Spatial Modeling of *P. ramorum* in California:

Strengths and weaknesses of rule-based, statistical, and cellular automata methods

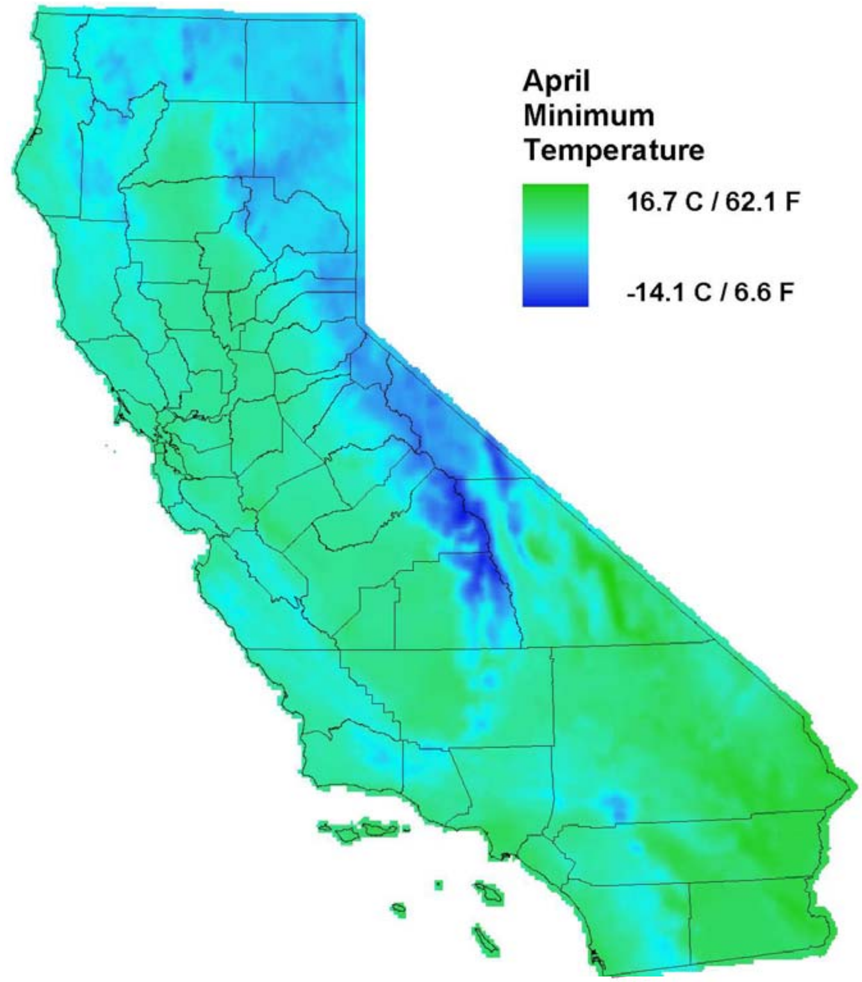
Ross K. Meentemeyer
University of North Carolina at Charlotte
Department of Geography & Earth Sciences

I. Rule-Based Modeling

- incorporates epidemiological and biological factors driving establishment and spread in California plant communities
    - Host susceptibility
    - Host epidemiology
    - Moisture
    - Temperature

Parameterizing the Model

- Scored 5 variables to encode magnitude & direction of their effect on disease spread

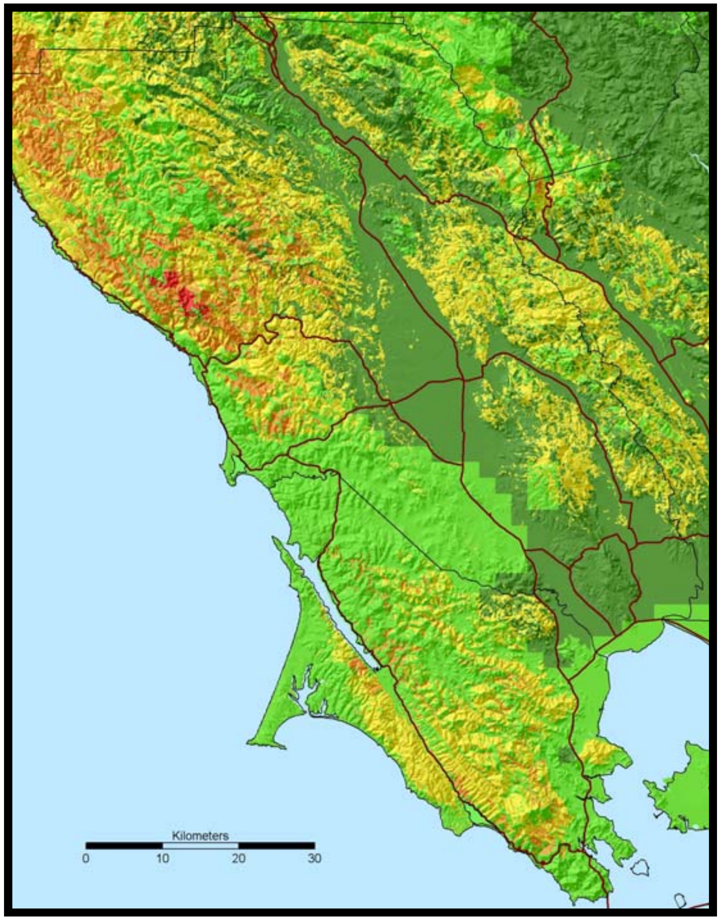


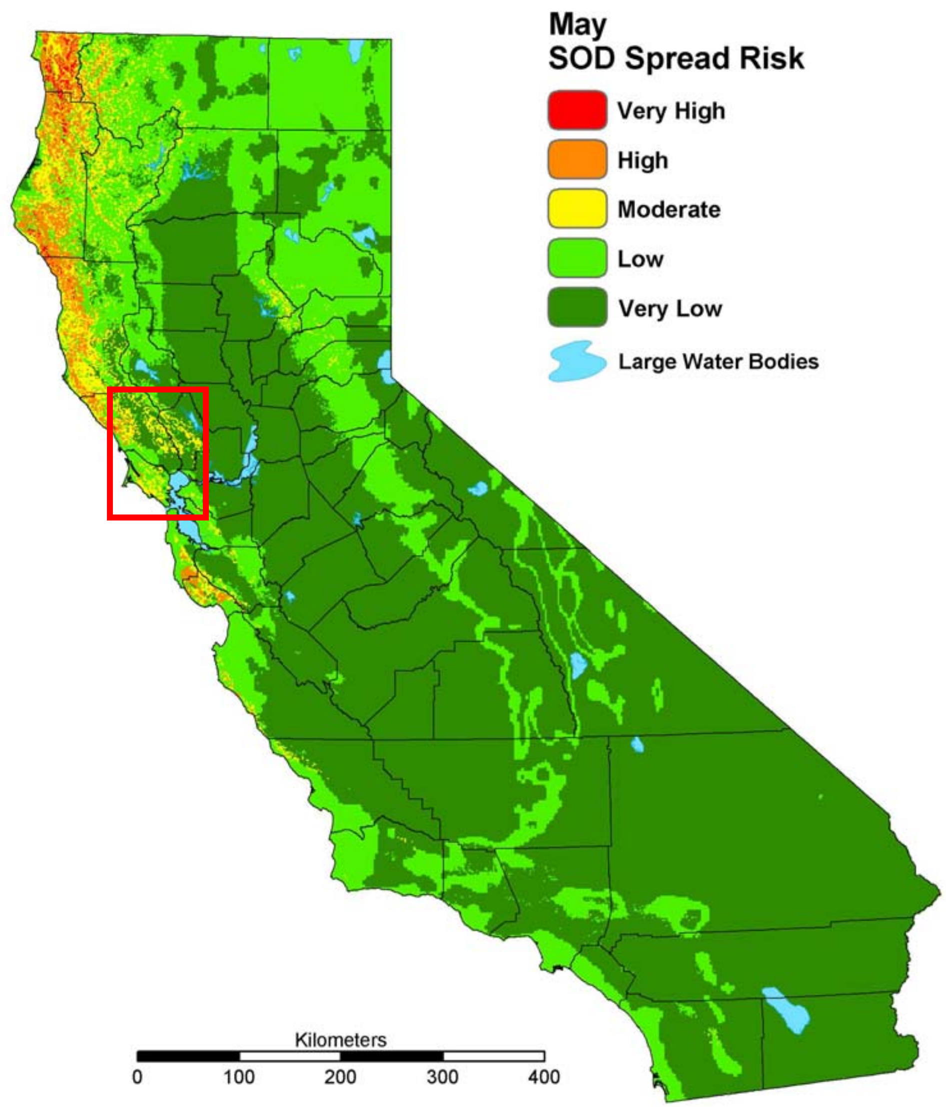
Variables

- Host species abundance
- Precipitation
- Relative Humidity
- Maximum Temperature
- Minimum Temperature

Model Results

Seasonal Changes: December - May

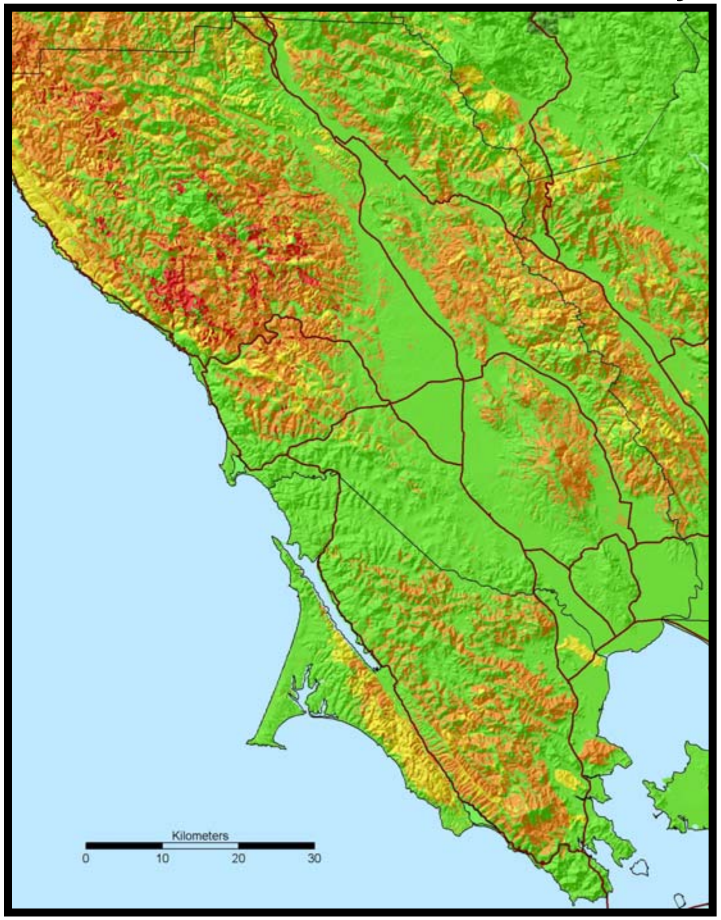


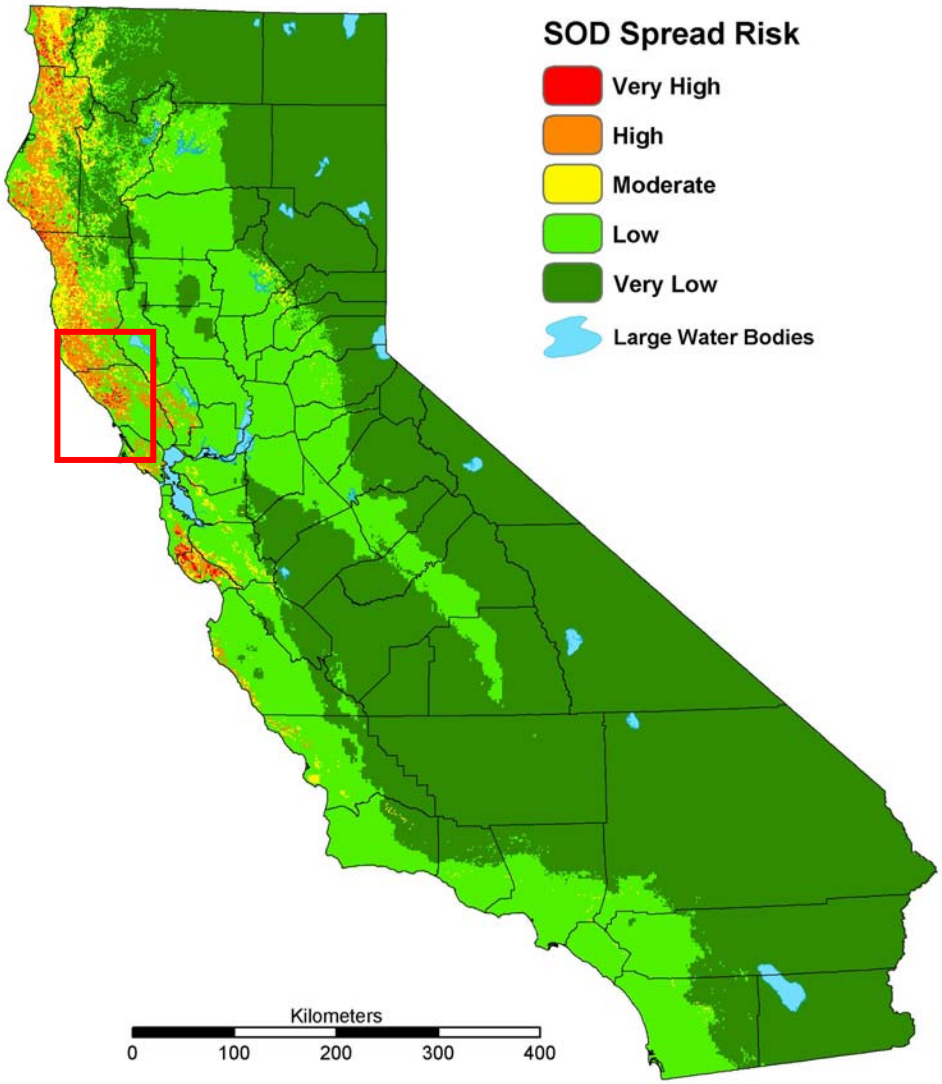


Meentemeyer et al. 2004 (*Forest Ecology & Management)*

Model Results

Cumulative Risk: December - May





Meentemeyer et al. 2004 (*Forest Ecology & Management*)

Rule-Based Modeling

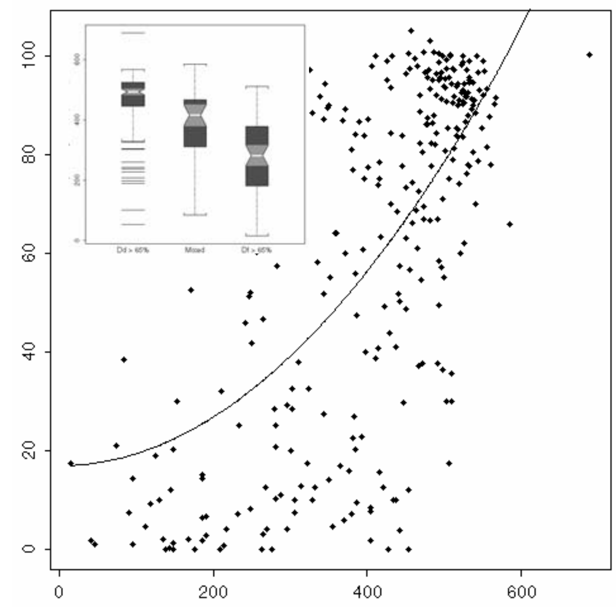
Strengths

- simple methods
- works well with limited data on pathogen survival and transmission

Limitations

- too simple
- may not reflect real field conditions
- risk-based (not probabilistic) and static

II. Statistical Modeling

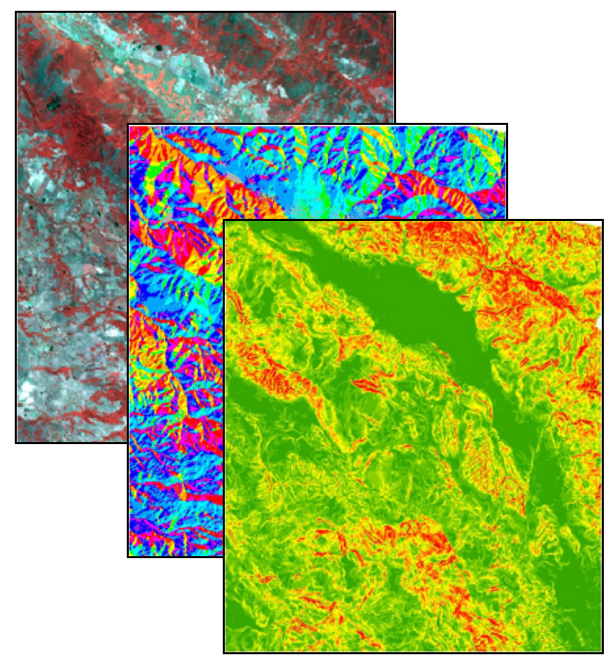


- based on observed relationships between pathogen occurrence, host composition, and environmental conditions

Example Methods

- Regression
- CART
- Neural Networks

(Regression, CART: require presence & absence data)



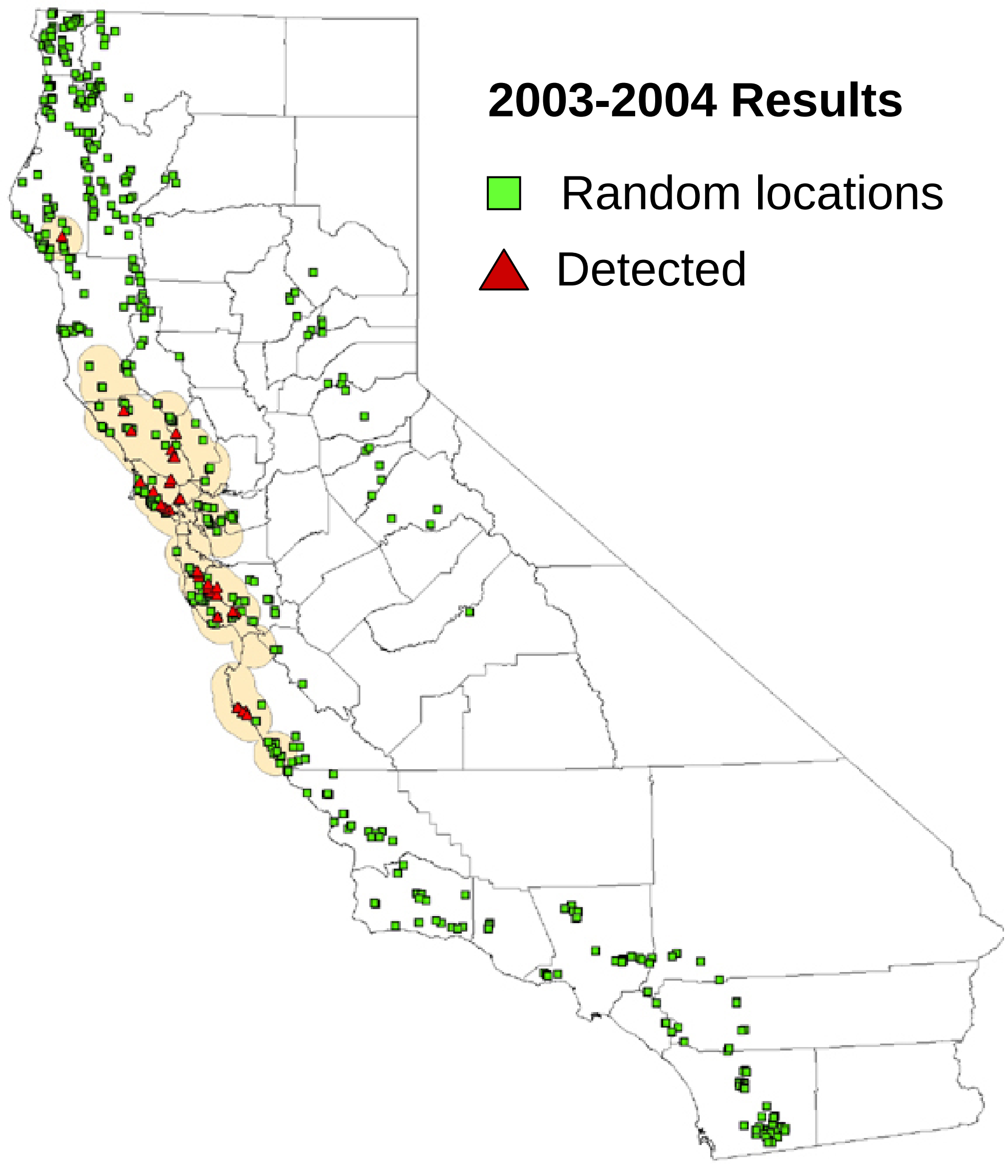
- once a model is developed, can apply its equation across mapped variables in a GIS

Response Variable

- extensive dataset on distribution *P. ramorum*
- within high-risk forests, we have surveyed 495 random locations (2003-2004)
- *P. ramorum* detected at 33 of the 495 sites.

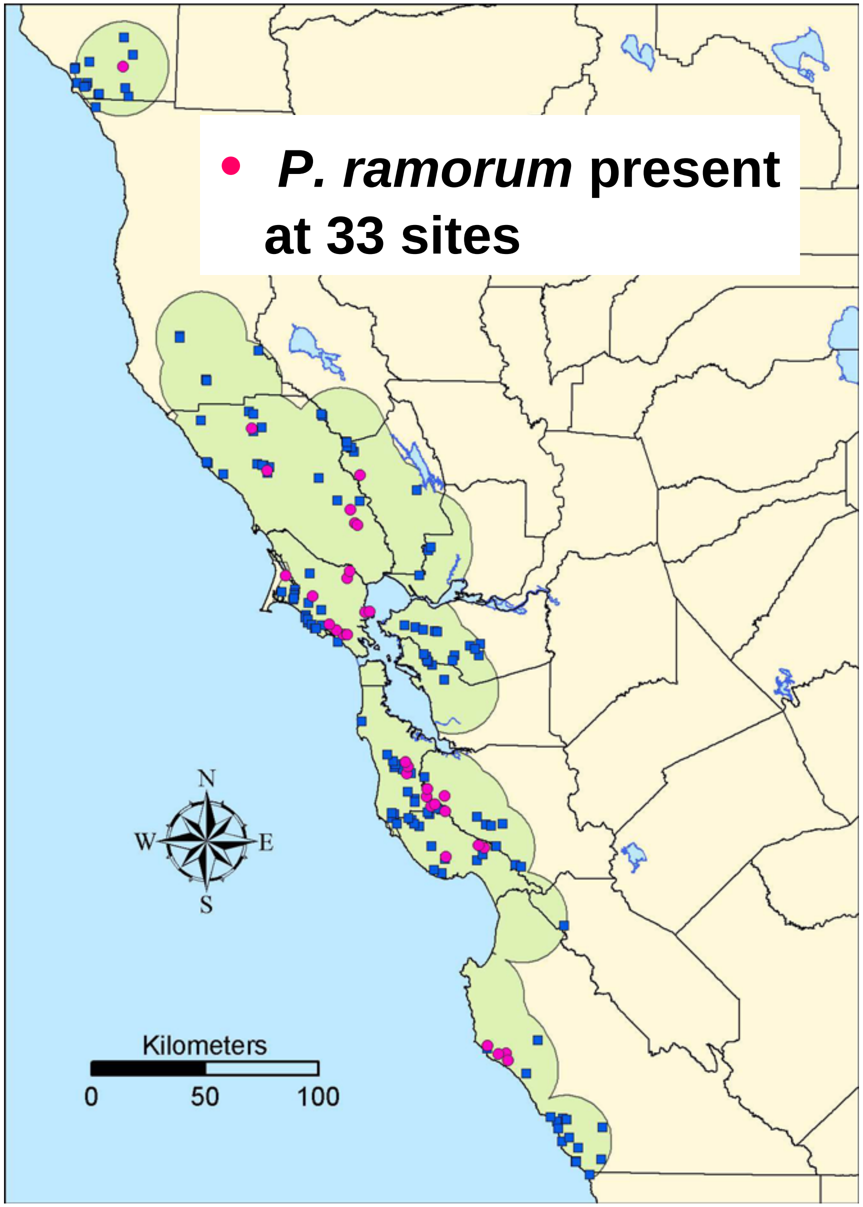
  n = 9 of 139 in 2003
  n = 24 of 347 in 2004

- 20 km was farthest detection



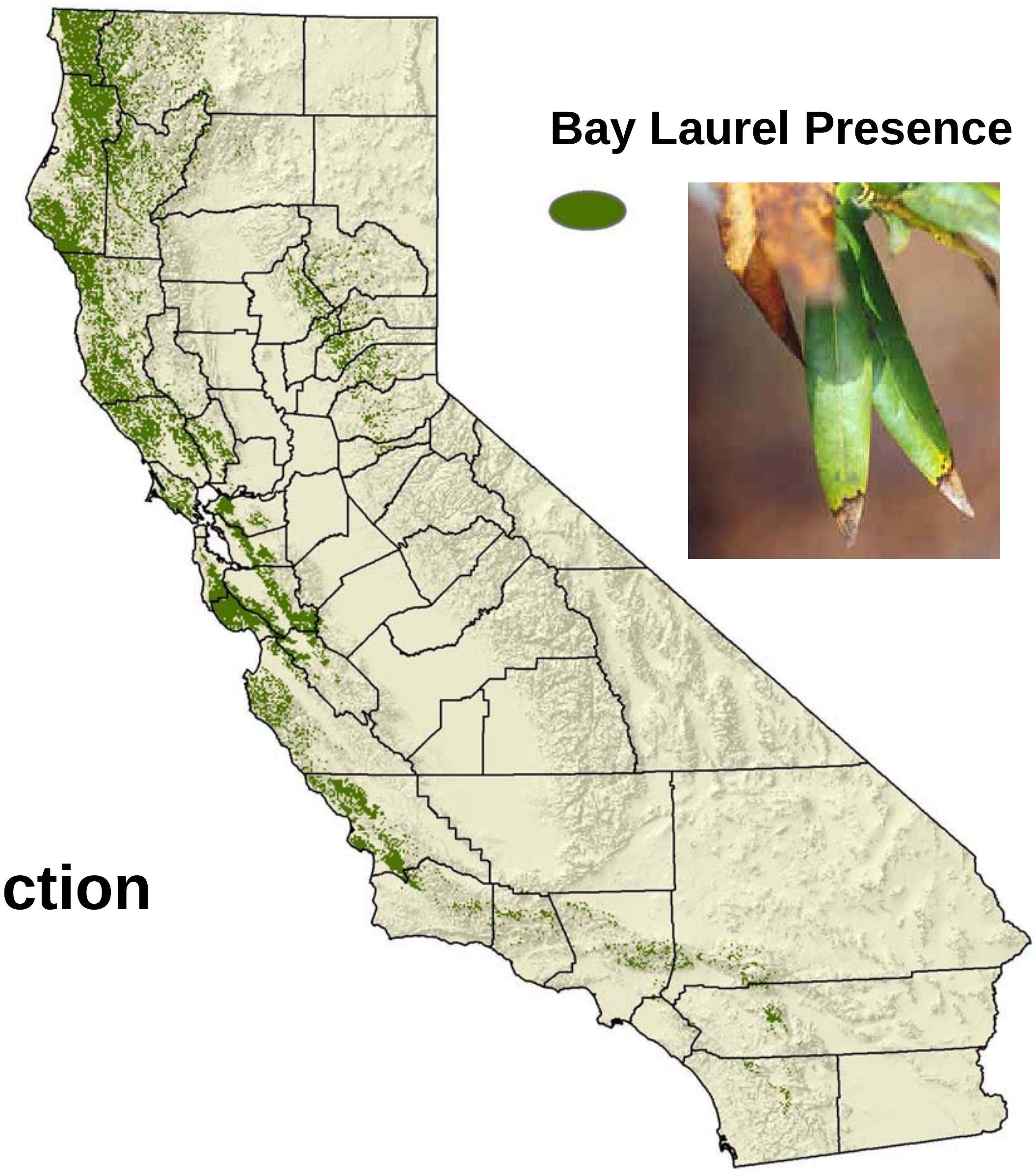
Defining the Analysis Region

- mapped COMTF sites
- limited to areas < 20 km of confirmed infection sites
- yielded **166** random sites from survey

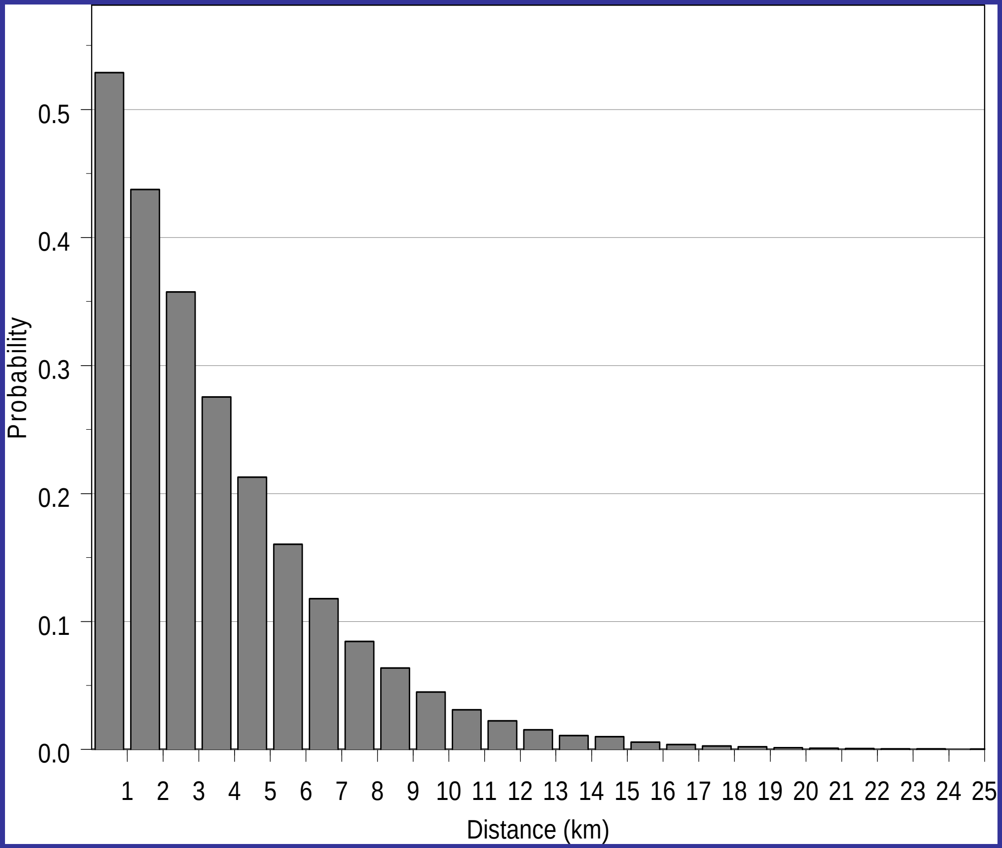


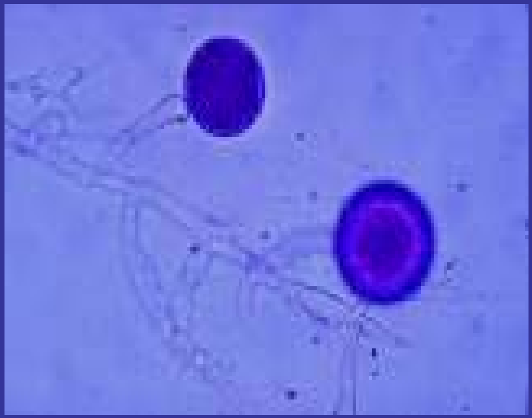
Predictor Variables

- Population pressure
- Climate conditions
- Fire history
- Presence of bay laurel & tanoak
- Distance to known infection



Statistically-Based Model





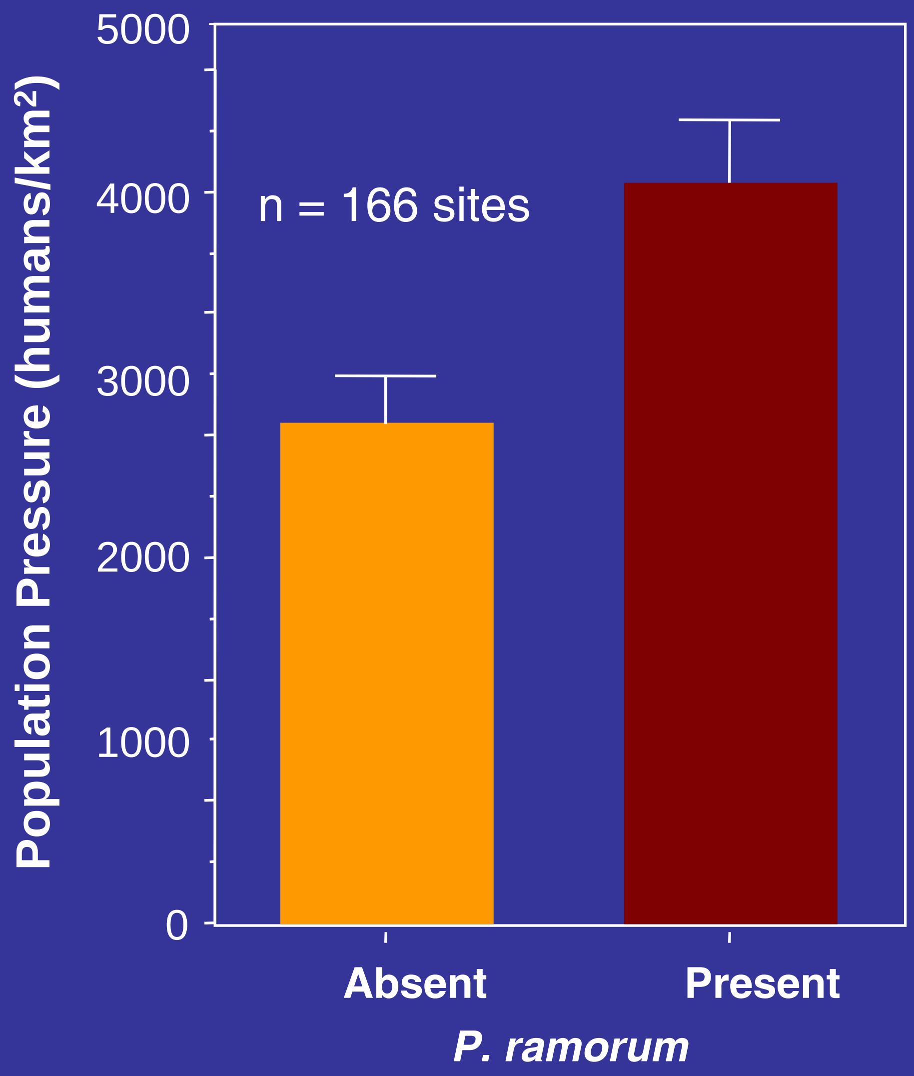
Logistic Regression (r^2 =0.51)

- Distance (-) p=0.0000
- Pop. pressure (+) p=0.0001
- Bay laurel (+) p=0.0002
- Precipitation (+) p=0.0001
- Pop. pressure *x* Precip. p=0.0001

Eliminated Variables

- Min. & Max. Temperature
- Relative Humidity
- Fire History
- Tanoak presence

Statistically-Based Model



Logistic Regression (r^2 =0.51)

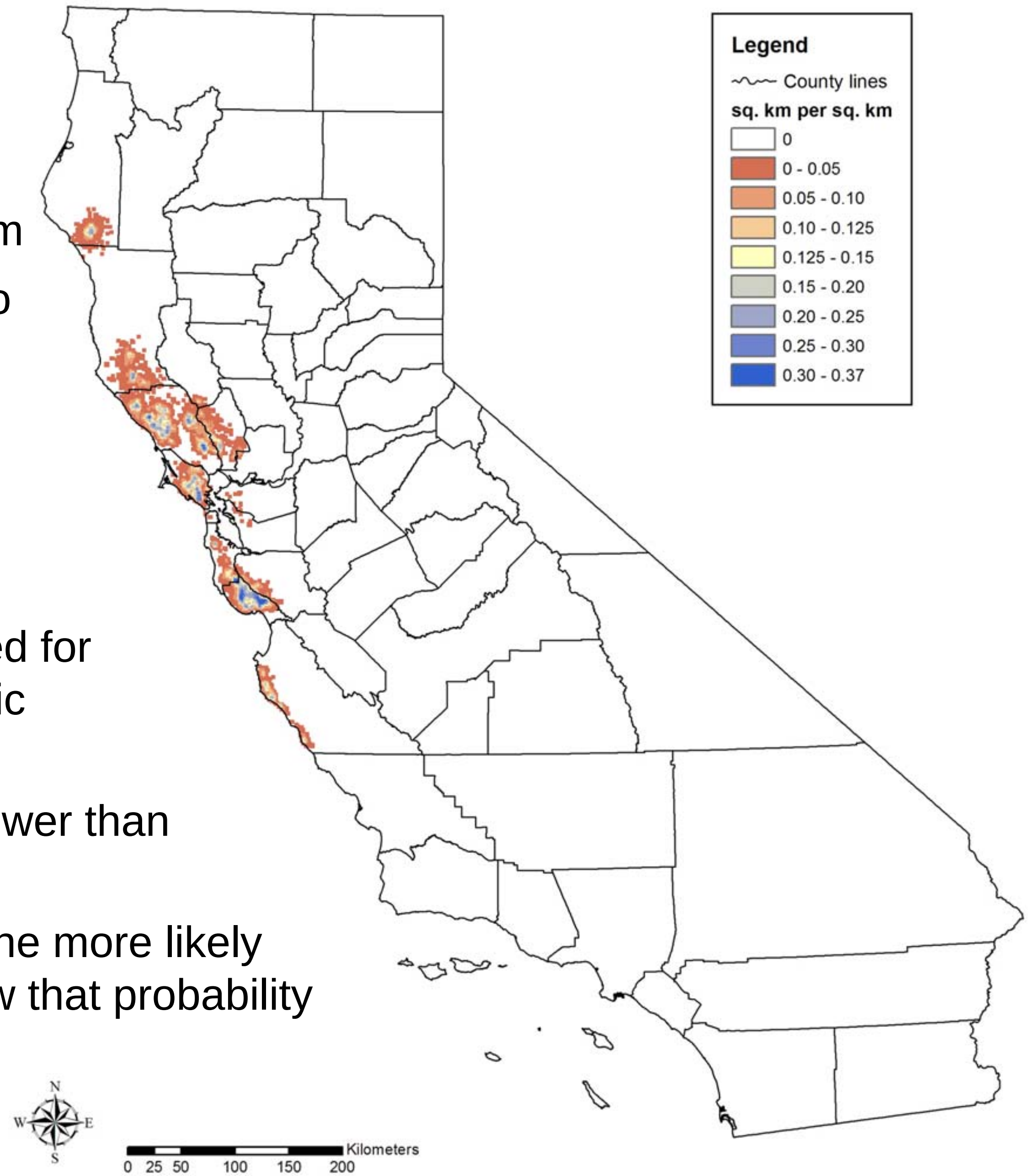
Distance (-)	p=0.0000
Pop. pressure (+)	p=0.0001
Bay laurel (+)	p=0.0002
Precipitation (+)	p=0.0001
Pop. pressure *x* Precip.	p=0.0001

Eliminated Variables

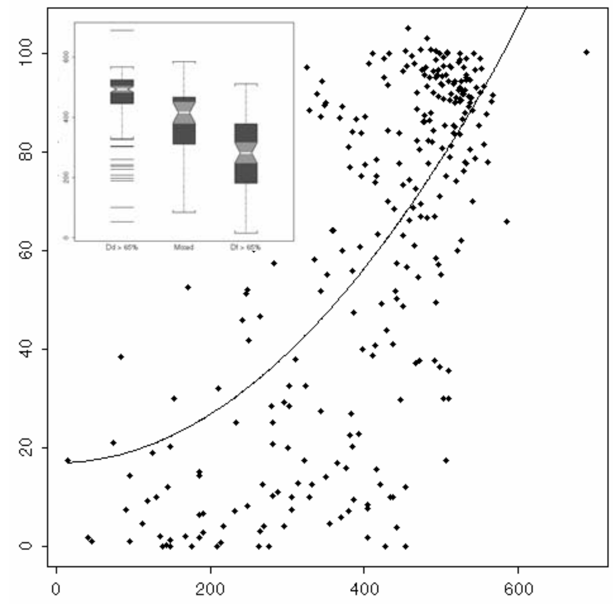
Min. & Max. Temperature
Relative Humidity
Fire History
Tanoak presence

Application and Simulation in GIS

- Mapped all variables at 50 x 50 m
- logistic regression was applied to each variable to produce a map of occurrence probability (0-1)
- used Monte Carlo simulation to predict distribution across unsampled regions
- a random number (0-1) generated for each cell and compared to logistic probability
- cell infected if random number lower than probability of infection
- higher the predicted probability the more likely the random number will fall below that probability



Statistical Modeling



Strengths

- established methods like regression
- empirical relationships across study system facilitate ecological and epidemiological understanding

Limitations

- susceptible sites have not yet been exposed (distance variables partially offset this)
- probabilistic, but still static (not dynamic)
- need dynamic models of establishment and spread driven by weather events and dispersal vectors

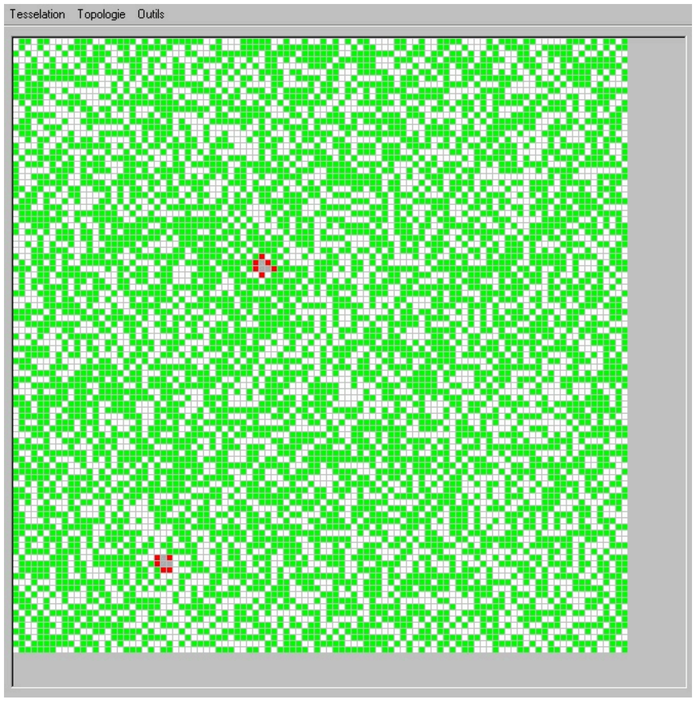
III. Cellular Automata Modeling

- powerful approach for modeling processes driving establishment & spread

- may conduct repeatable non-destructive experiments at large scales
  - effect of extreme weather
  - efficacy of management practices

- requires parameterization of system components across a grid and specified time steps (e.g. 1hr, 1day)

Spread of a Wildfire



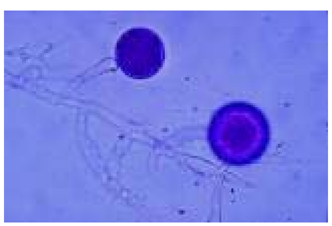
System Components of *P. ramorum*

HOST

>20 plant spp., 12 families

POTENTIAL VECTORS

DISEASE



P. ramorum

PATHOGEN

Moisture
Mild temps

ENVIRONMENT

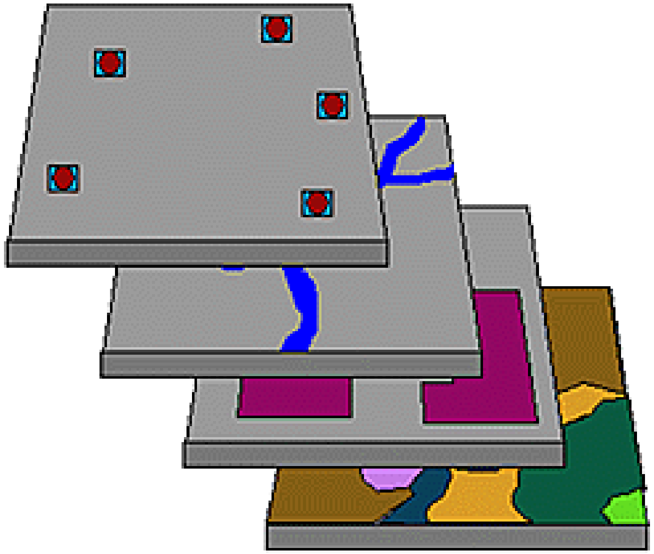
System Components

In the model

$$P_{t+1} \approx \frac{S * P * T * I * H}{D}$$

P = Probability of infection

S = Host susceptibility
T = Temperature
P = Precipitation
I = Surrounding disease levels
D = Distance
H = Human vectors



Model Approach

- Probability of infection modeled for each cell based on system components
  - Number between zero and one

Model landscape
Time step 1

0.08	0.21	0.11	0	0
0	0.15	0.50	0.01	0
0.04	0.02	0.13	0.18	0
0.42	0	0.26	0.06	0.22
0	0.26	0	0.04	0

Model Approach

- Probability of infection modeled for each cell based on system components
  - Number between zero and one
- Weighted random process
  - Generate random numbers between zero and one

Model landscape
Time step 1

0.08 0.15	0.21 0.23	0.11 0.79	0	0
0	0.15 0.10	0.50 0.52	0.01 0.34	0
0.04 0.07	0.02 0.86	0.13 0.25	0.18 0.97	0
0.42 0.64	0	0.26 0.29	0.06 0.11	0.22 0.73
0	0.26 0.84	0	0.04 0.12	0

Model Approach

- Probability of infection modeled for each cell based on system components
  - Number between zero and one
- Weighted random process
  - Generate random numbers between zero and one
- Compare probability to random numbers
  - Infected – *Probability > Random*
  - Uninfected – *Probability < Random*
  - No host vegetation (0)

Model landscape
Time step 1

0.08 0.15	0.21 0.23	0.11 0.79	0	0
0	0.15 0.10	0.50 0.52	0.01 0.34	0
0.04 0.07	0.02 0.86	0.13 0.25	0.18 0.97	0
0.42 0.64	0	0.26 0.29	0.06 0.11	0.22 0.73
0	0.26 0.84	0	0.04 0.12	0

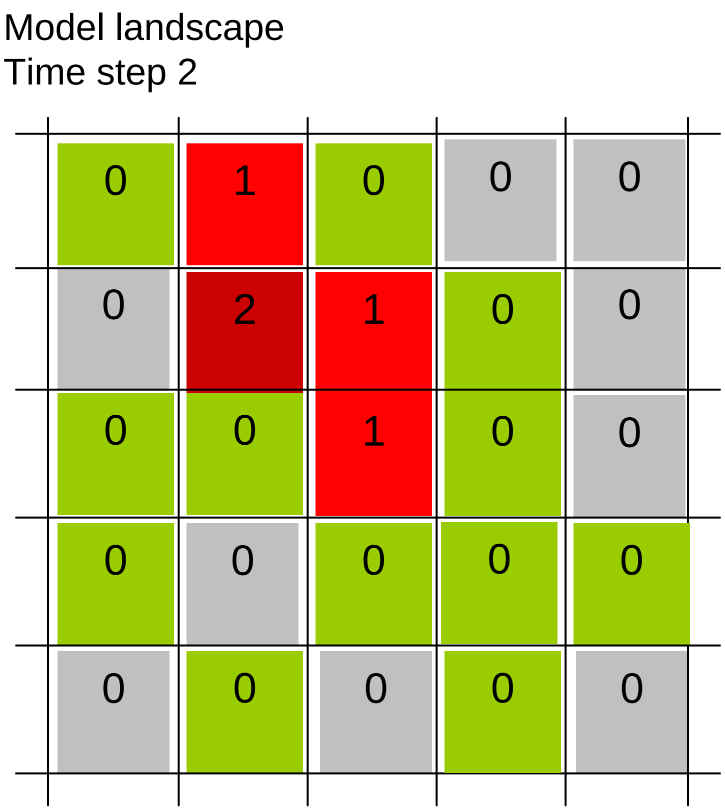
Model Approach

- Probability of infection modeled for each cell based on system components
  - Number between zero and one
- Weighted random process
  - Generate random numbers between zero and one
- Compare probability to random numbers
  - Infected – *Probability > Random*
  - Uninfected – *Probability < Random*
  - No host vegetation (0)
- Outputs disease intensity as count of times cell is infected

Model landscape
Time step 1

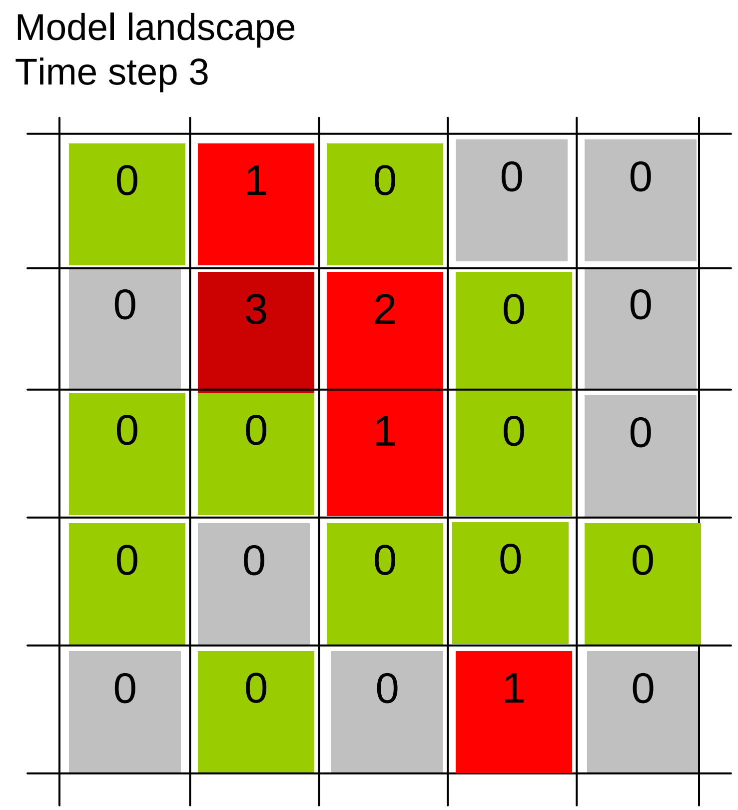
Model Approach

- Probability of infection modeled for each cell based on system components
  - Number between zero and one
- Weighted random process
  - Generate random numbers between zero and one
- Compare probability to random numbers
  - Infected – *Probability > Random*
  - Uninfected – *Probability < Random*
  - No host vegetation (0)
- Outputs disease intensity as count of times cell is infected
- Iterates on a weekly time step



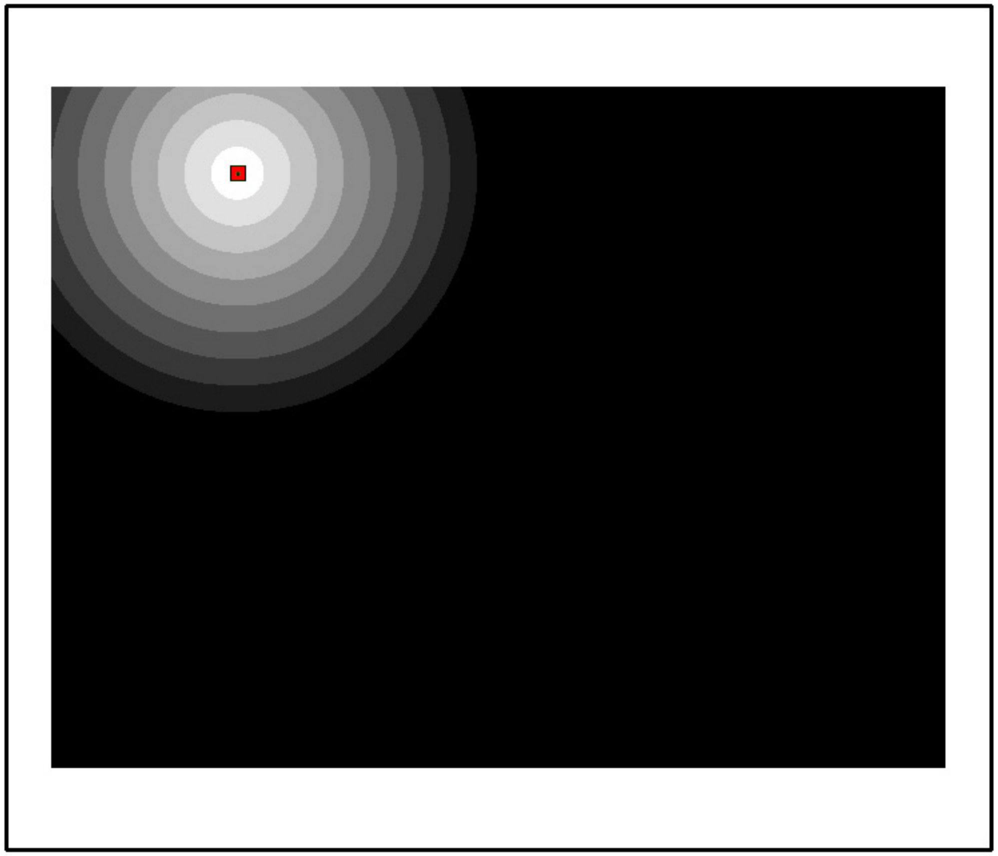
Model Approach

- Probability of infection modeled for each cell based on system components
  - Number between zero and one
- Weighted random process
  - Generate random numbers between zero and one
- Compare probability to random numbers
  - Infected – *Probability > Random*
  - Uninfected – *Probability < Random*
  - No host vegetation (0)
- Outputs disease intensity as count of times cell is infected
- Iterates on a weekly time step



Model Approach

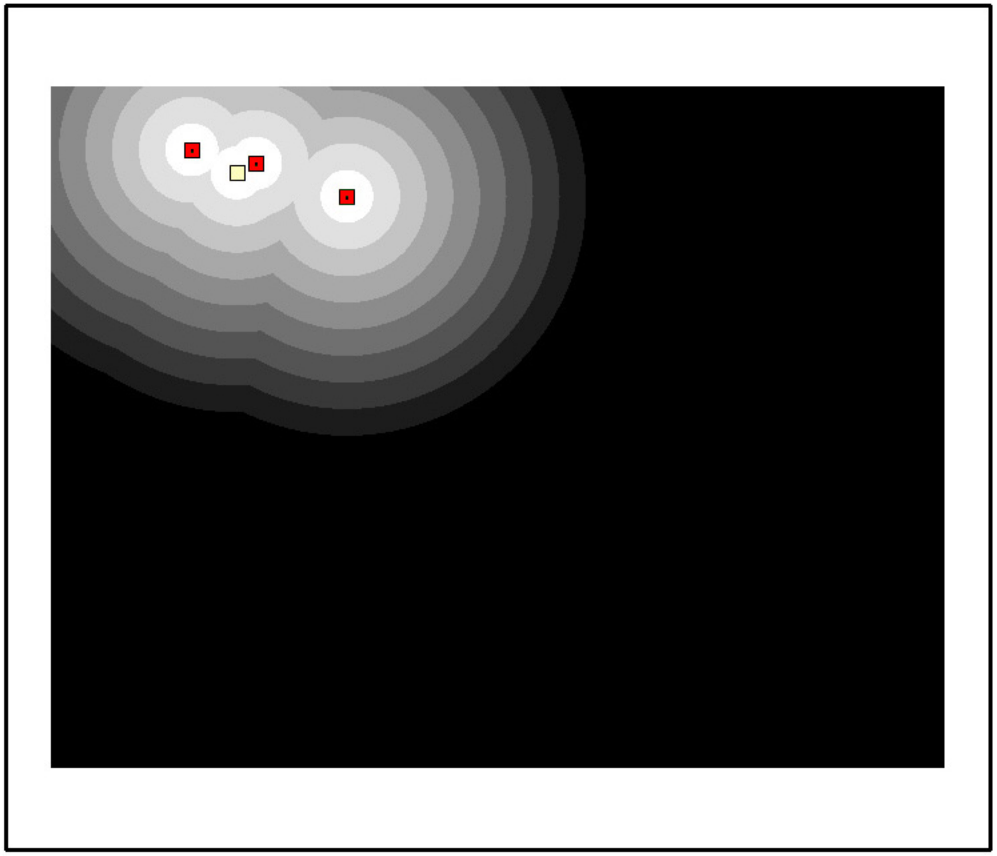
Example Simulation



introduction → establishment → dispersal → invasive spread

Model Approach

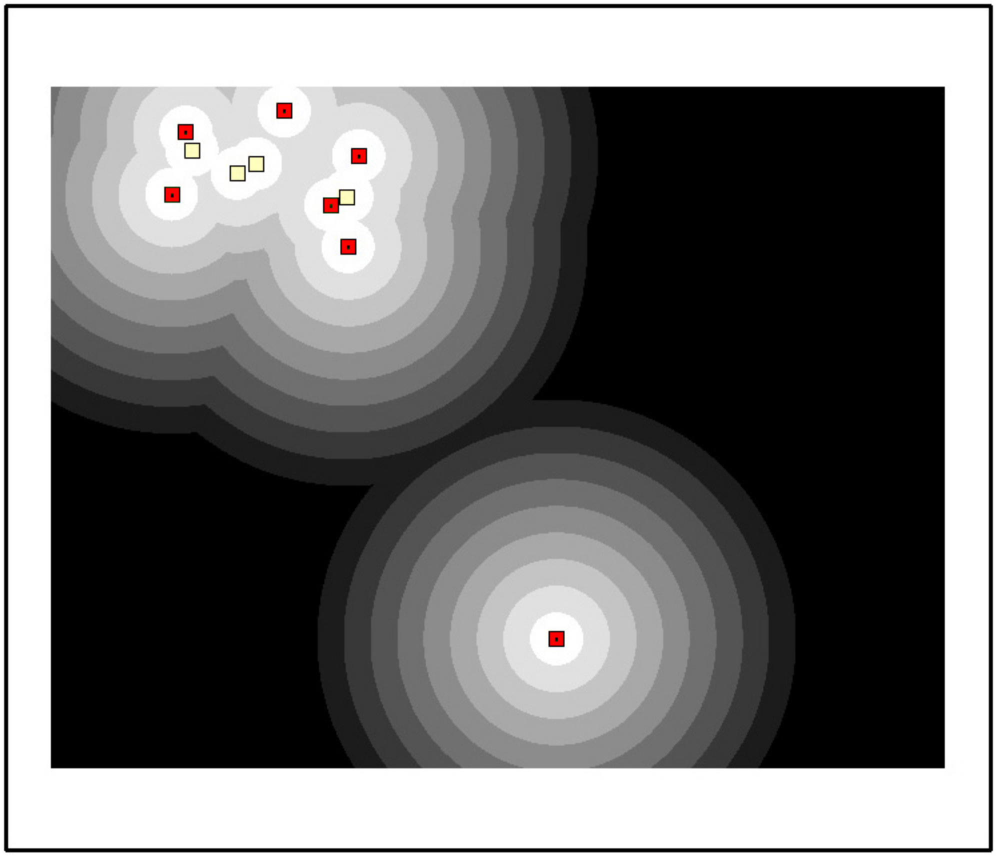
Example Simulation



introduction → establishment → dispersal → invasive spread

Model Approach

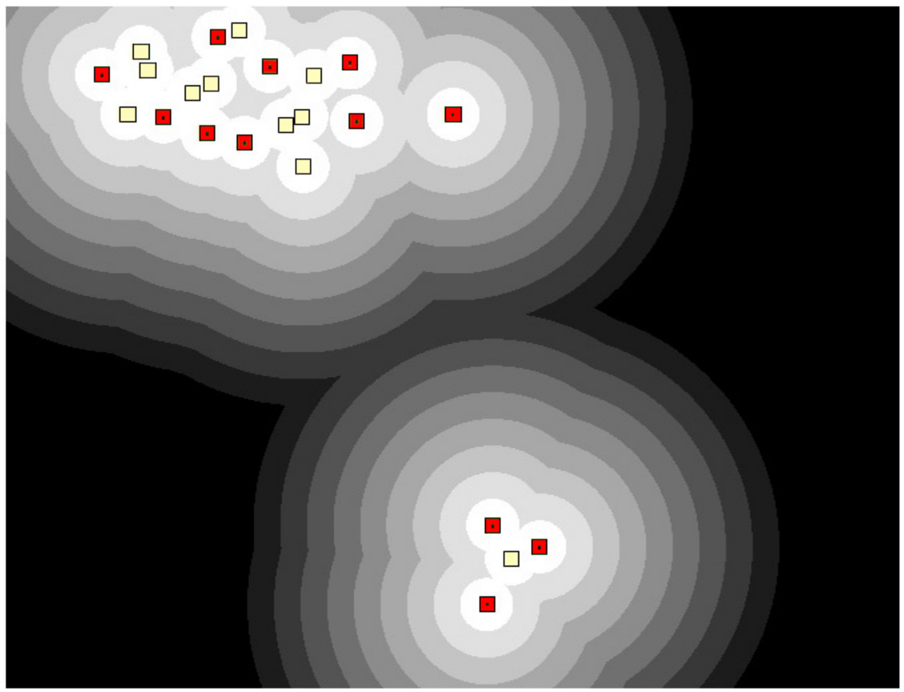
Example Simulation



introduction → establishment → dispersal → invasive spread

Model Approach

<u>Example Simulation</u>



introduction → establishment → dispersal → invasive spread

Mapping and Parameterizing System Components

Spatial Grain = 250 m cells
Time Step = 1 week

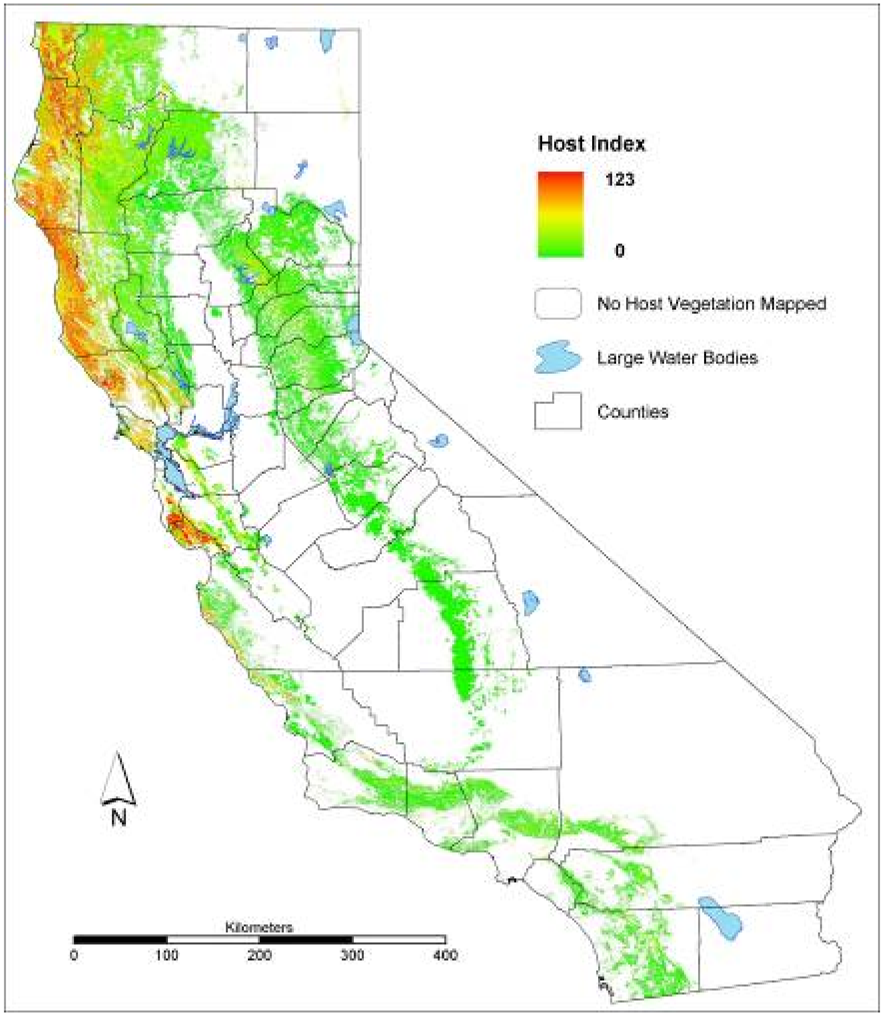
Host Susceptibility

Parameterization

Host Index (*S*)

- Meentemeyer *et al.* 2004
- Ranked susceptibility of each species
- Using CALVEG data, each cell received a composite score based on susceptibility rank and abundance of all species present
- Output scaled to 100 points

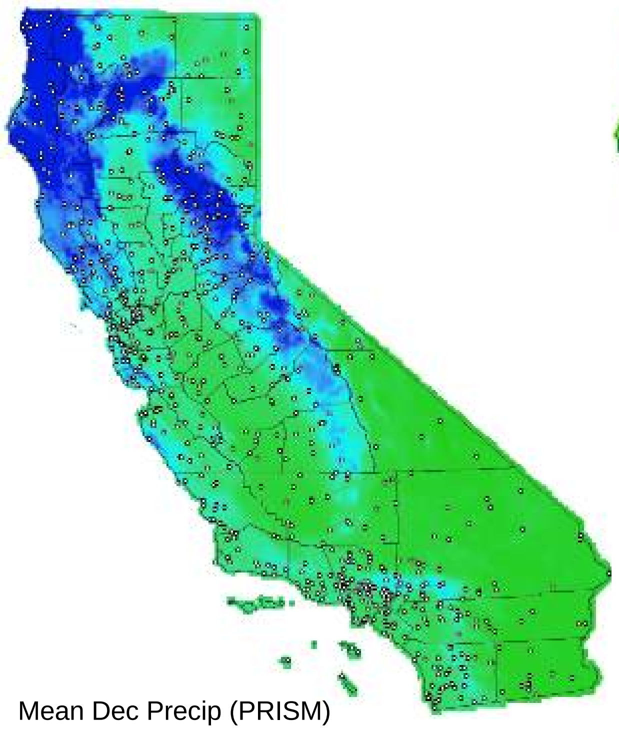
$$P_{t+1} \approx \frac{S * P * T * I * H}{D}$$



Weather Events

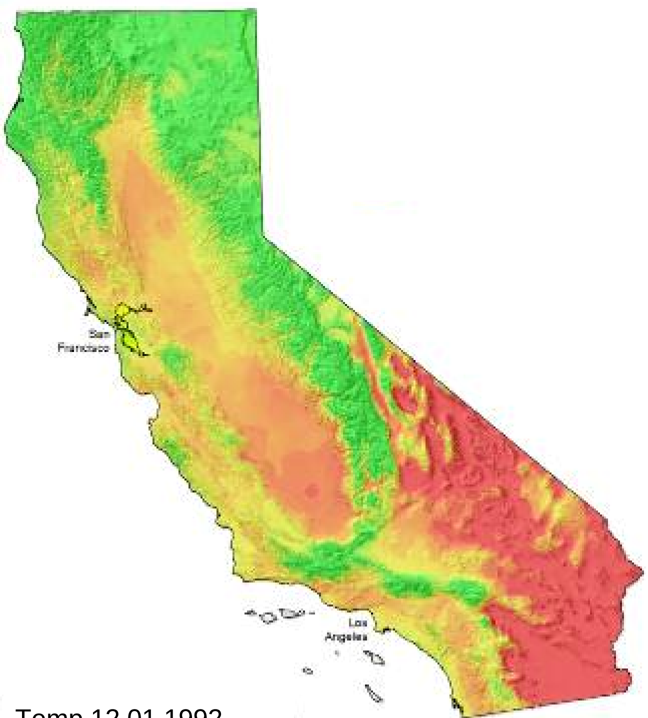
<u>Mapping Methods</u>

- Daily precip & temp at NCDC-NOAA point locations
- Integrate point data with long-term average grids (PRISM)



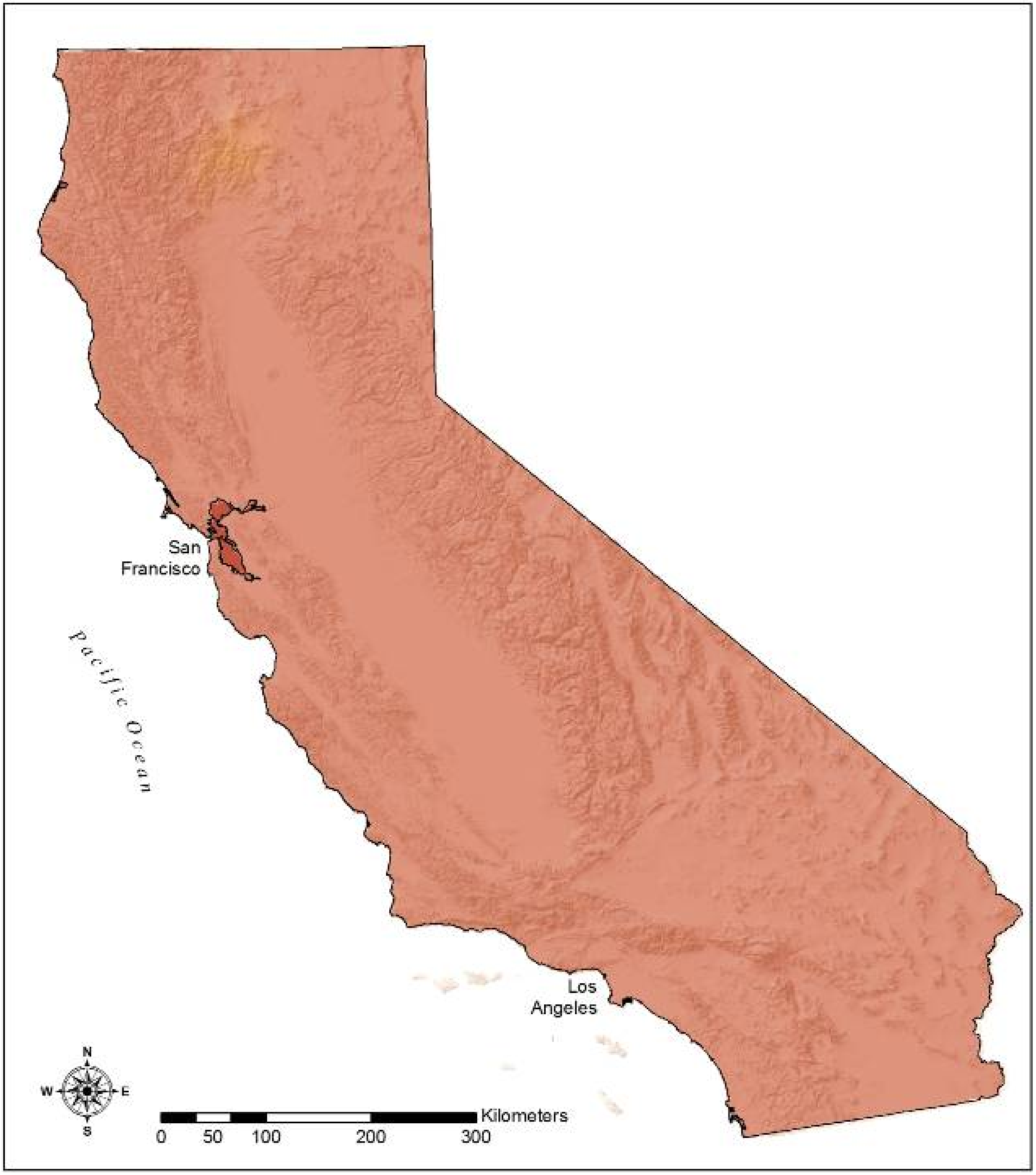
Mean Dec Precip (PRISM)
NCDC Base Stations

Precip 12.01.1992



Temp 12.01.1992

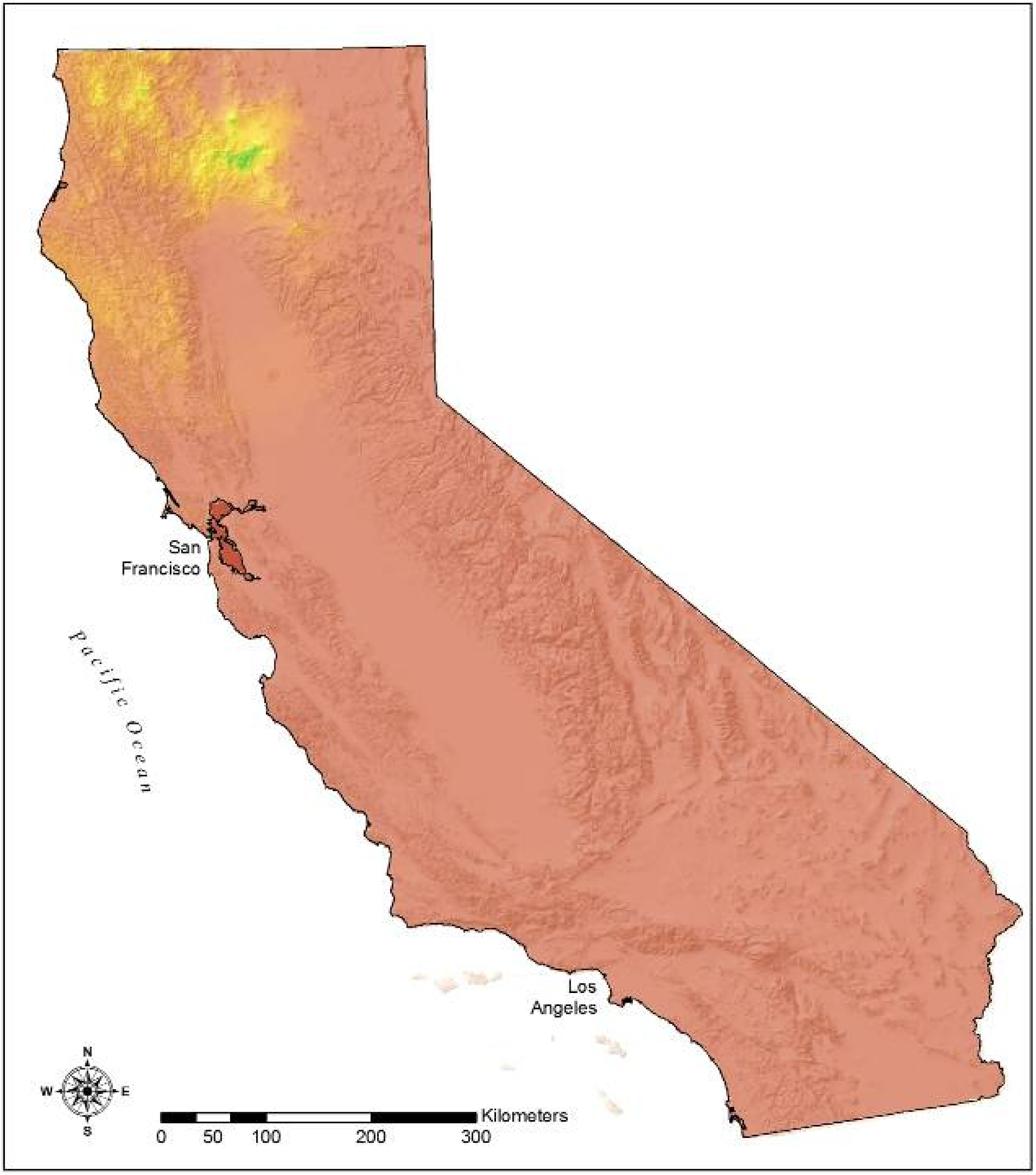
Hunter & Meentemeyer (in press) *J. Appl. Meteor.*



PRECIPITATION

March 20, 1998

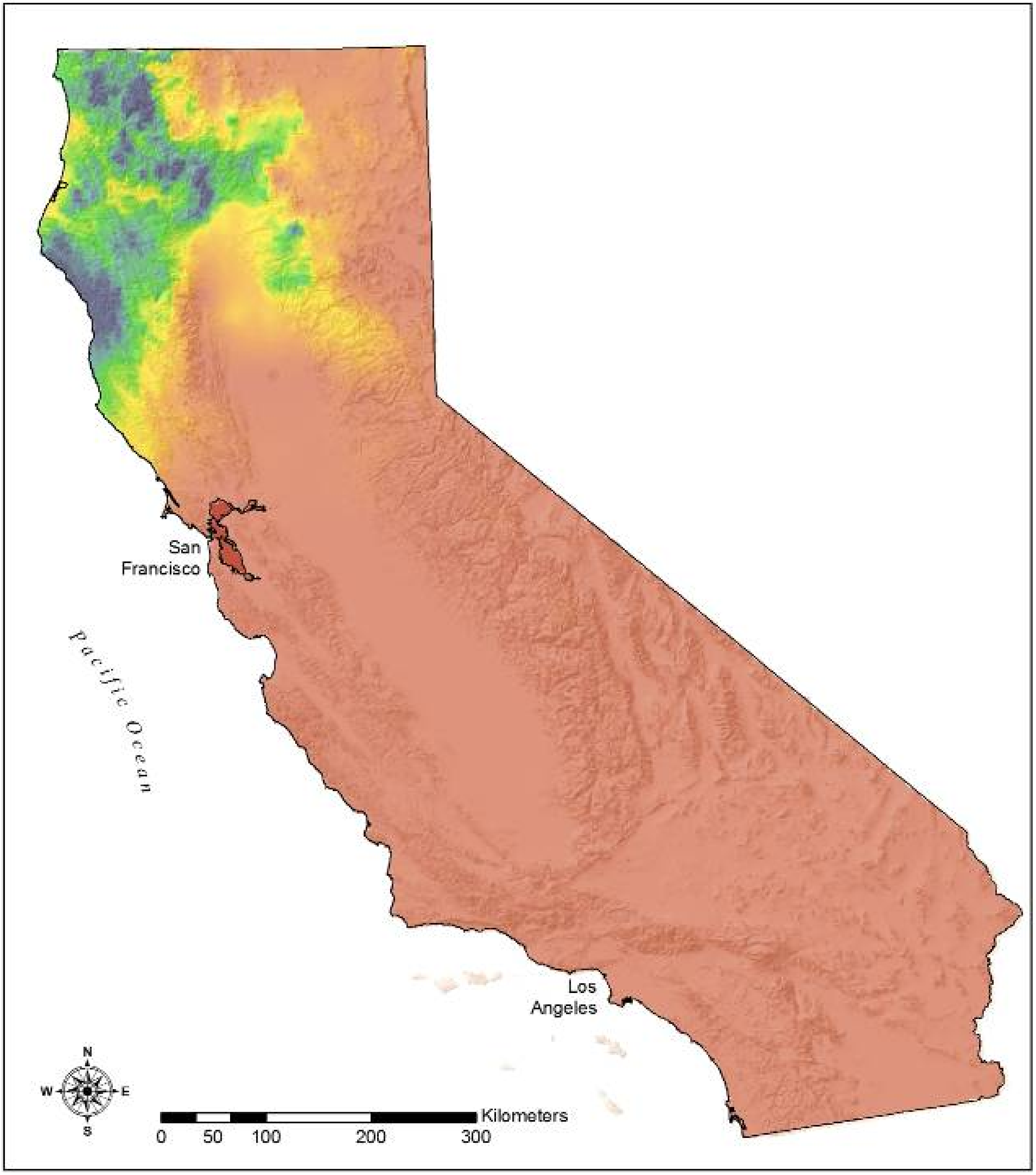
Hunter & Meentemeyer (in press)
J. Appl. Meteor.



PRECIPITATION

March 21, 1998

Hunter & Meentemeyer (in press)
J. Appl. Meteor.



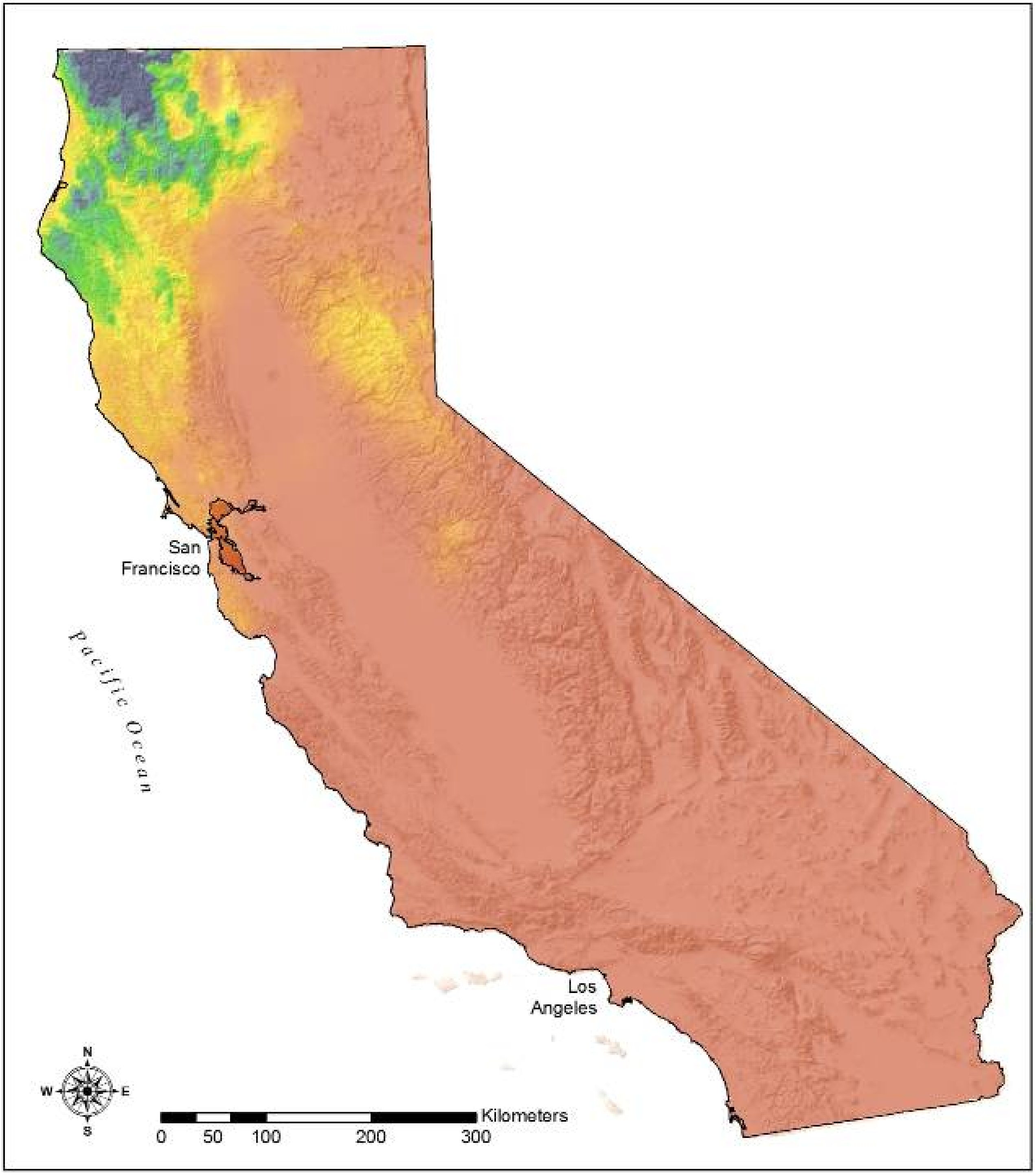
PRECIPITATION

March 22, 1998

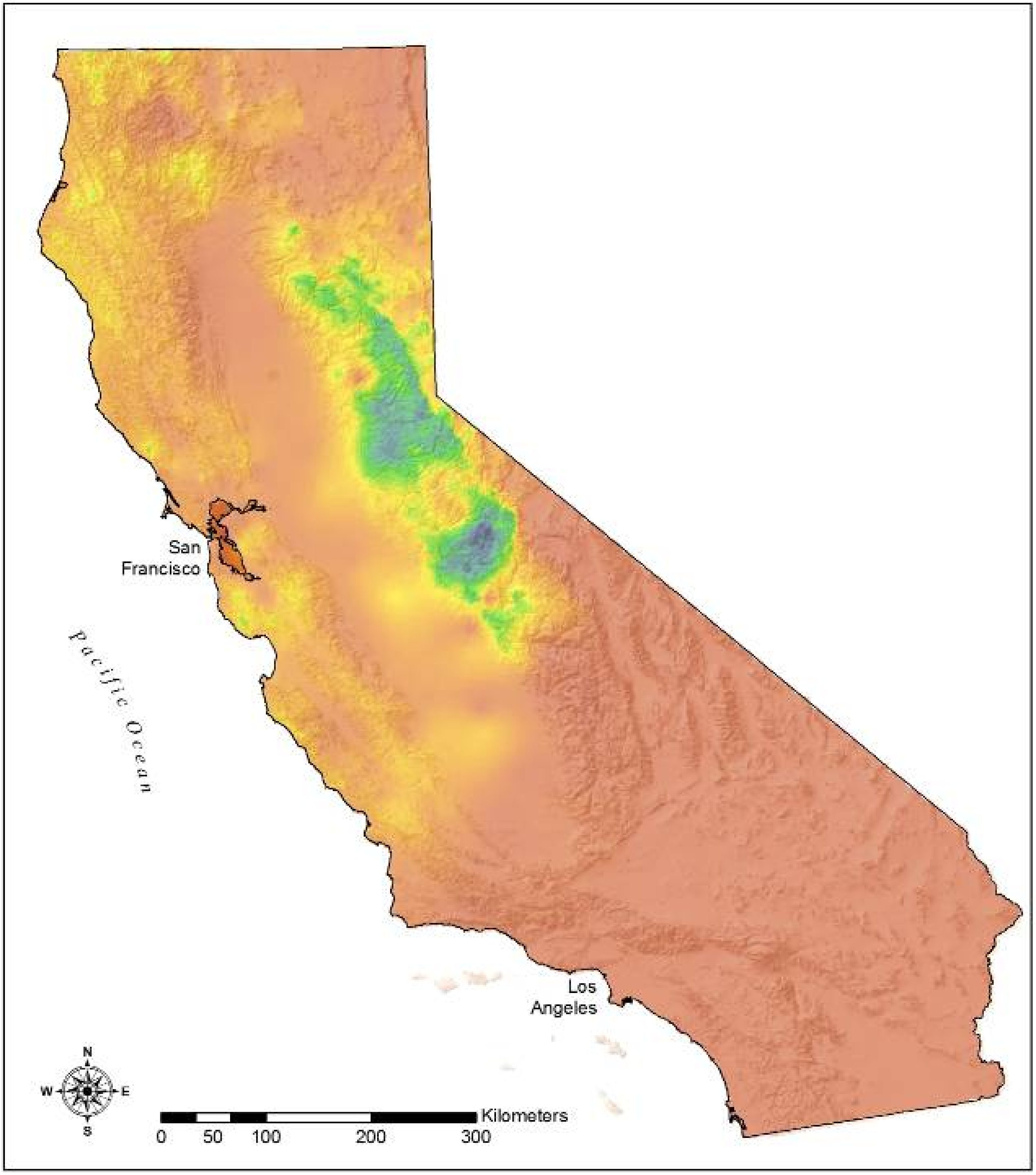
Hunter & Meentemeyer (in press)
J. Appl. Meteor.

PRECIPITATION

March 23, 1998



Hunter & Meentemeyer (in press)
J. Appl. Meteor.



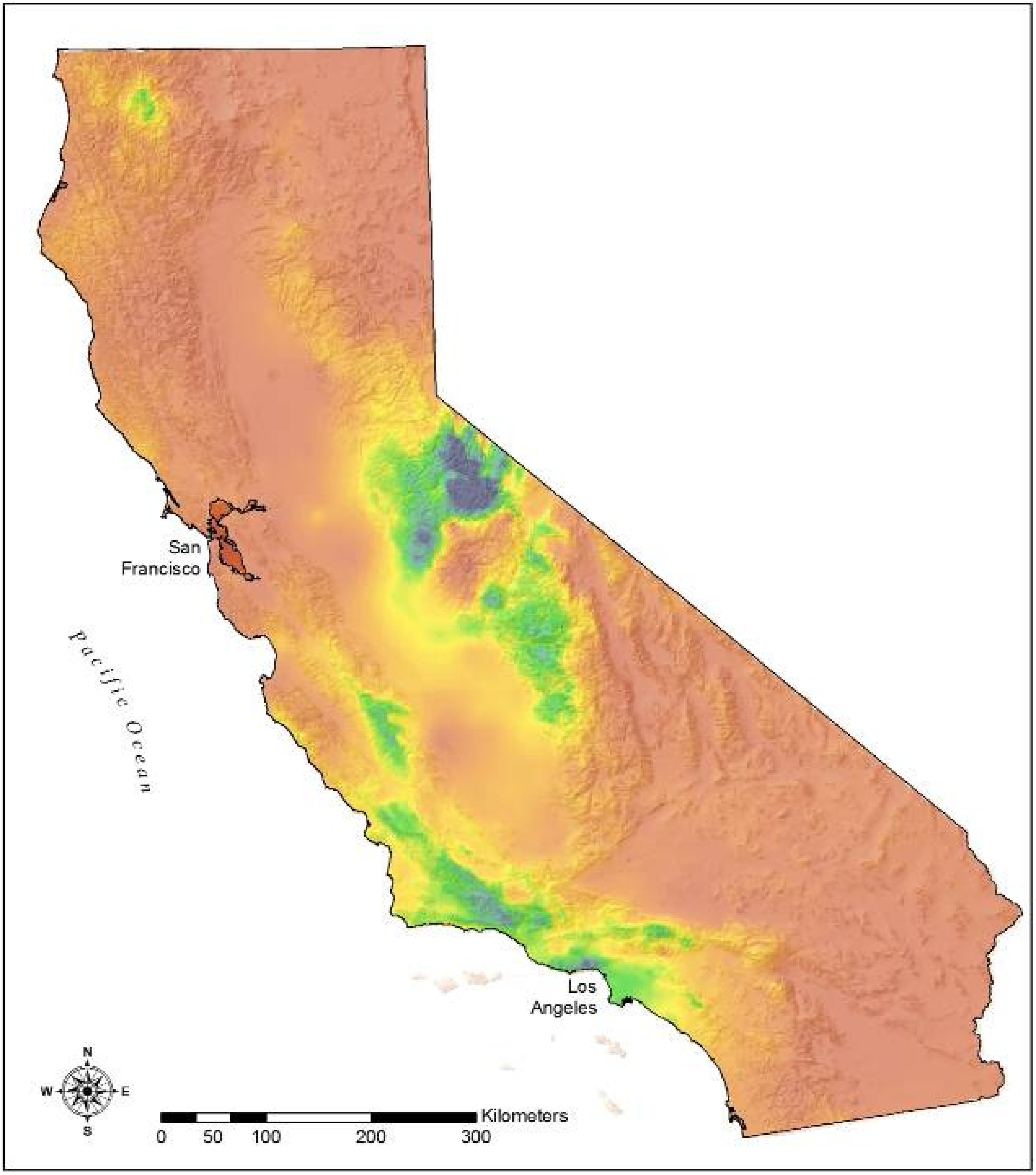
PRECIPITATION

March 24, 1998

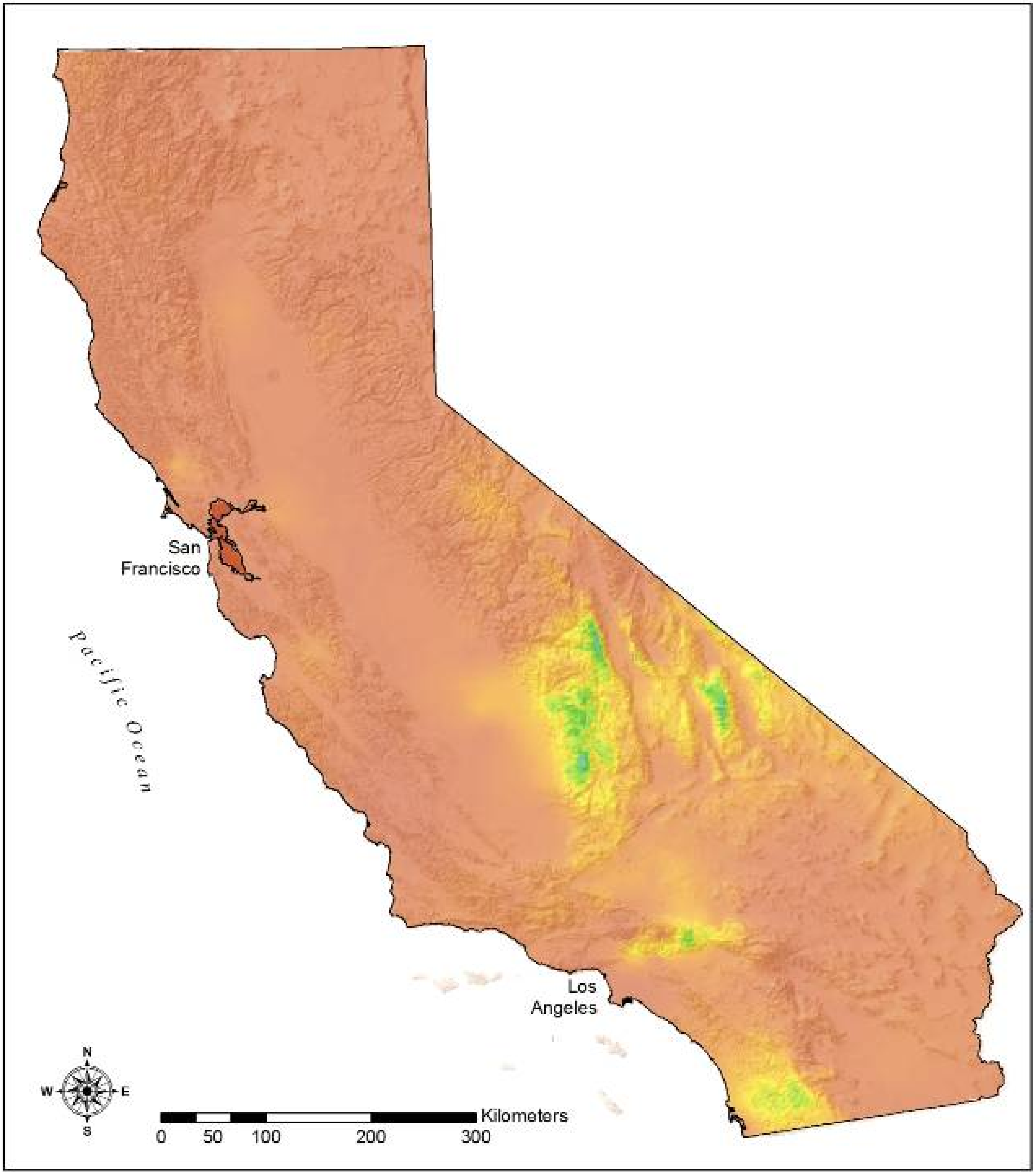
Hunter & Meentemeyer (in press)
J. Appl. Meteor.

PRECIPITATION

March 25, 1998



Hunter & Meentemeyer (in press) *J. Appl. Meteor.*



PRECIPITATION

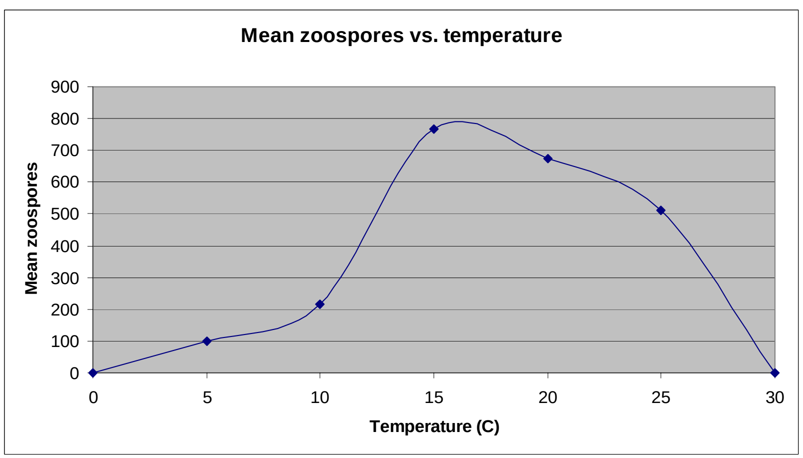
March 26, 1998

Hunter & Meentemeyer (in press)
J. Appl. Meteor.

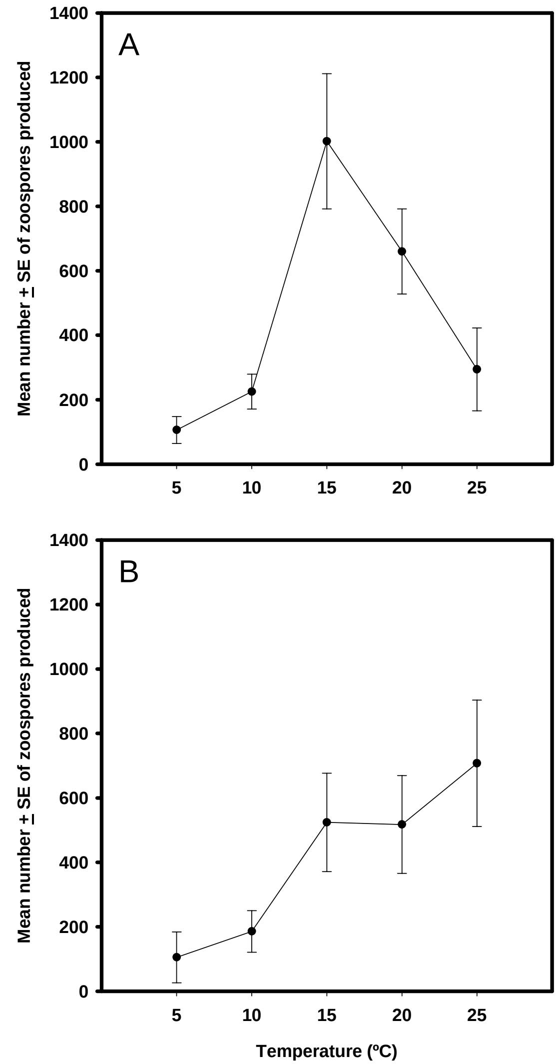
Weather Events

Parameterizing Temperature

- Davidson *et al.* (2005) lab studies
  - Measured zoospore production on bay
  - 7 set temps, 2 trials, 2 sets of 10 leaves each trial
- In the model
  - Average 2 trials
  - Fit curvilinear equation, scaled to 10 point scale
- Coded into model using weekly mean, max daily



$$P_{t+1} \approx \frac{S * P * T * I * H}{D}$$

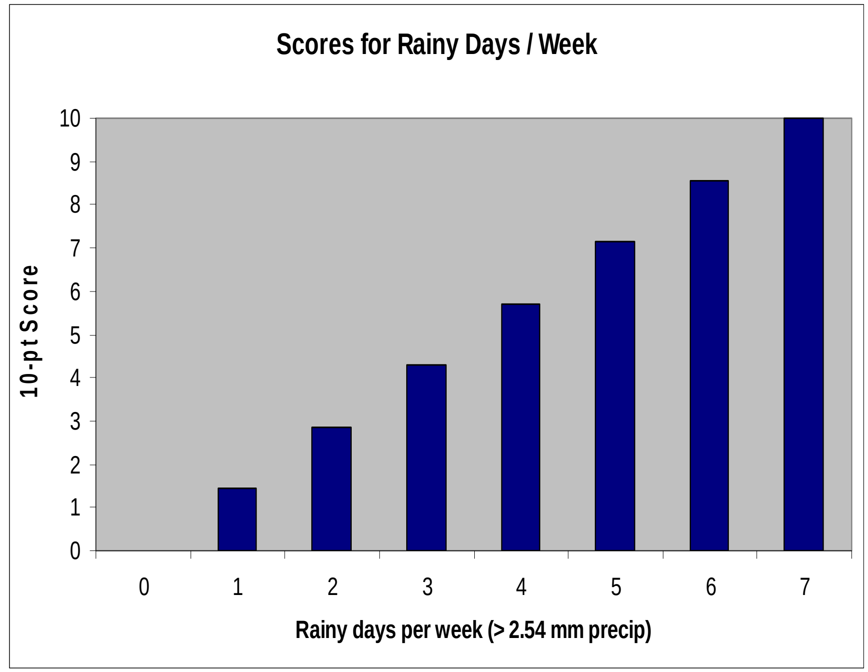


Davidson *et al.* 2005 *Phytopathology*

Weather Events

Parameterizing Precipitation

- Garbelotto / Davidson studies
  - 9-12 hrs of leaf moisture required for significant infection
- In the model
  - Rainy days are >2.5 mm precip
  - Assume linear relationship between potential sporulation and # of wet days per week
  - Scale # of wet days per week to 10 point scale

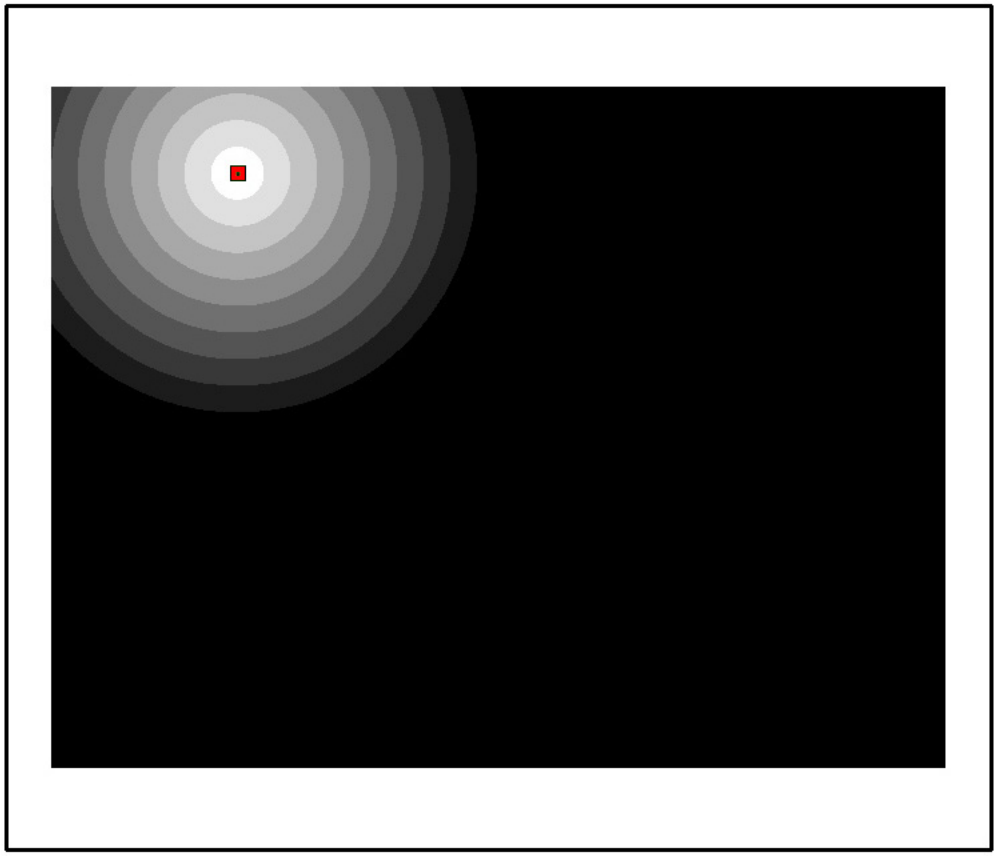


$$P_{t+1} \approx \frac{S * P * T * I * H}{D}$$

Dispersal distance

Mapping Methods

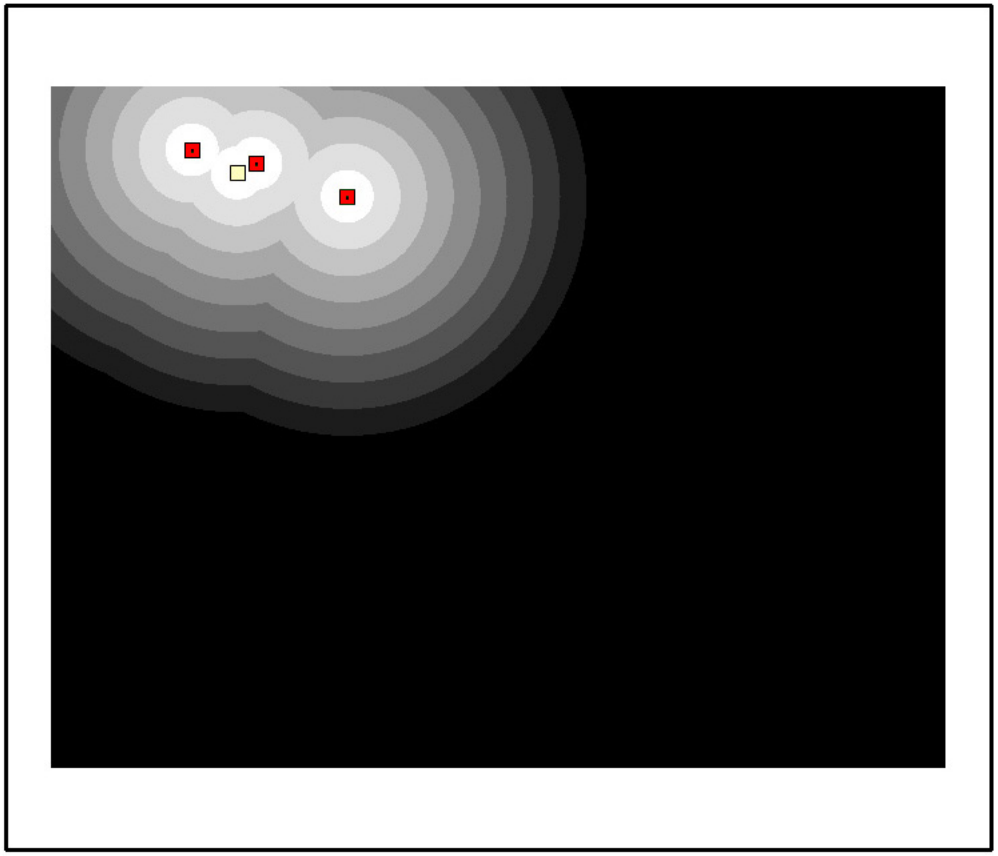
- Straight-line distance to nearest infected site
- Recalculated for each time step



Dispersal distance

Mapping Methods

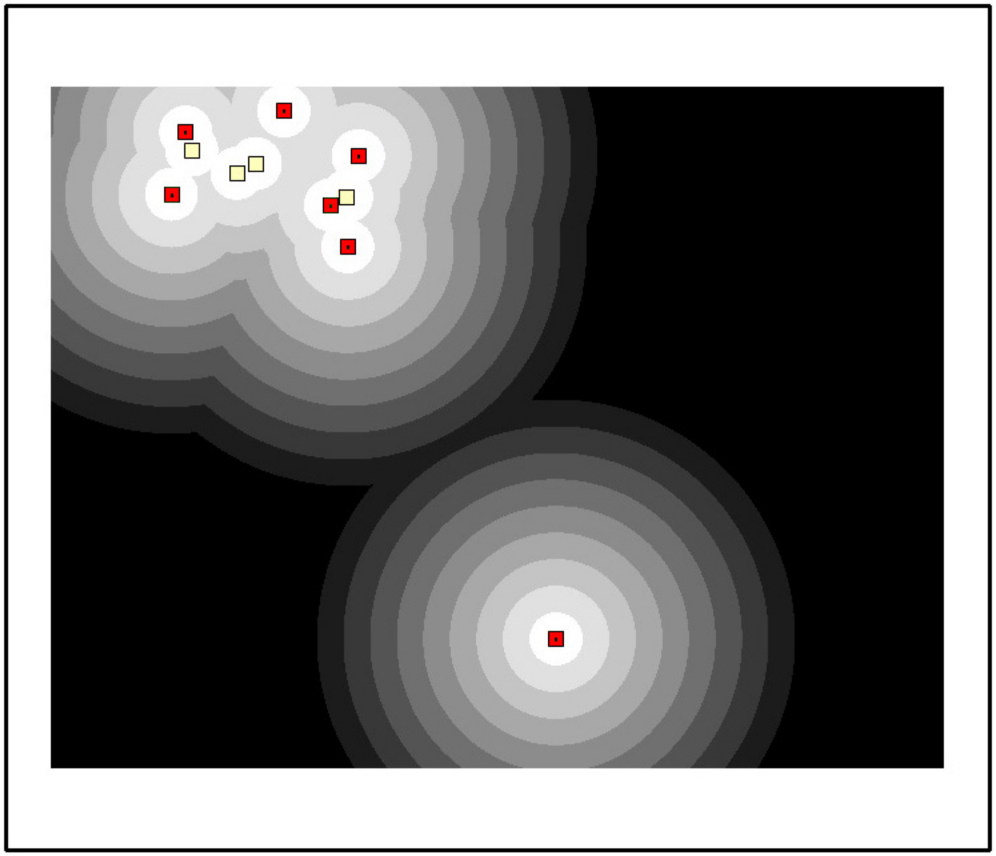
- Straight-line distance to nearest infected site
- Recalculated for each time step



Dispersal distance

Mapping Methods

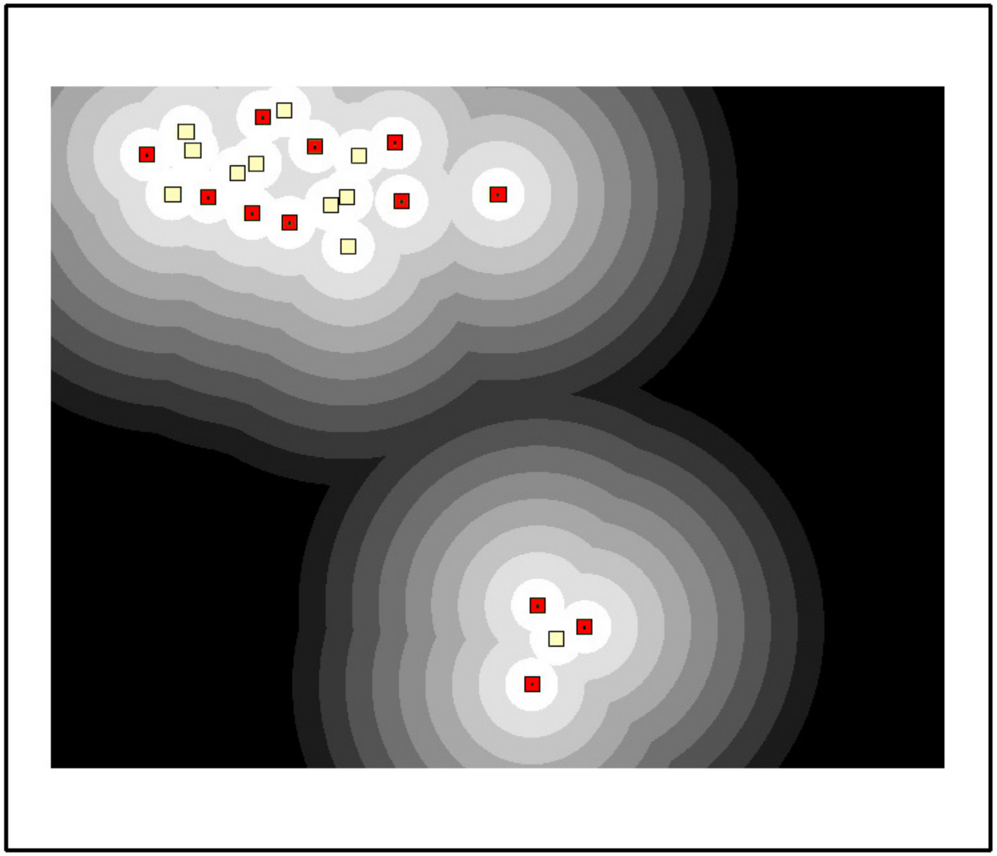
- Straight-line distance to nearest infected site
- Recalculated for each time step



Dispersal distance

Mapping Methods

- Straight-line distance to nearest infected site
- Recalculated for each time step



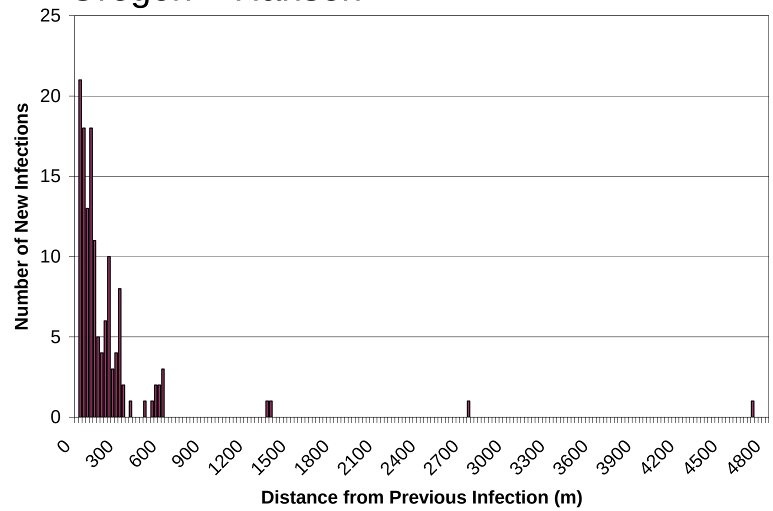
Dispersal distance

Parameterization

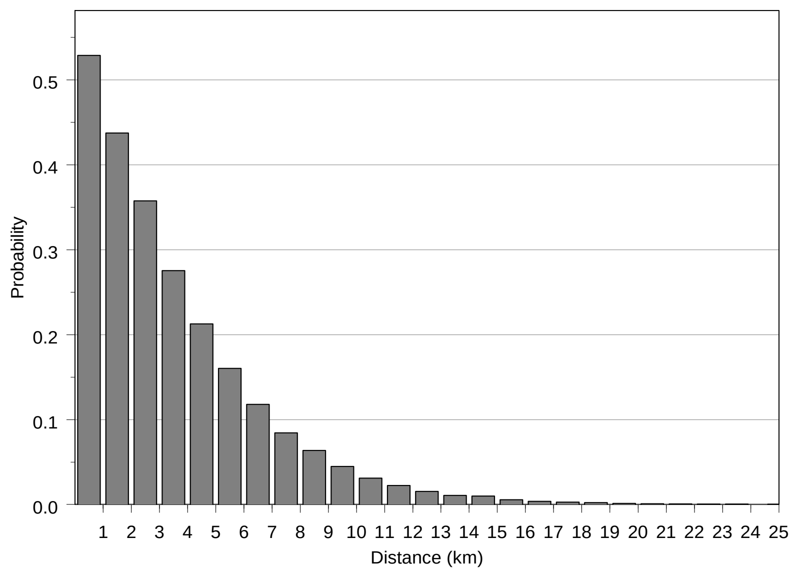
- 2 distance decay studies are being explored
- Model is programmed based on these curves and other potential curves

$$P_{t+1} \approx \frac{S*P*T*I*H}{D}$$

Oregon – Hansen



California – Meentemeyer



Surrounding disease levels (I)

- **Assumption:** higher pathogen abundance in surrounding cells increases chance of infection
- In the model
  - Output an index of pathogen abundance at each time step

    i.e. the count that each cell has been infected
  - To account for surrounding disease levels, sum the counts within a 1250m rectangle and scale to 10 pts

250m

0	1	0	0	0
0	3	2	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	1	0

Output from previous example

=

----------------1250m---------------

		8		

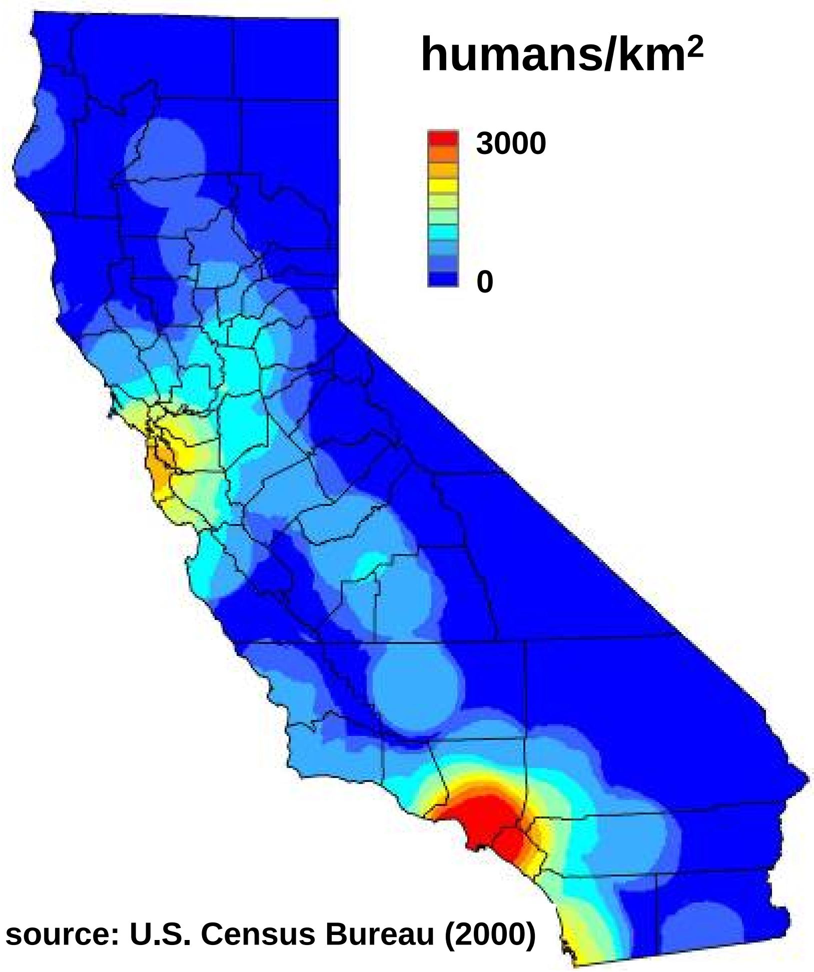
Output of surrounding pathogen abundance for an individual cell

$$P_{t+1} \approx \frac{S * P * T * I * H}{D}$$

Potential long-distance vectors

Mapping Methods

- *P. ramorum* disperses locally through several physical pathways
  - i.e. rain splash, wind-driven rain, stream water
- Human activity poses highest risk for long-distance dispersal
- Human population density within 50km radius areas



source: U.S. Census Bureau (2000)

Potential long-distance vectors (H)

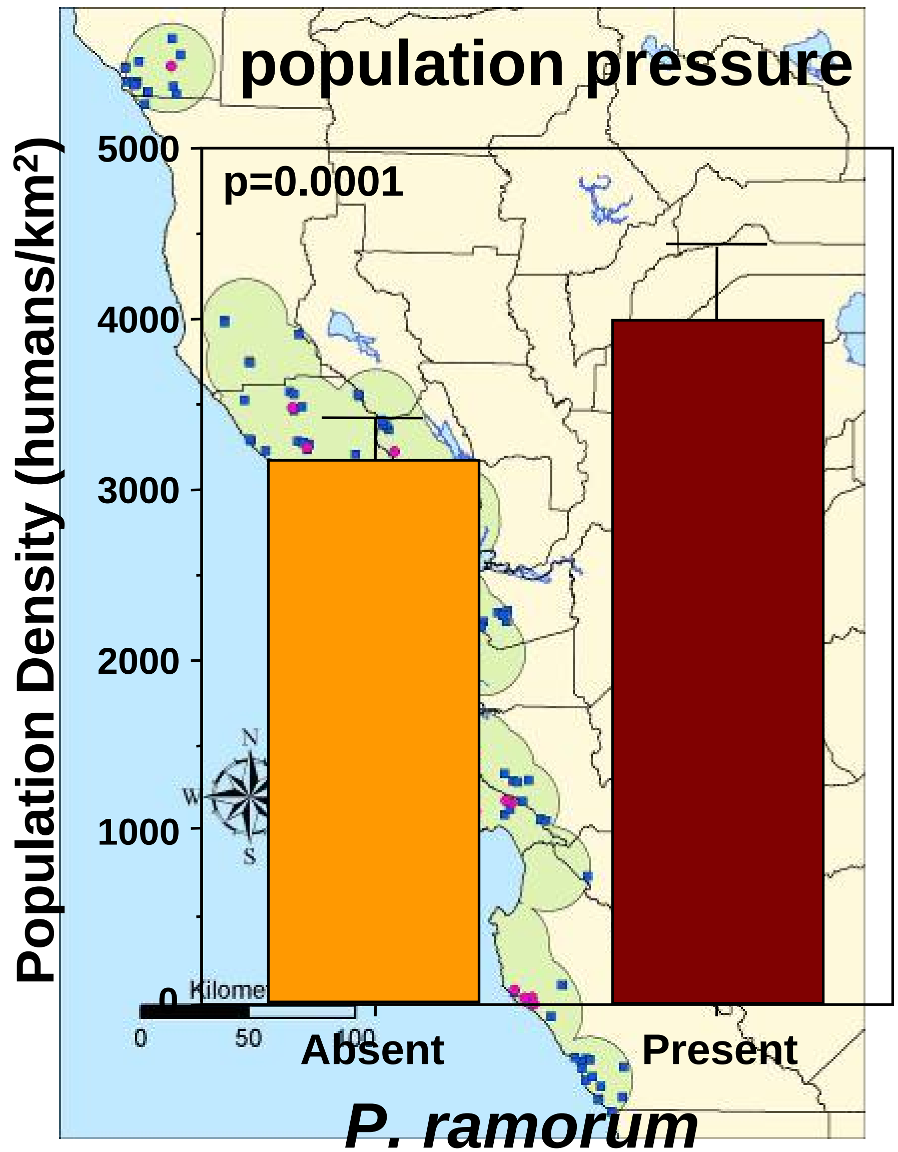
<u>Parameterization</u>

- Logistic regression
- Response variable
  - Presence/absence
  - Statewide foliar survey data (Meentemeyer) n=166

- Determine relationship between pop density and infection probability while controlling for host vegetation and climate

- Scale probabilities to 10 point scale for integration with other parameters

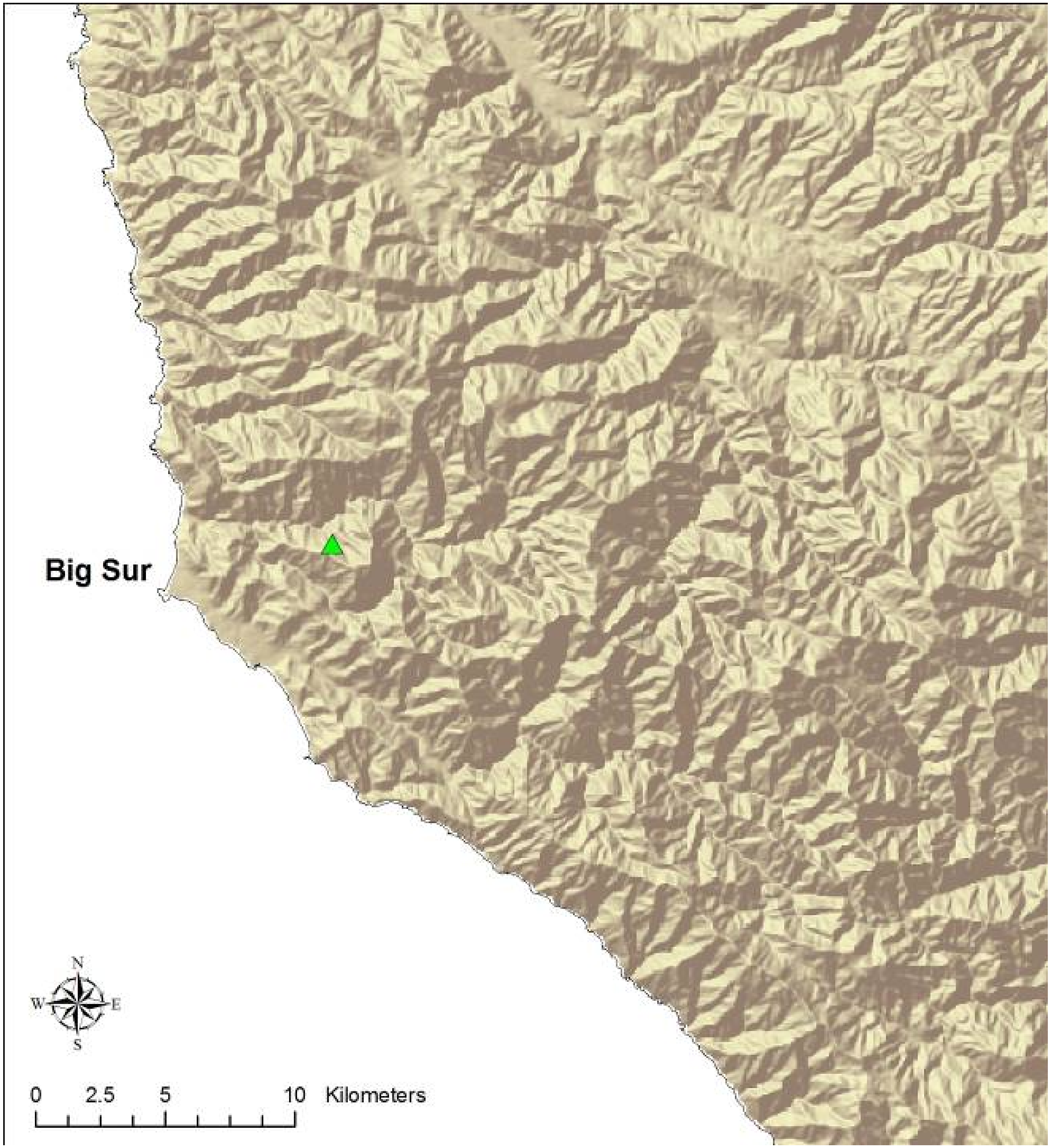
$$P_{t+1} \approx \frac{S*P*T*I*H}{D}$$



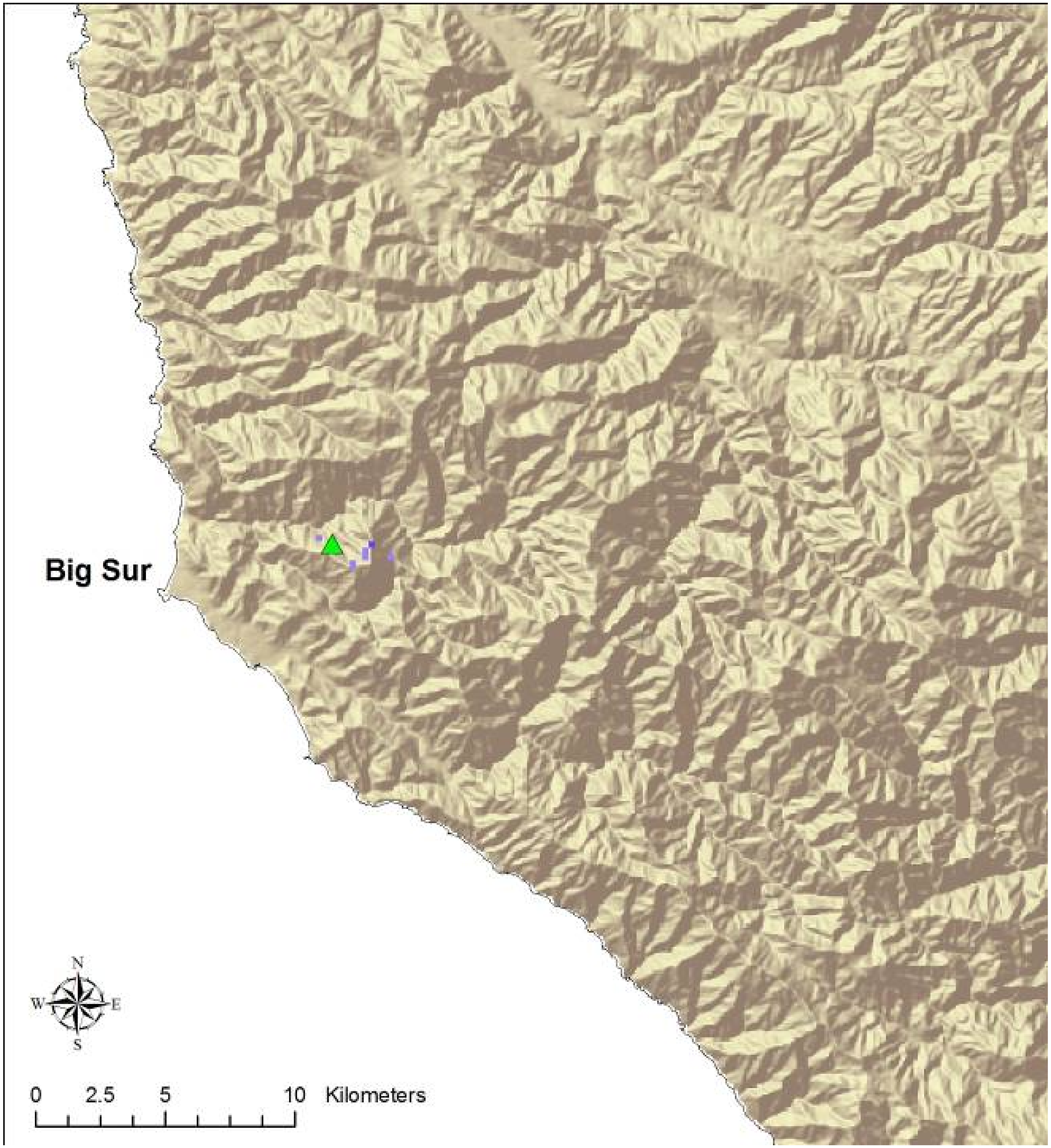
Applying the model

Example model run

- Initiated spread with a single infected cell
  - Large patch of highly susceptible vegetation near Big Sur
- Initiated in 1985
  - Weekly time step through Dec 2003

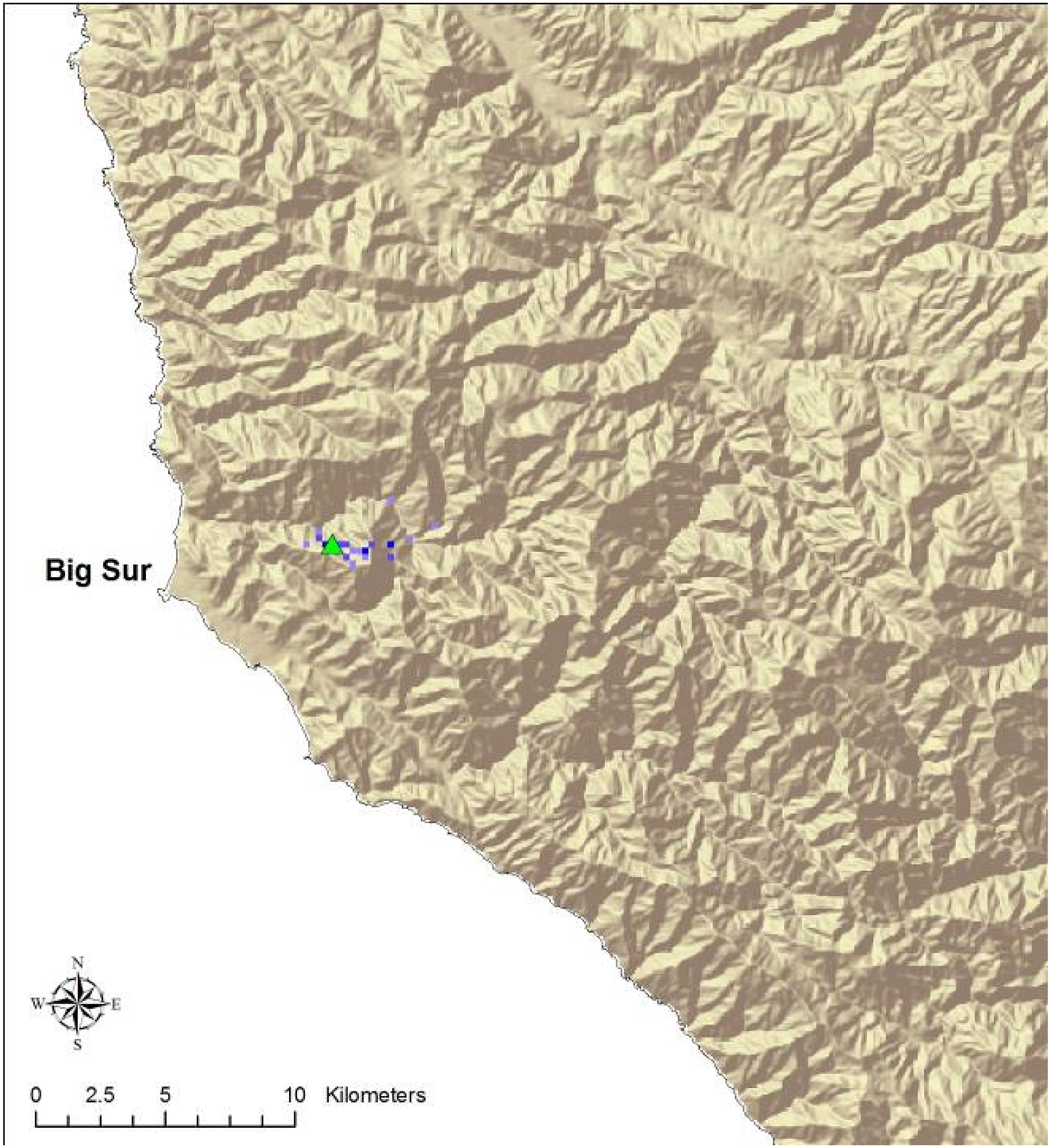


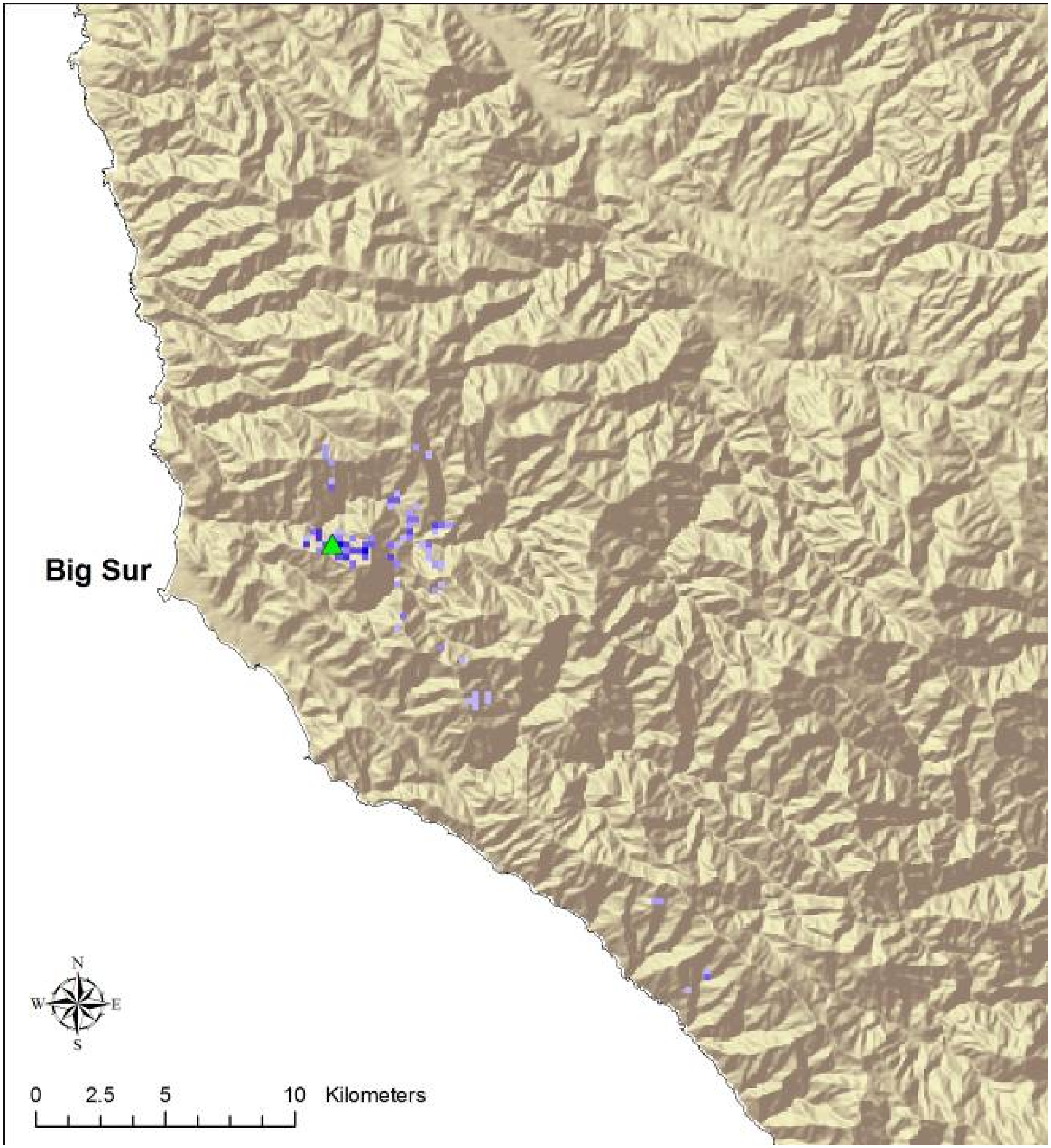
1985



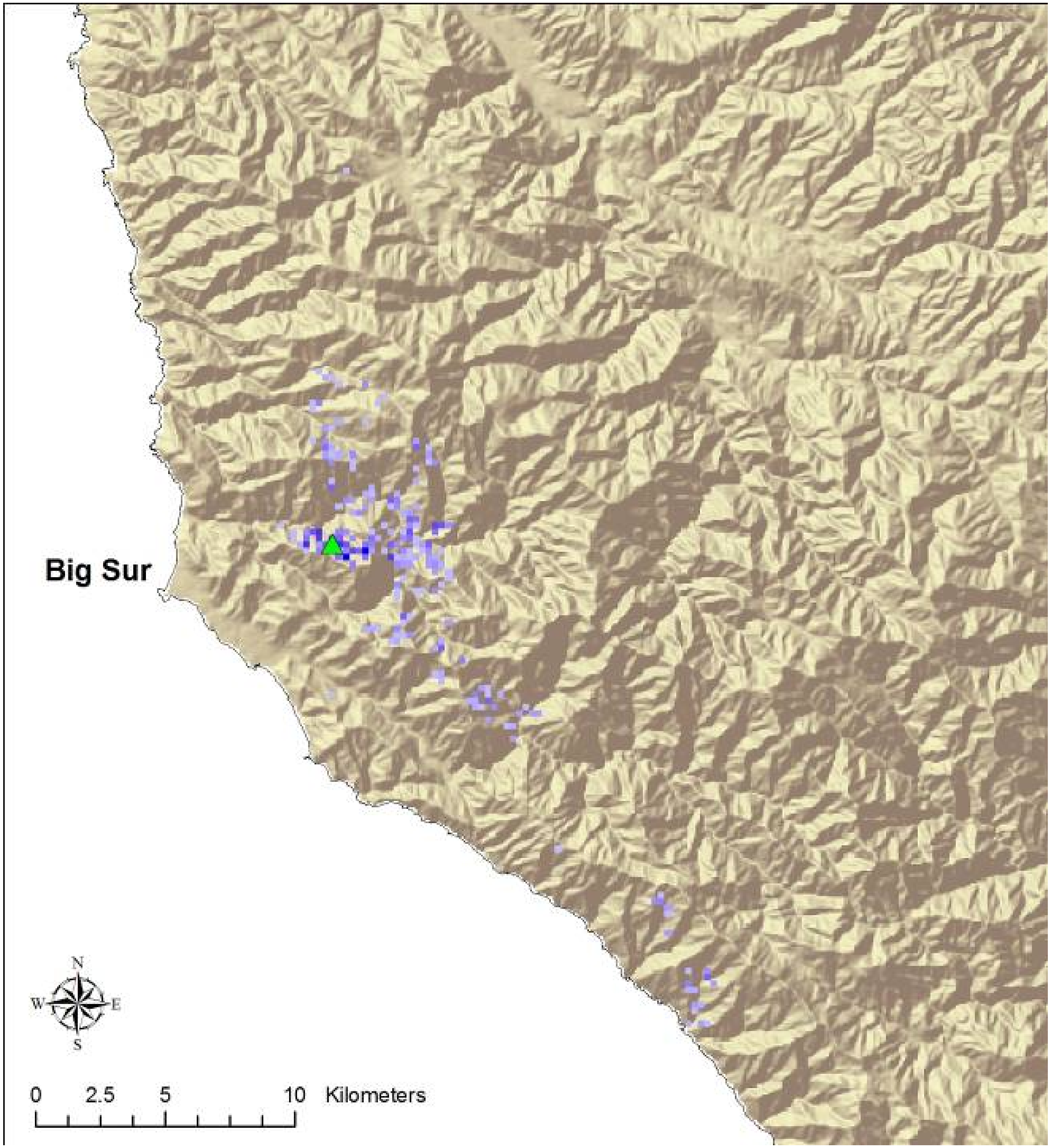
1988

1991

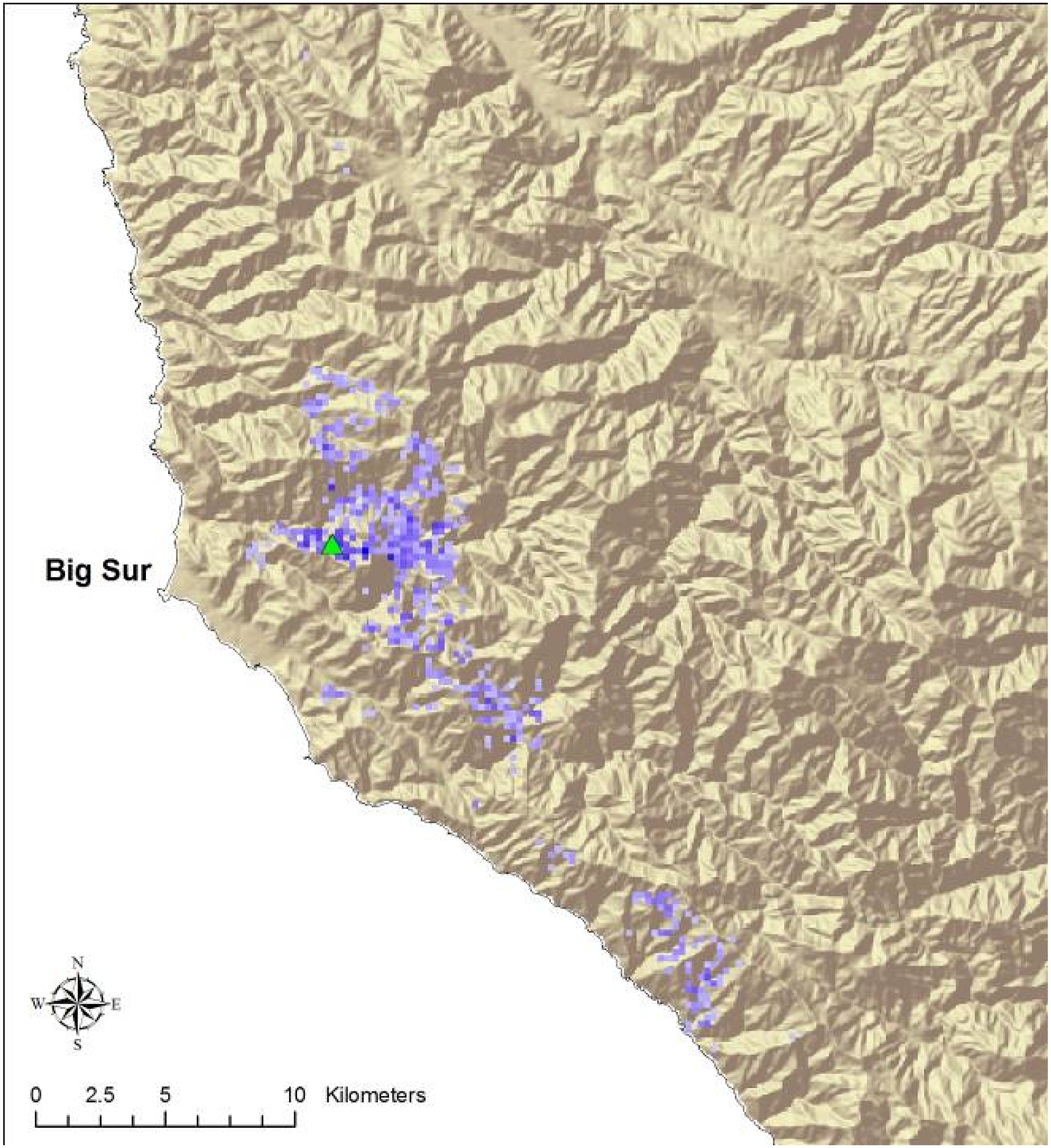




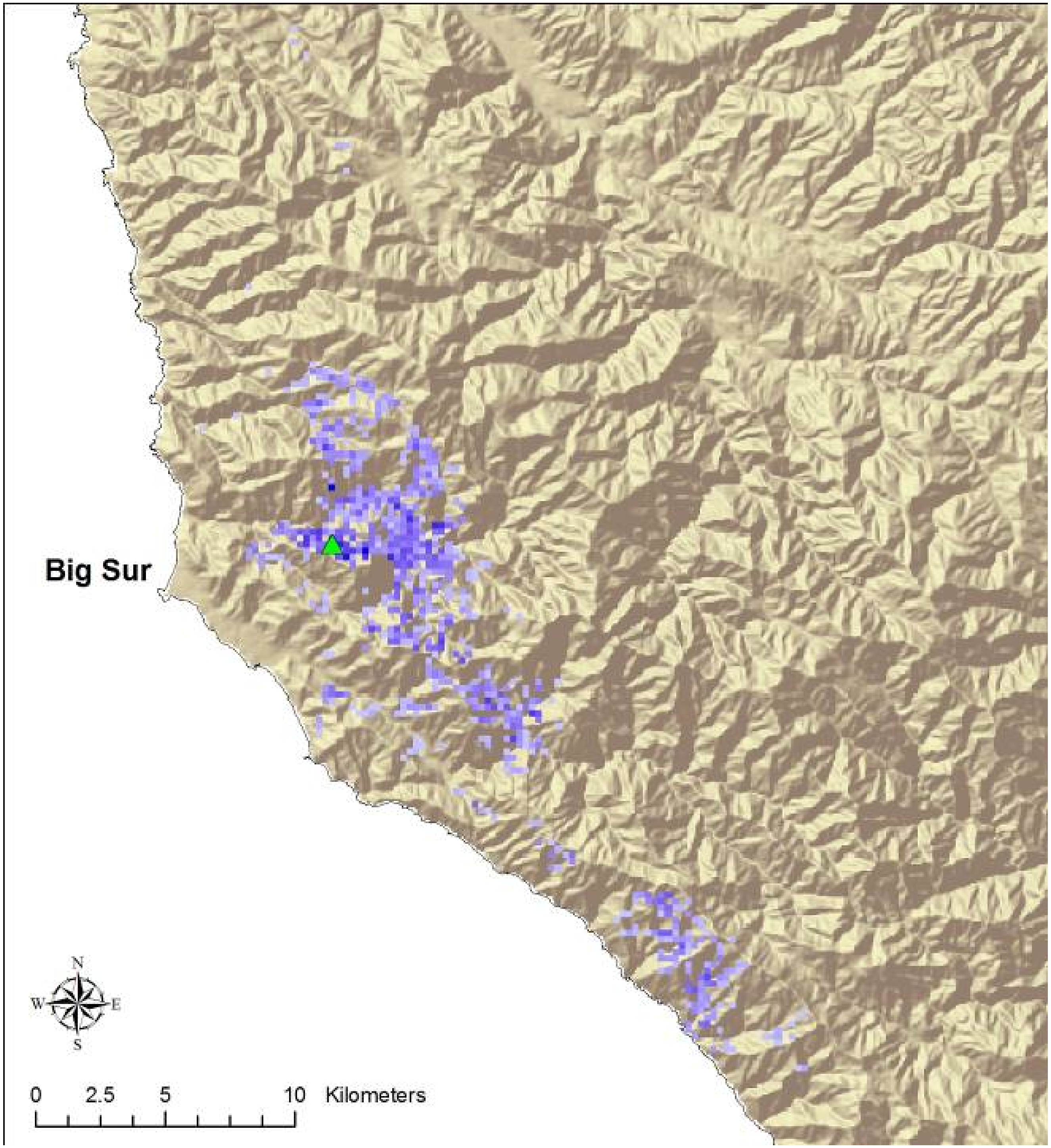
1994



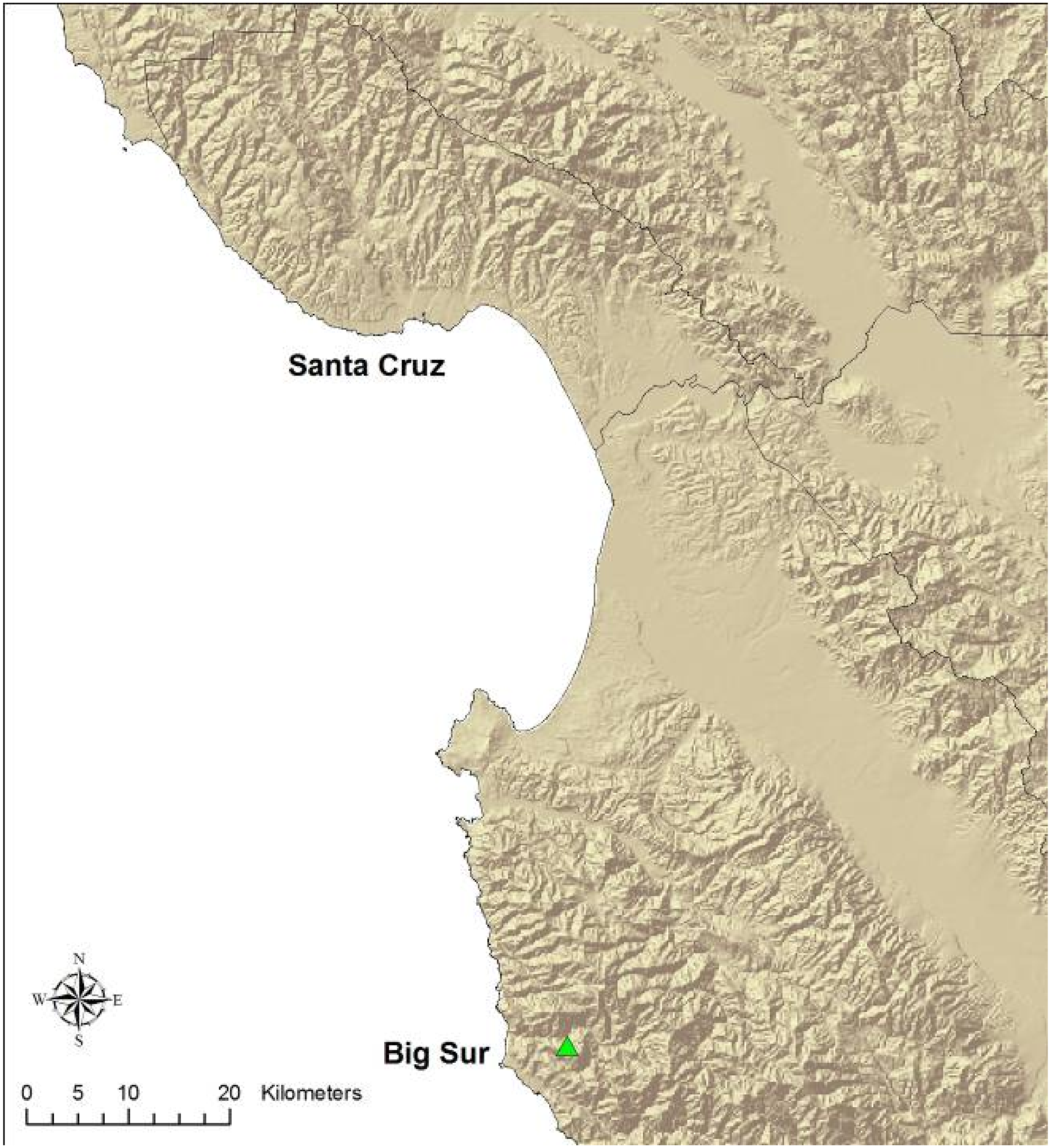
1997



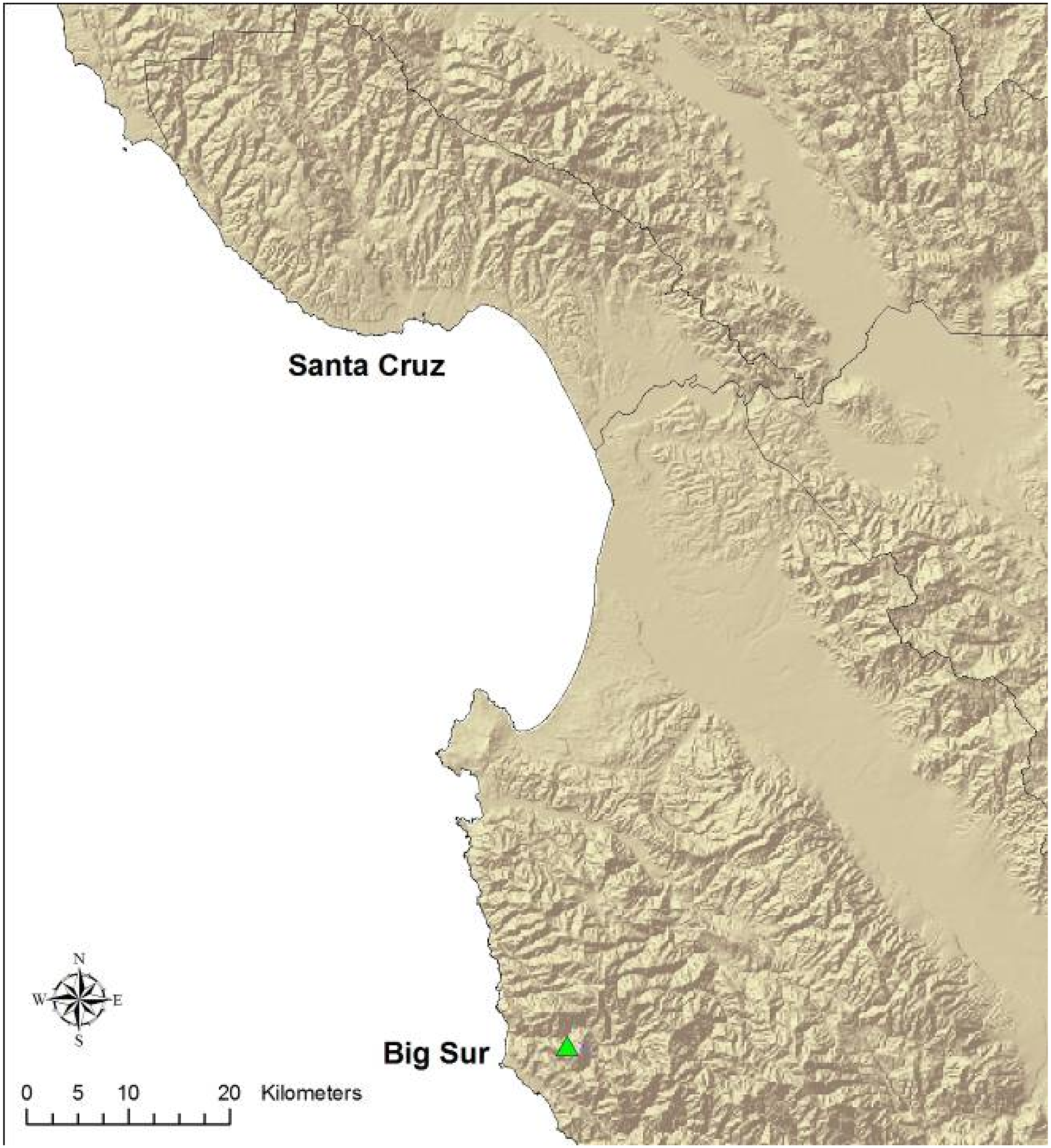
2000



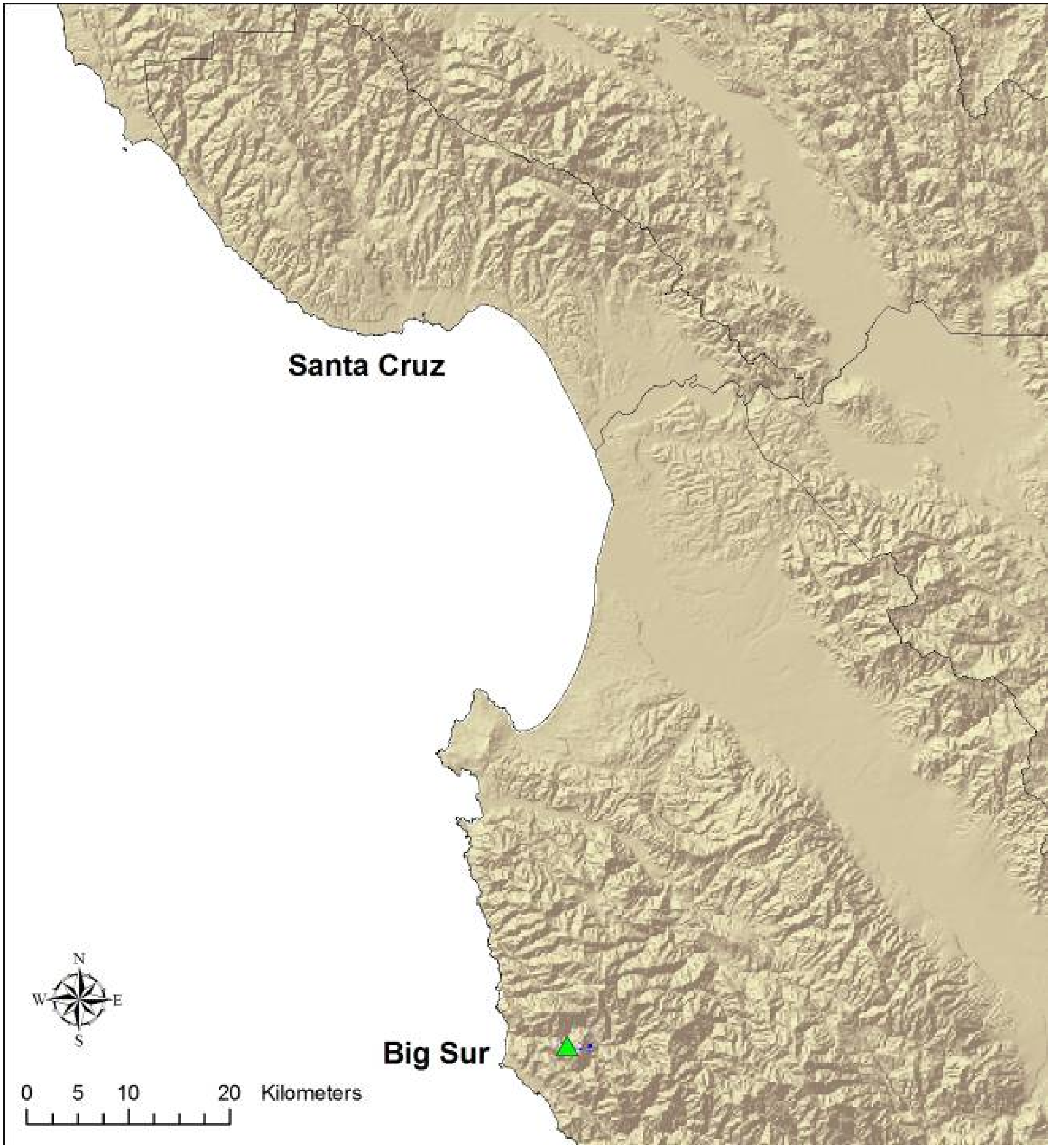
2003



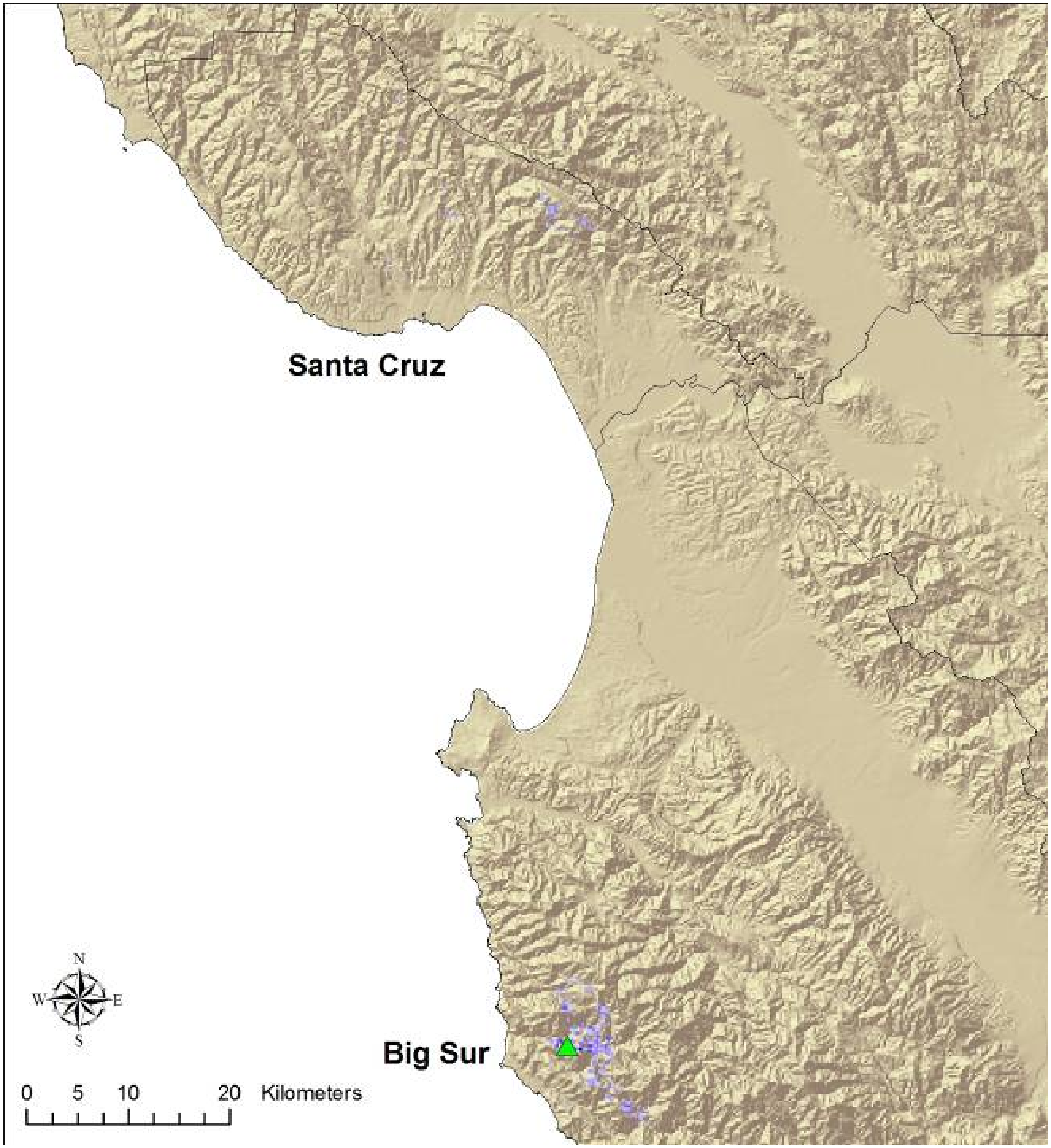
1985



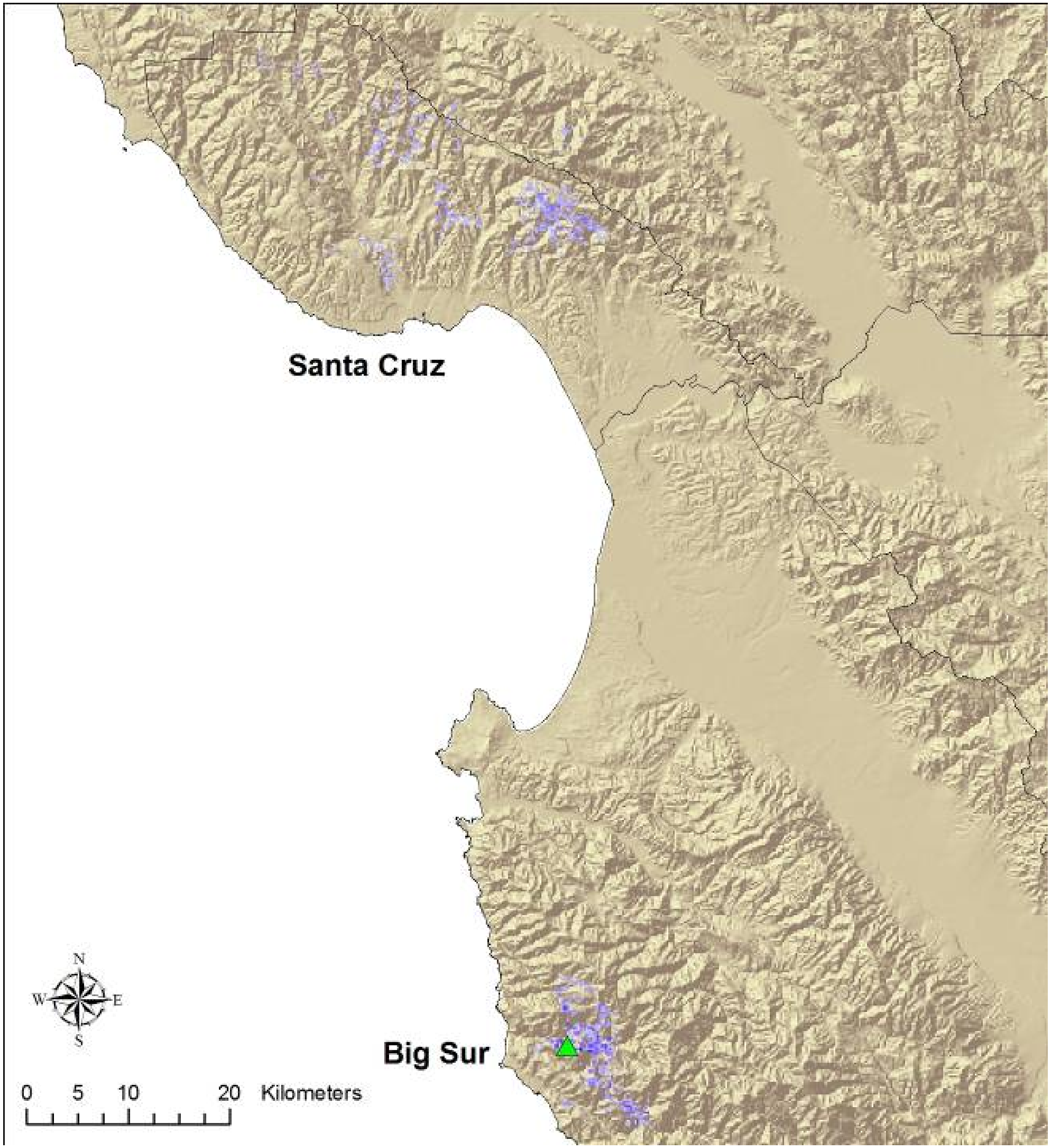
1988



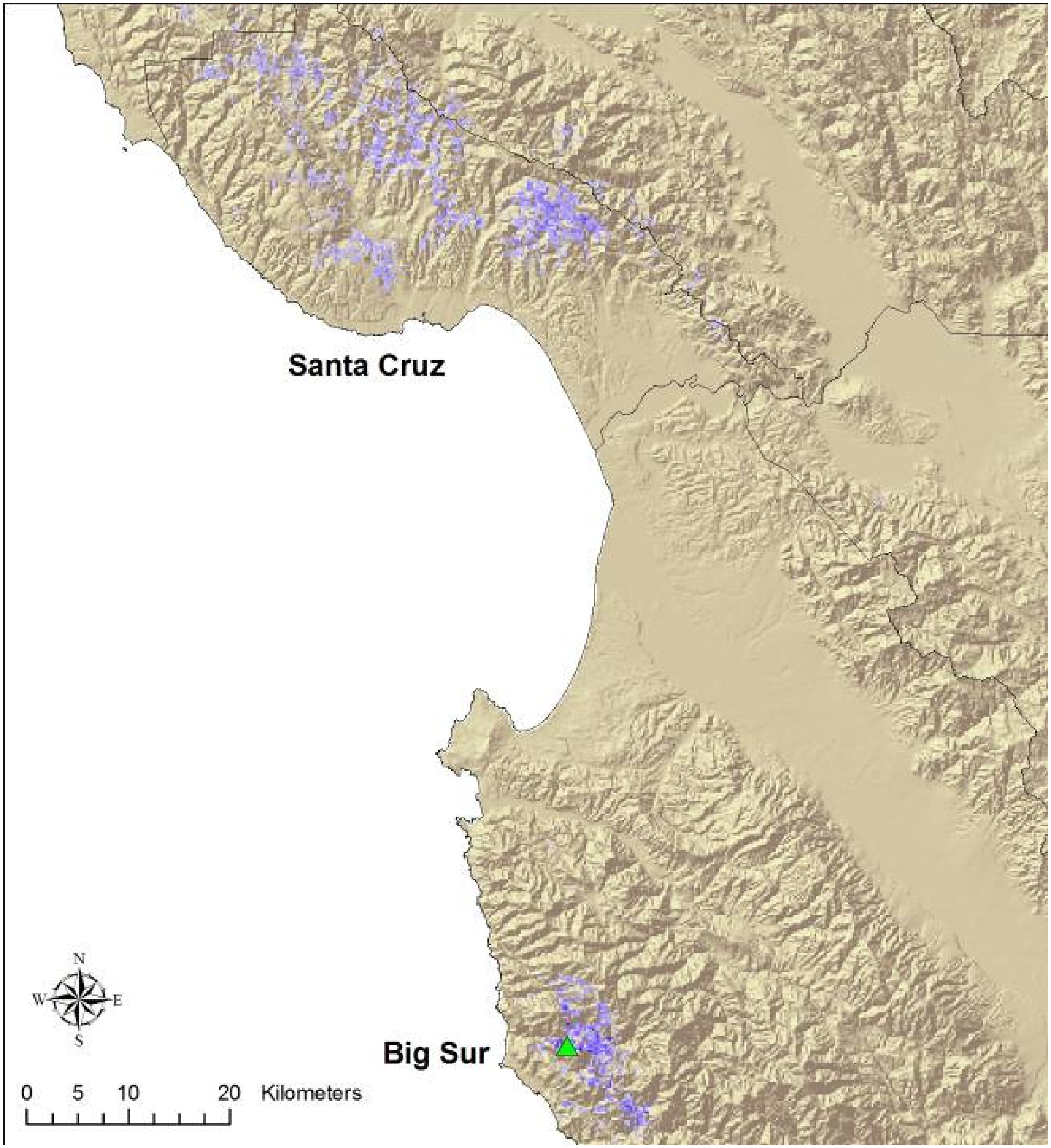
1991



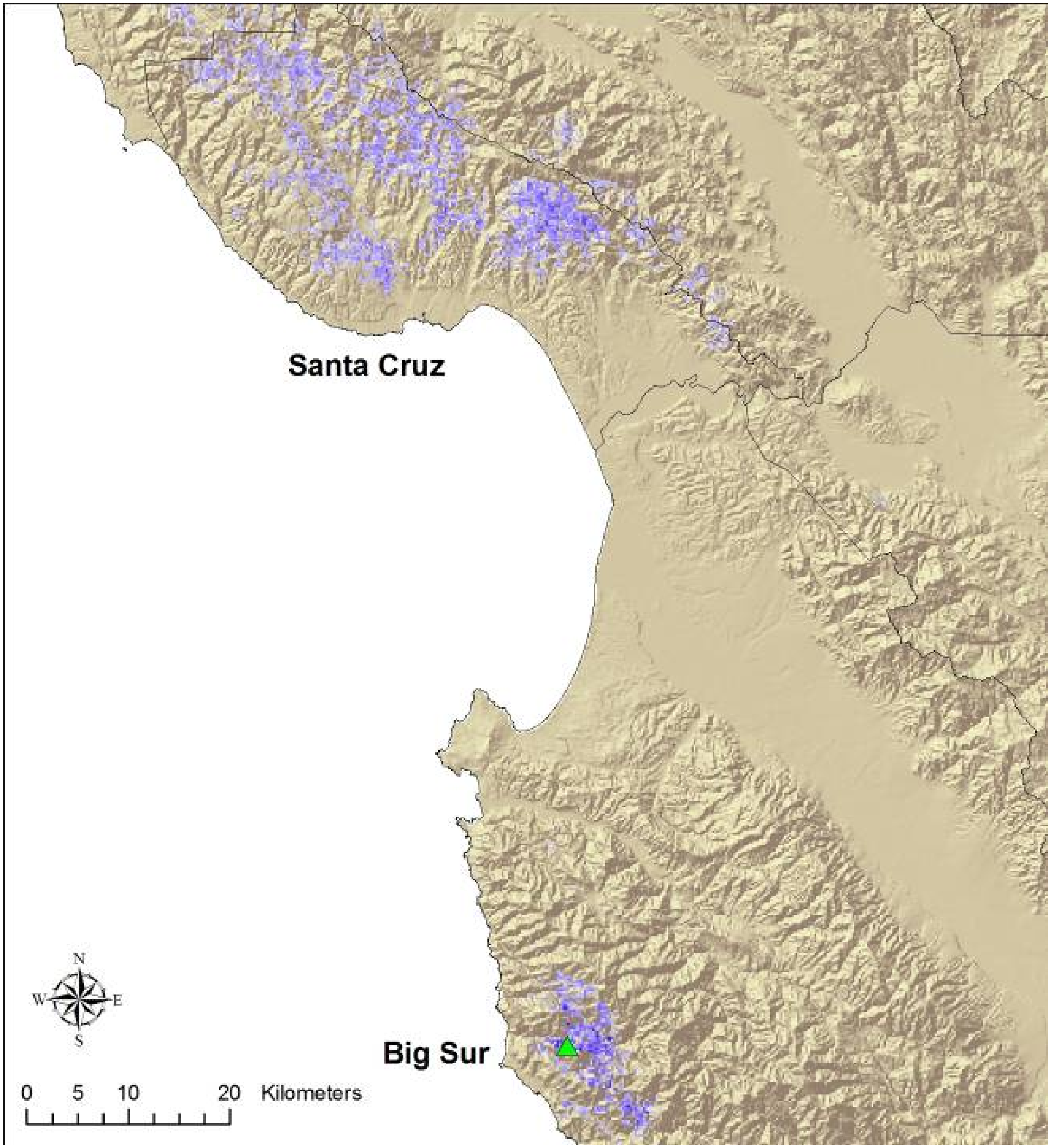
1994



1997



2000



2003

Model Evaluation

- Compare output maps with observed disease distribution
  - Run model 10 times
  - Assess predicted vs. observed infected area within 100 5k blocks
- Evaluate magnitude and direction of error and determine predictive accuracy
- Conduct sensitivity analysis of parameters

Additional research questions

- Test hypotheses about dispersal and infection processes
  - Influence of weather events
  - Simulate management scenarios

Significance

- Critical for predicting and slowing the spread of *P. ramorum* and other invasive species
  - Incorporate spatial and temporal heterogeneity of environmental conditions into spread modeling
  - Focus management efforts on high-risk landscapes
  - Test alternate management scenarios