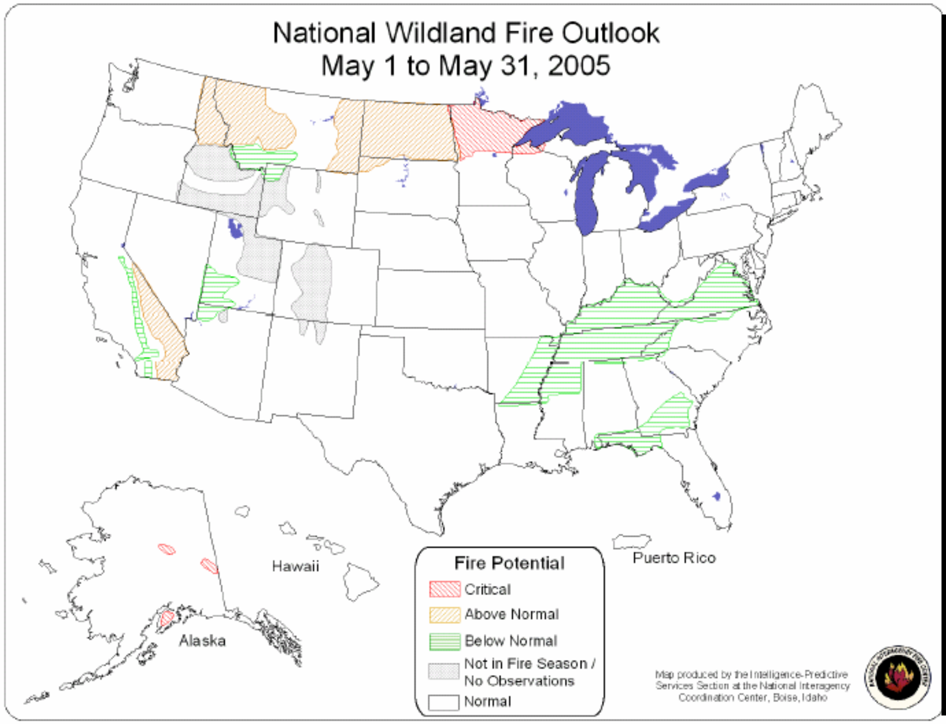
Forecast for May 2005

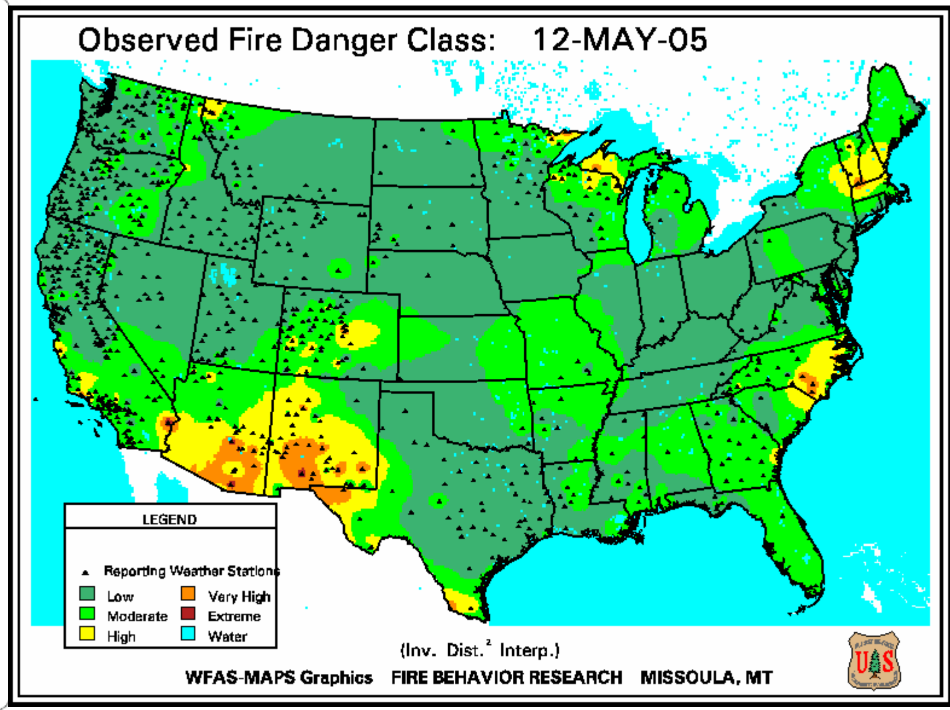


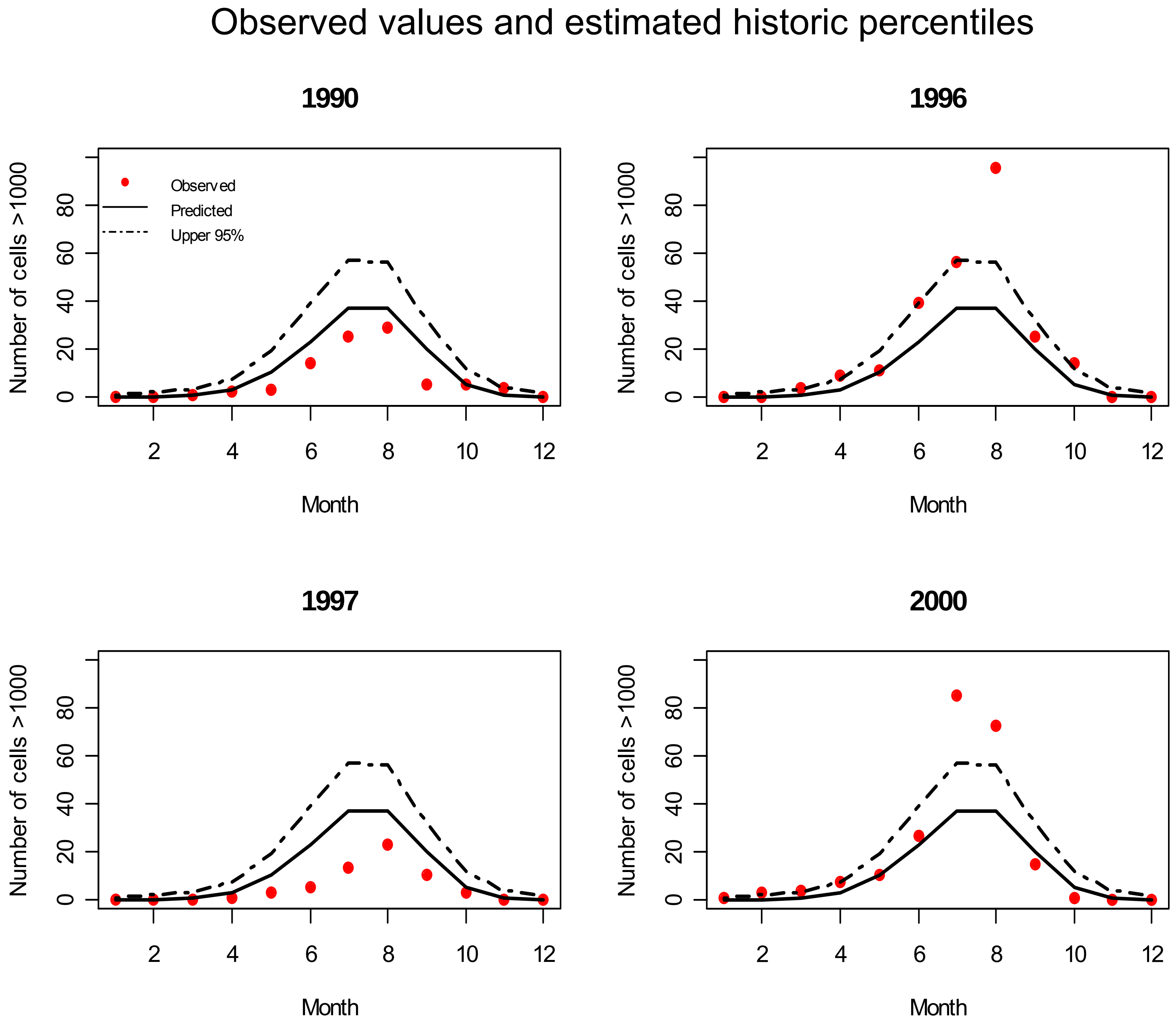
Forecast for May 2005

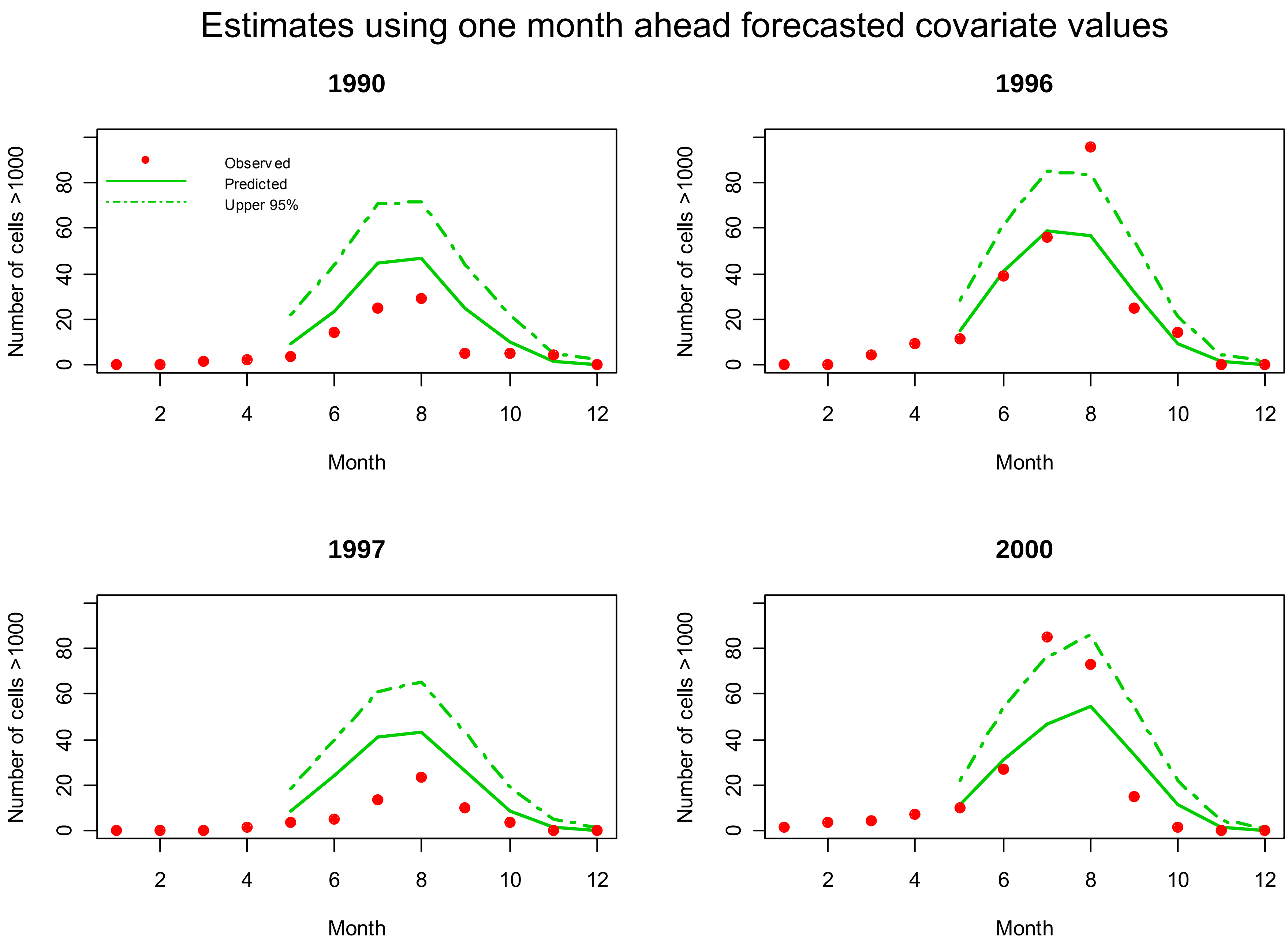




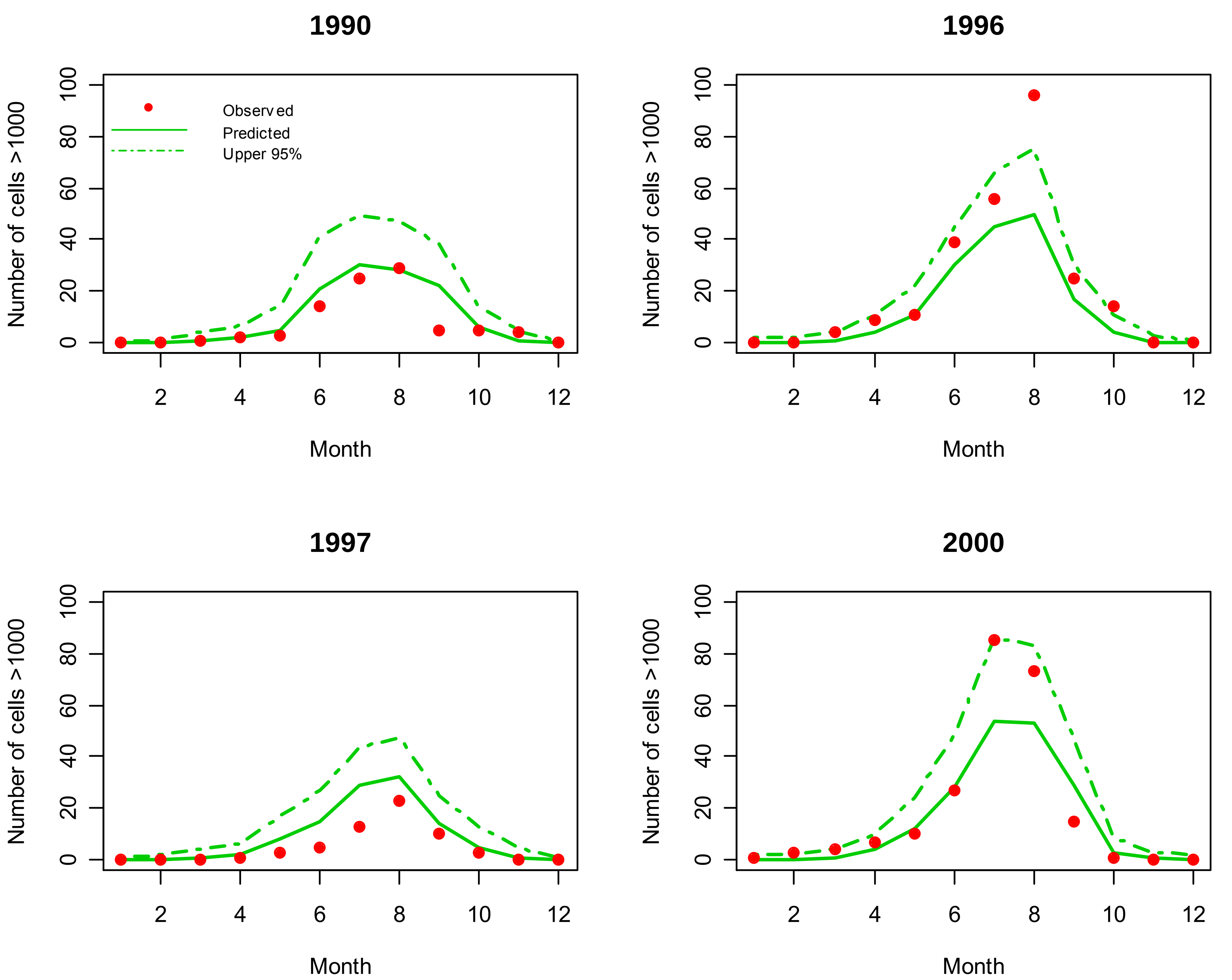




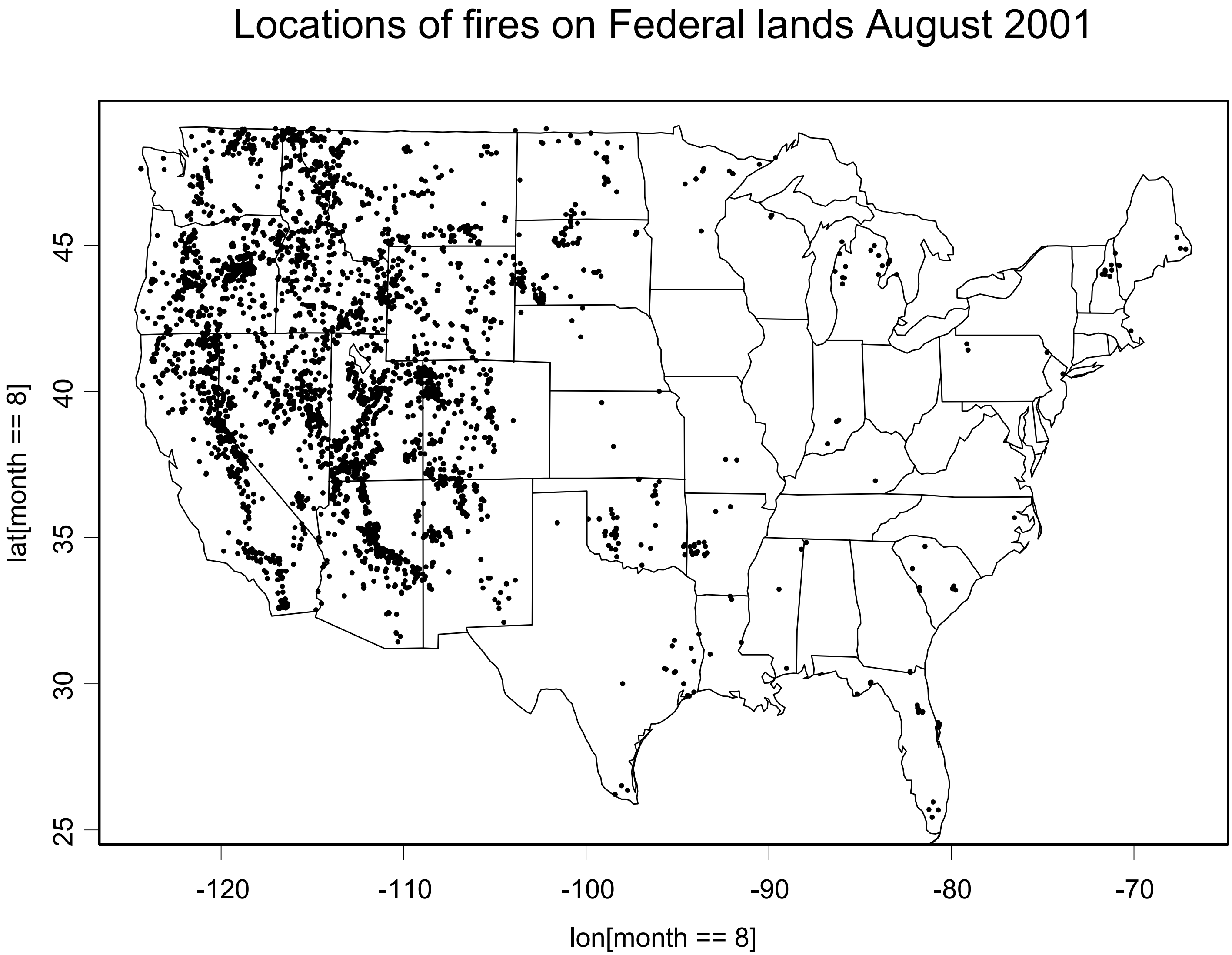

Observed values and estimated historic percentiles
1990
1996
1997
2000
Observed
Predicted
Upper 95%
Number of cells >1000
Month
0
20
40
60
80
2
4
6
8
10
12


Estimates using one month ahead forecasted covariate values
1990
1996
1997
2000
Observed
Predicted
Upper 95%
Number of cells >1000
Month

Estimates using observed covariate values



3) **km x km x day level**



Model : spatial-temporal point process

Assume a conditional intensity function

$$\lambda(x, y, t \mid \theta) = \Pr\text{ob}\{dN(x, y, t) = 1 \mid H_t\} / dxdydt$$

N = number of fires in $(x,x+dx] \times (y,y+dy] \times (t,t+dt]$
H_t = history up to time t

Estimation

Let

$N_{x,y,t} = 1$ if there is a fire in (x,y,t) - voxel

$= 0$ otherwise.

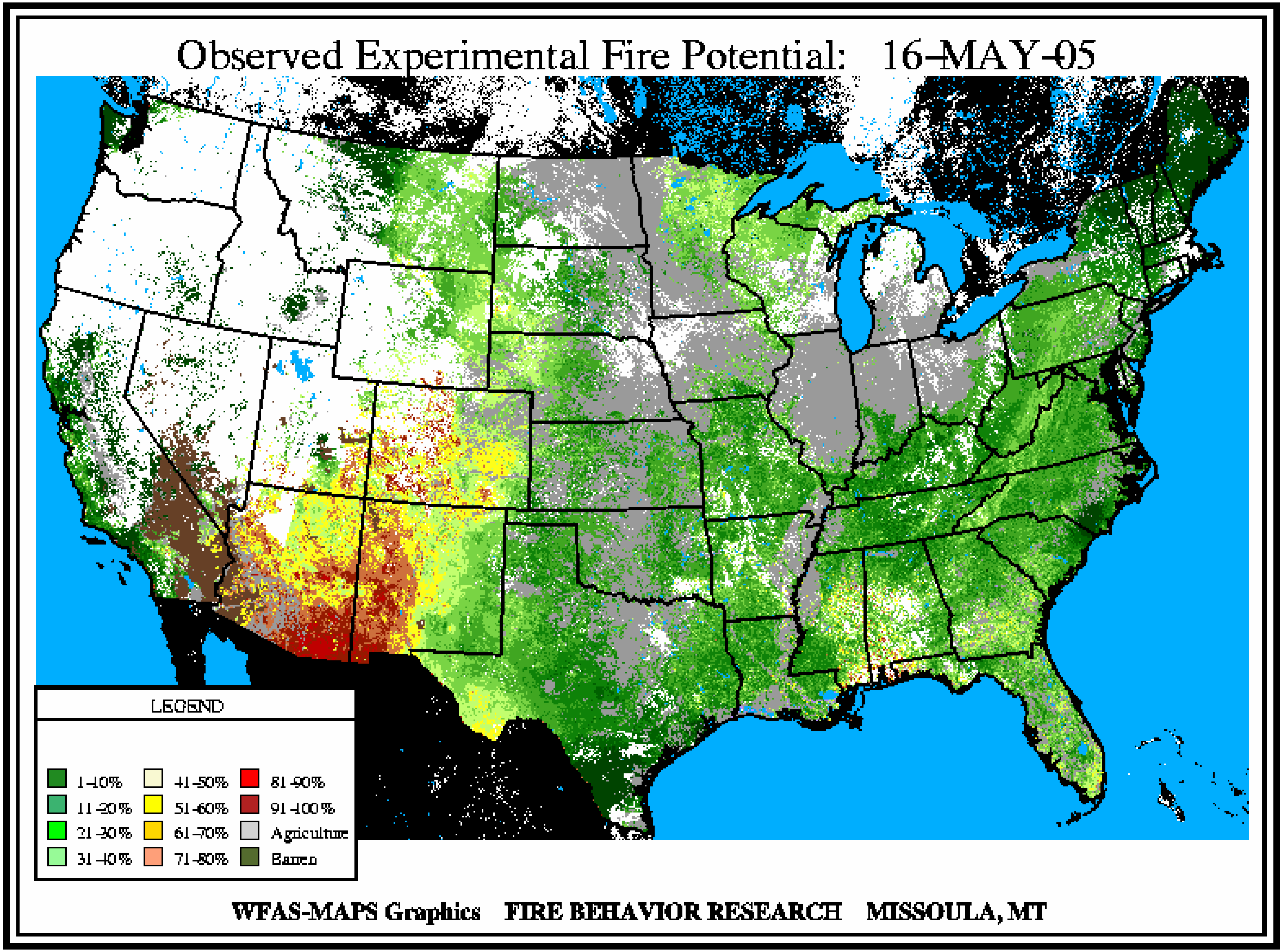
voxel size = *km* x *km* x *day*

Bernoulli log-likelihood

$$\sum_{x,y,t} N_{x,y,t} \log(\lambda_{x,y,t}) + \sum_{x,y,t} (1 - N_{x,y,t}) \log(1 - \lambda_{x,y,t})$$

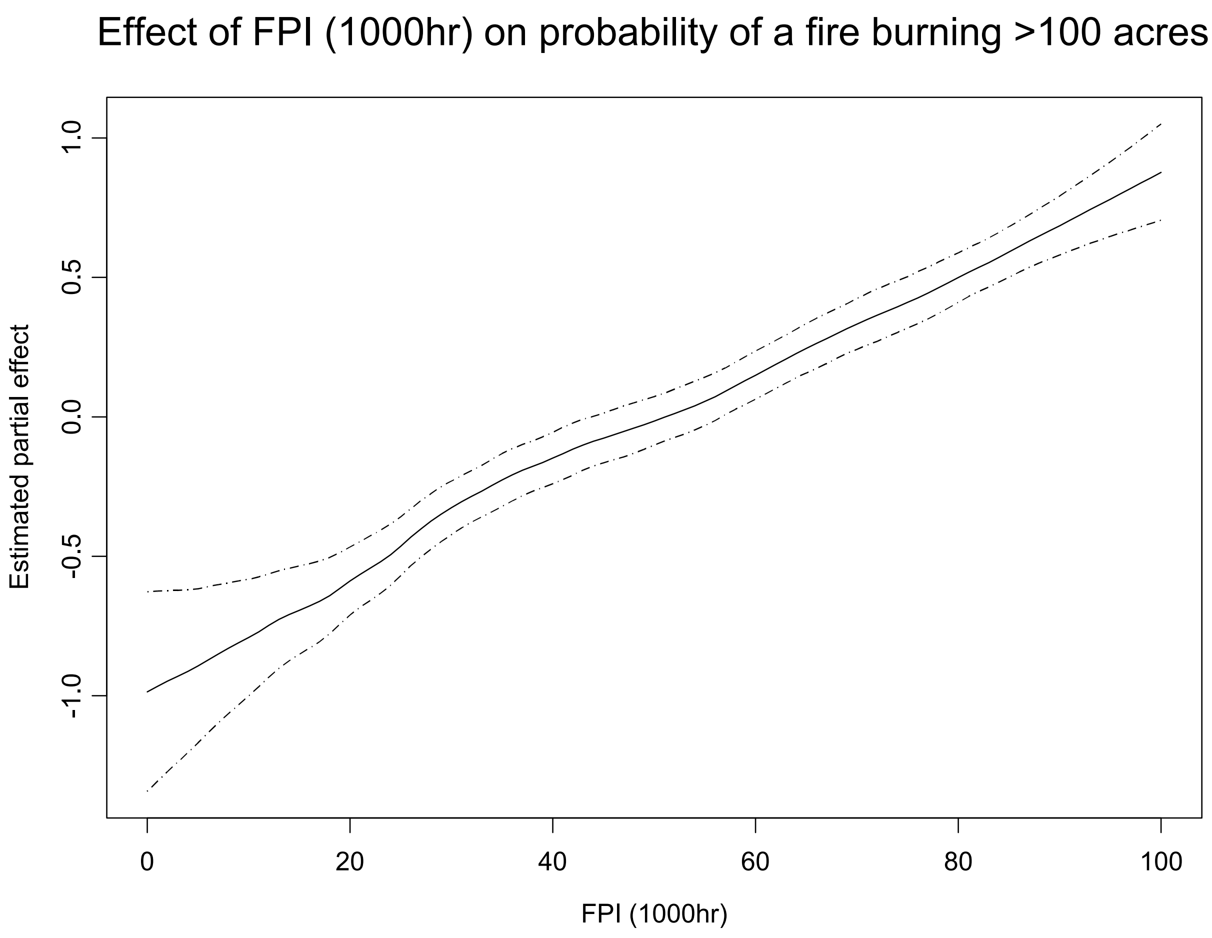
Data for US Forest lands one year:

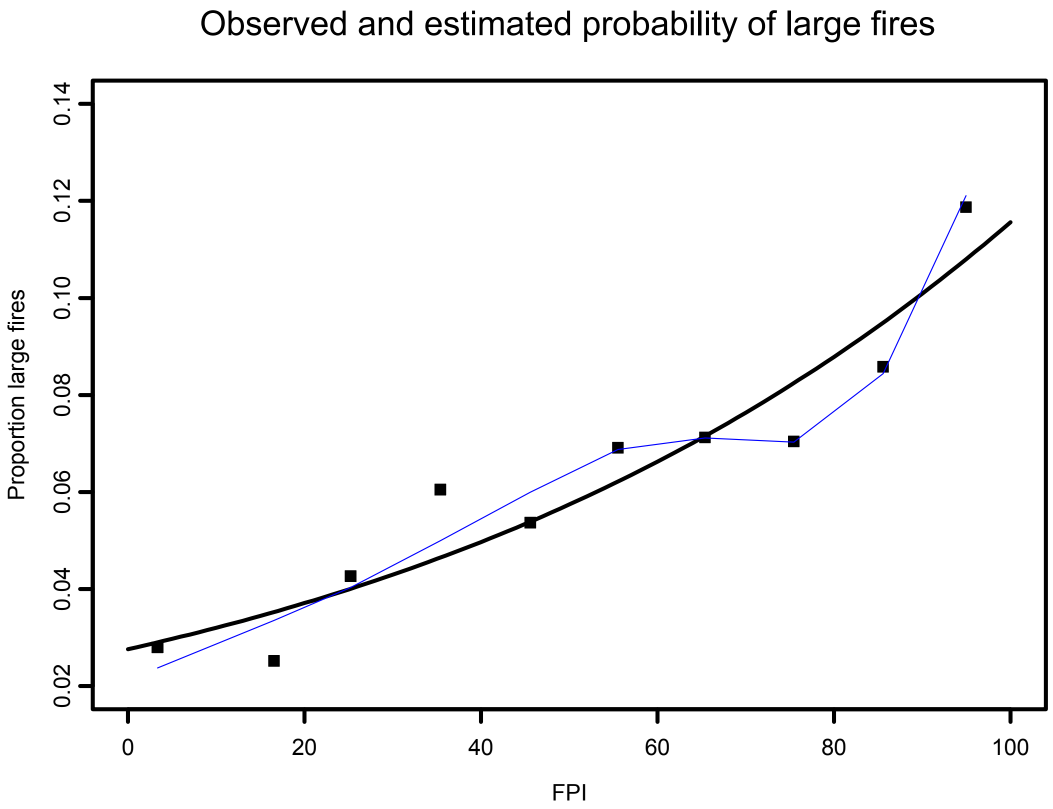
sum over 730×10^6 voxels



Bob Burgan (Missoula, MT)
Jacie Klaver (EROS)

Pr[large fire (> 100 acres) | ignition] = π(lon,lat,day,FPI)





The odds of a fire becoming large is estimated to increases by 2 fold (1.84-2.25)
when FPI increase by 40 points.

■ Observed proportion of large fires

Fitted probability (model with FPI alone)

Fitted probability (model with FPI, spatial location and day)

4) At the fire level

Fire Behavior Models

Characterize propagation and spread of fires under various environmental conditions.

Two common approaches used to model fire spread:

1. Discrete approach
2. Continuous approach

Fire Behavior Models

1) *Discrete approach*

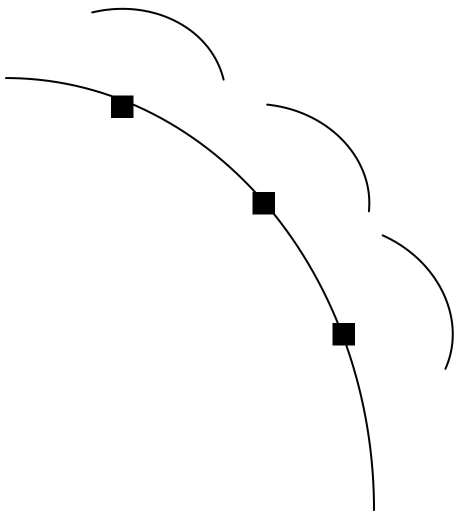
Forest is divided into a set of cells arranged on a regular grid with each cell having a probability of burning that depends on conditions in the cell and in surrounding cells.

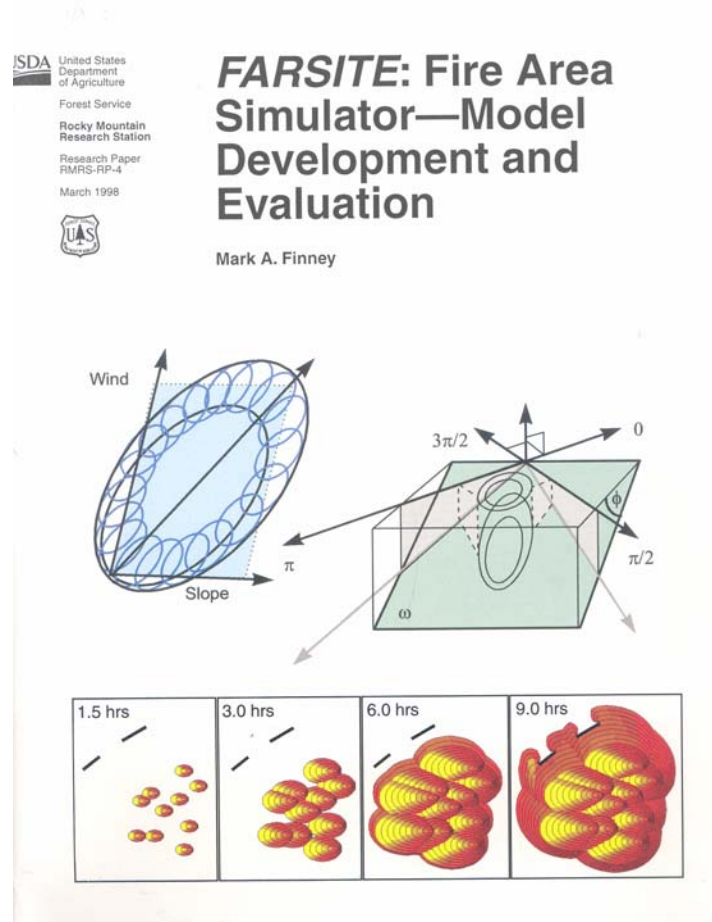
Fire Behavior Models (cont.)

2. *Continuous approach*

Based on Huygen's principle concerning propagating light waves:

All points along a wave front act as if they were point sources for a new wave




USDA United States Department of Agriculture

Forest Service

Rocky Mountain Research Station

Research Paper RMRS-RP-4

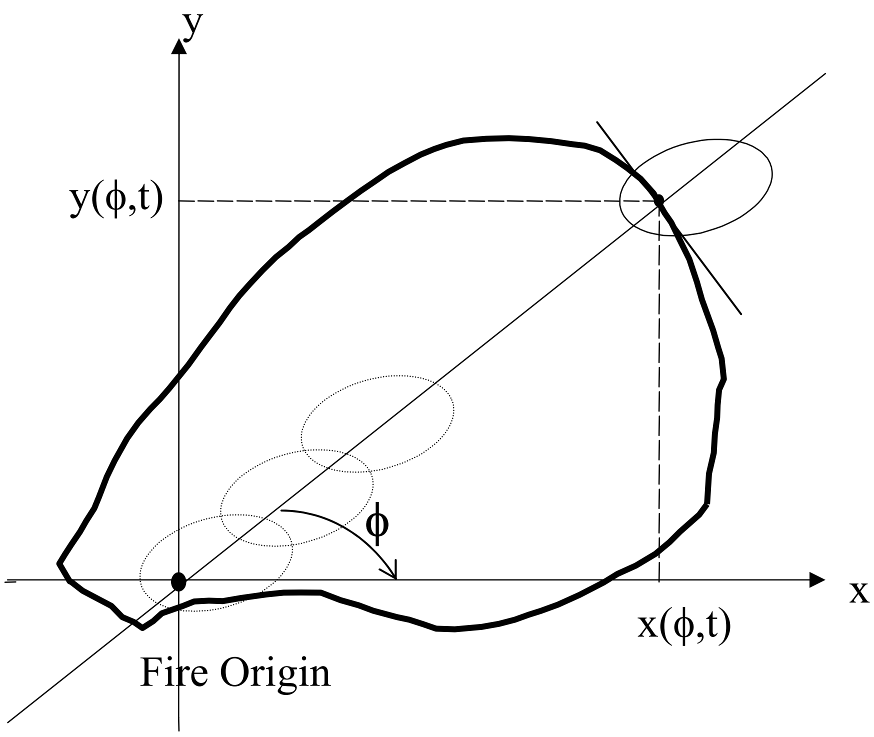
March 1998

FARSITE: Fire Area Simulator—Model Development and Evaluation

Mark A. Finney

Stochastic partial differential equations

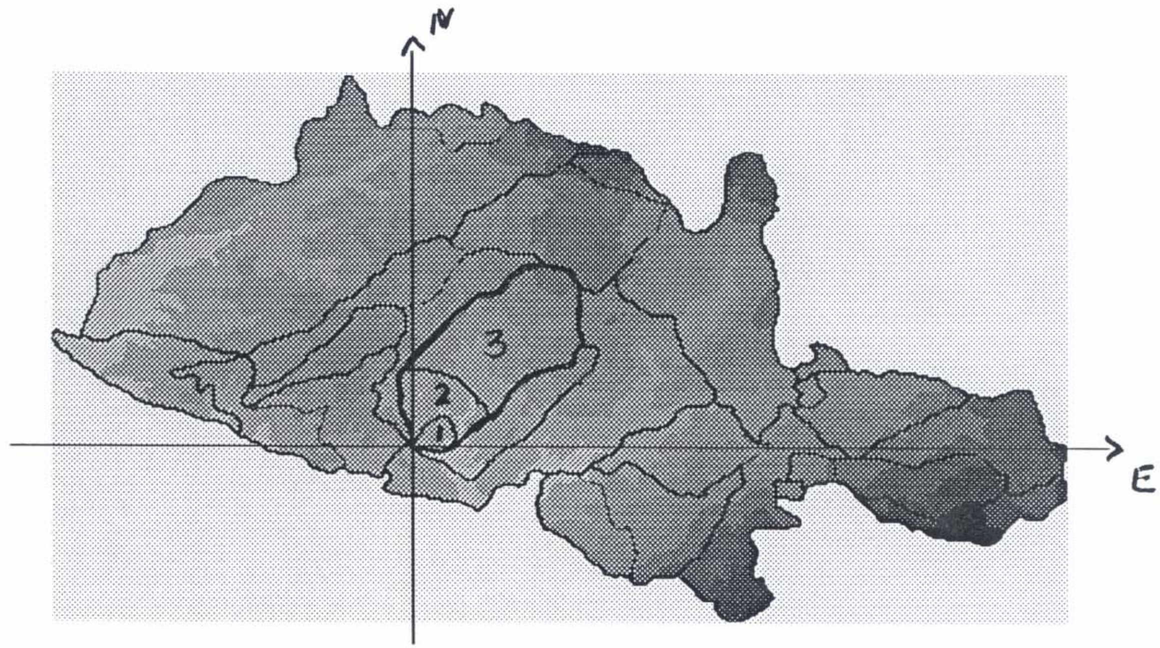
$$\partial x_t(s) = R_t(s)\cos(\phi_t(s))\partial t + K_t(s)\sin(\phi_t(s))\partial t + \sigma_1 dB_1(t)$$
$$\partial y_t(s) = R_t(s)\sin(\phi_t(s))\partial t + K_t(s)\cos(\phi_t(s))\partial t + \sigma_2 dB_1(t)$$



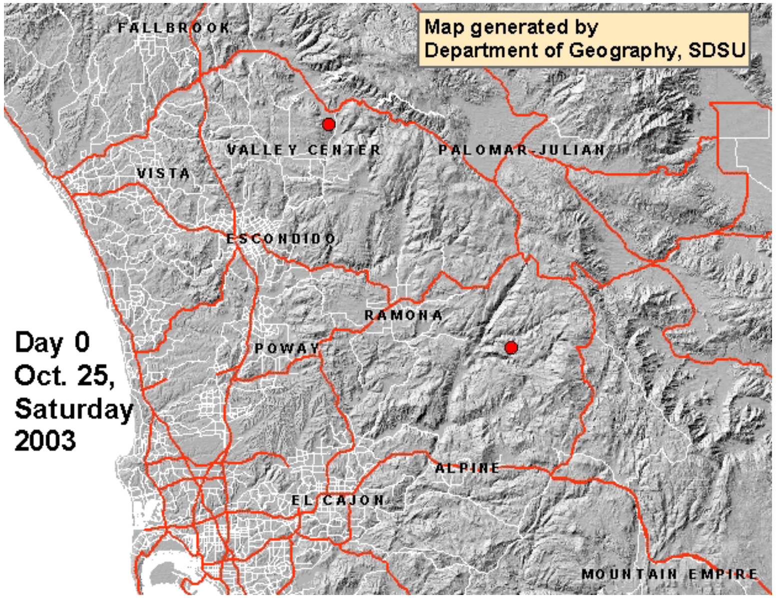
$R_t(s)$ = rate of spread in direction normal to perimeter

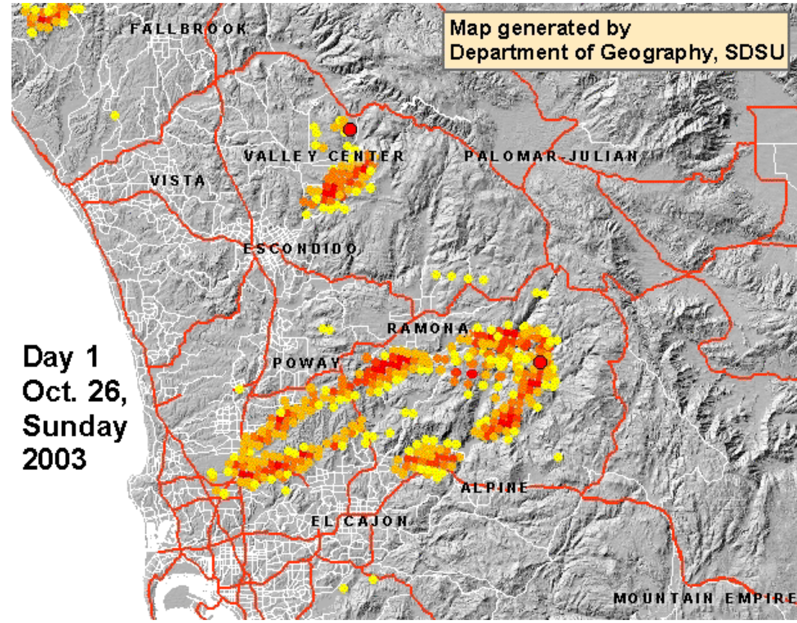
$K_t(s)$ = rate of spread in tangential direction

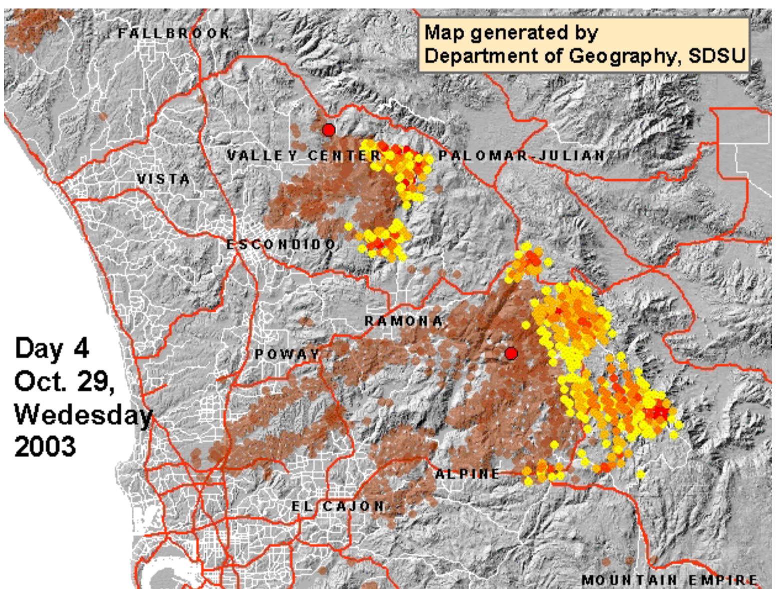
B(t) = Brownian process



Fire perimeter map for Bee Fire, San Bernardino National Forest, California, USA. June 30 to July 30, 1996.







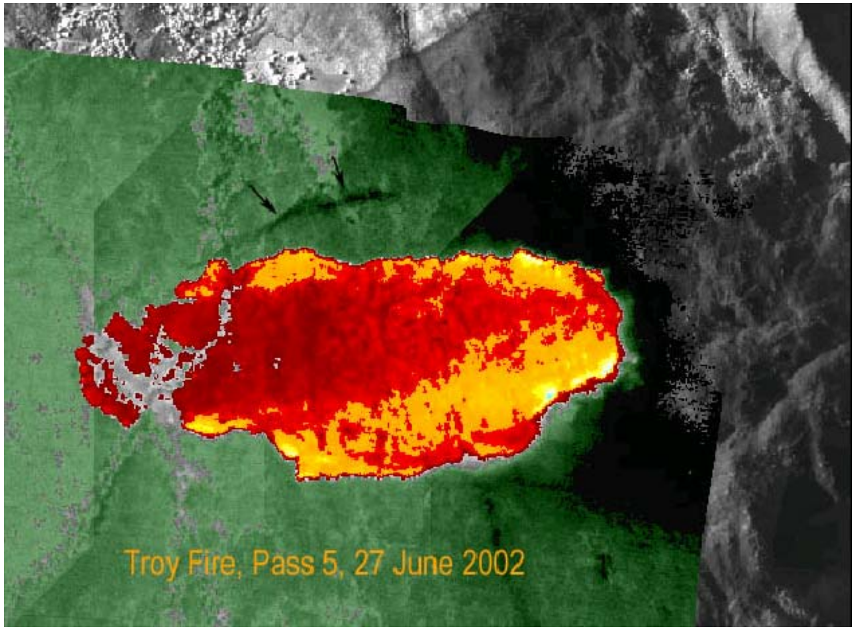
The maps are based on MODIS data two passes per day.
Maps created by Ming Tsou

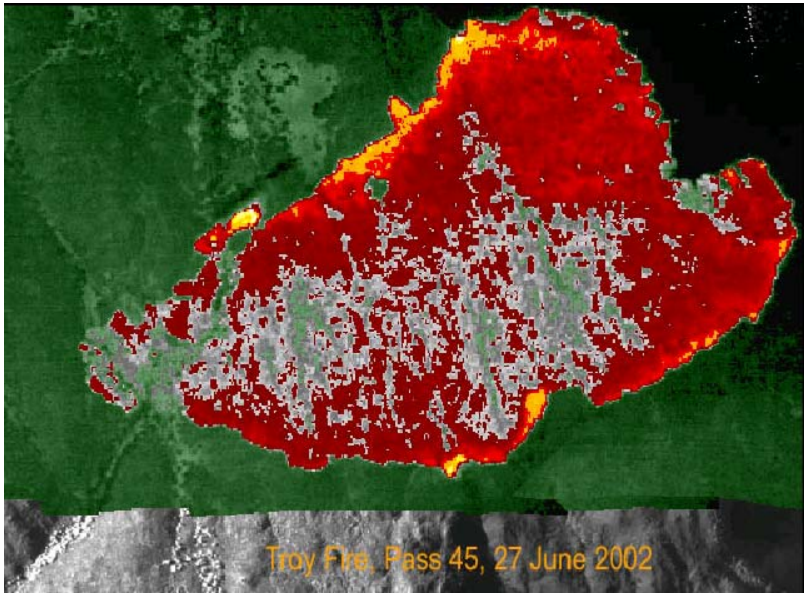
Piper Navajo remote sensing aircraft

Fire Science: Wildfire Monitoring and Prediction

FireMapper measures the radiance of emitted thermal-infrared light, which readily penetrates smoke

Phillip Rigan
PSW Forest Service
Riverside Fire Laboratory


Troy Fire, Pass 5, 27 June 2002


Troy Fire, Pass 45, 27 June 2002

Summary

1. Probability models useful for forecasting and studying goodness-of-fit (usefulness of Fire Danger Indices)
2. R statistical package (analysis and mapping)
3. Local regression techniques (loess, splines, basis function)
4. Space-time point processes – basic to all fire models (marked point process – acres or cost)
5. Data management and acquisition (automate updating data from the web)
6. State-space and interacting particle system models for fire spread