Model-based control of spatio-temporal epidemics using latent processes

Gavin J Gibson

Maxwell Institute for Mathematical Sciences

Heriot-Watt University

JOINT WORK WITH: Hola Adrakey, Chris Gilligan, Nik Cunniffe (Cambridge), Tim Gottwald (USDA), George Stretaris (HWU)

ModStatSAP Meeting, Paris, 19th March 2018

LEADERS IN IDEAS AND SOLUTIONS

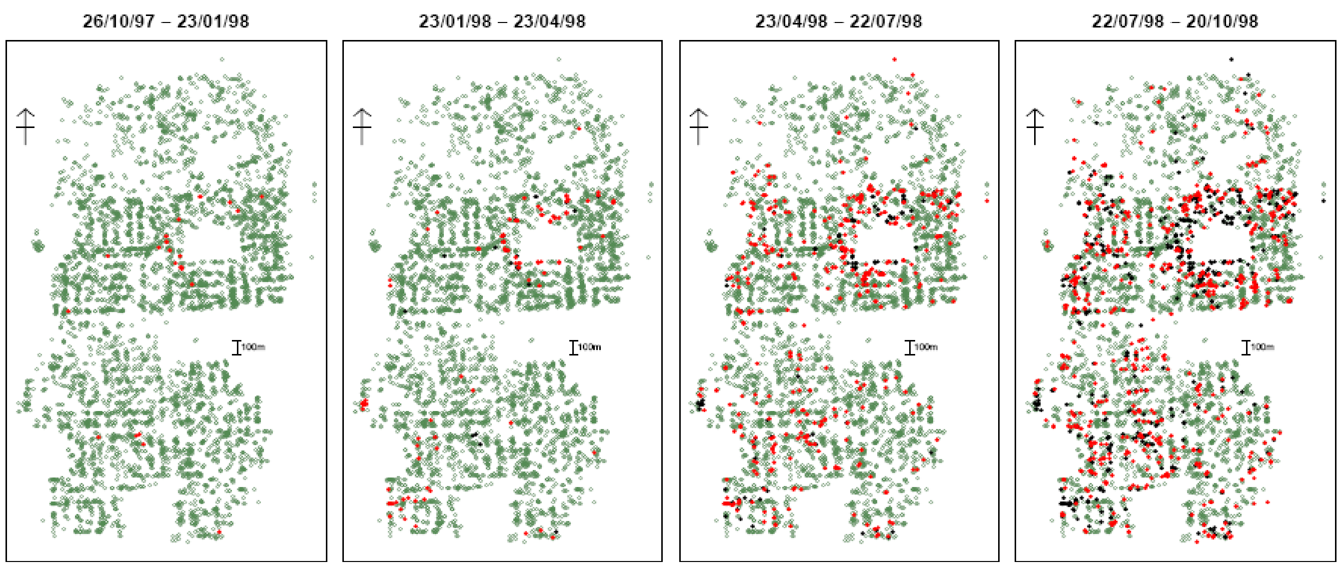
Summary

- Spatio-temporal stochastic models for informing control strategies
- Formulation of posterior measures for guiding control strategy
- Use of functional-model representations (non-centered parameterisations) for efficient comparison
- Conclusions – where to look for infection to maximise impact of control?

Generic problem

Observation of emerging epidemic

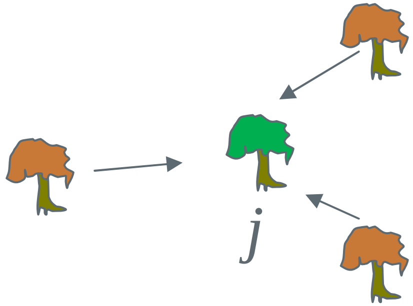
How should subsequent survey/control be designed in order to achieve a desired goal given available resources?



***Citrus canker epidemic**: Dade County, Miami, Florida*

SEIR spatio-temporal model

$S \rightarrow E$: If j is in state S at time t, then



$\Pr(j \text{ exposed } (t, t+dt)) = (\varepsilon + \beta \Sigma_i K(d_{ij}, \alpha))dt + o(dt)$

$E \rightarrow I$: $T_E^j \sim \pi_{\theta_E}^E$ (random sojourn time in E)

$I \rightarrow R$: $T_I^j \sim \pi_{\theta_I}^I$ (random sojourn time in I)

Parameters: $\theta = (\varepsilon, \beta, \kappa, \theta_E, \theta_I)$

Here we focus on simpler SI model with cryptic infections – infections only become symptomatic after fixed (known) period Δ (c.f. Neri et al (2014)).

Eradication strategy – ring culling

- Attempts to control Miami urban epidemic used a 1900ft eradication radius
- Model-based predictions of effective radius strongly dependent on choice of spatial kernel

Model fitting in Bayesian framework

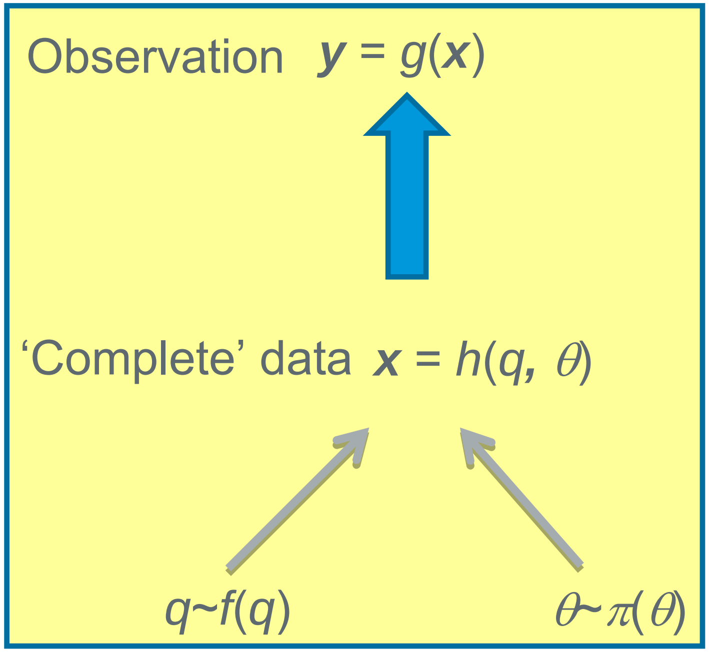
- For 'complete' data $\boldsymbol{x}(T)$ (e.g. times and nature of all transitions up to time T) $\pi(\boldsymbol{x} \mid \theta)$ tractable
- Given censored/filtered/noisy data $\boldsymbol{y}$, $\pi(\boldsymbol{y} \mid \theta)$ typically intractable
- Use data augmentation and sample from

  $\pi(\boldsymbol{\theta}, \boldsymbol{x}(T) \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\, \pi(\boldsymbol{x}(T), \boldsymbol{y} \mid \theta)$

  using e.g. MCMC
- Updating $\boldsymbol{x}$ often requires reversible-jump techniques given variable dimension

(See e.g. GJG, 1997, O'Neill & Roberts, 1999, Streftaris & GJG, 2004, Forrester *et al*., 2007, GJG *et al*., 2006, Chis-Ster *et al.* 2008, Starr *et al.* 2009, Jewell & Roberts, 2007, Neri *et al*., 2014, Lau *et al*., 2015)

Functional-model representations



Functional models (Dawid & Stone, 1983)

Consider outcome as deterministic function $h(q, \theta)$ where q has known distribution independent of θ.

In model choice q can be used as a latent residual process.

Investigating $\pi(\theta, q \mid \boldsymbol{y})$ rather than $\pi(\theta, \boldsymbol{x} \mid \boldsymbol{y})$ facilitates model assessment via latent classical tests.

Here we extend the idea to formulate models for epidemic dynamics in the presence of control $\boldsymbol{d}$, so that $\boldsymbol{x} = h^*(q, \theta, \boldsymbol{d})$.

Sellke Construction (Sellke, 1983)

Assigns threshold q_i to each individual. If $R_i(t)$ denotes infectious challenge to i at time t, infection time x_i occurs when integrated challenge reaches threshold

$$q_i = \int_0^{x_i} R_i(t)\,dt \sim \mathrm{Exp}(1)$$

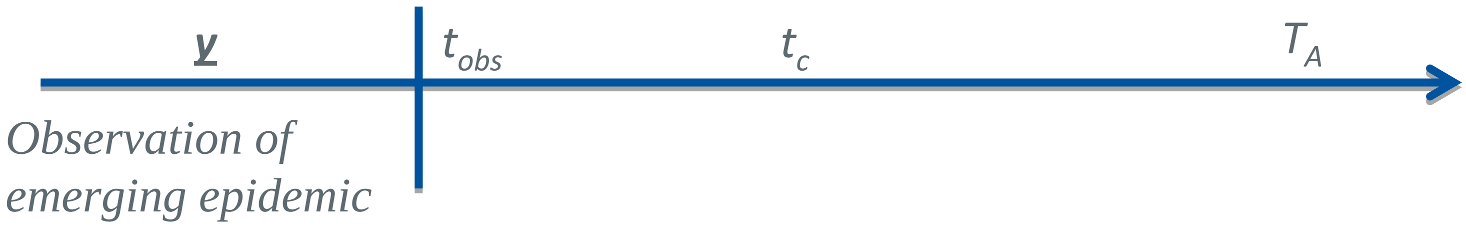
- Epidemic dynamics specified (for SI with cryptic) by $\underline{q}$ (vector of Sellke Thresholds) and θ, i.e. ($\boldsymbol{x} = h(\underline{q}, \theta)$).

- For controls $\boldsymbol{d}$, <u>based on removal of infected individuals</u>, it follows that $\boldsymbol{x} = h^*(\underline{q}, \theta, \boldsymbol{d})$.

- Gives a means of <u>coupling</u> epidemic trajectories <u>under different control strategies</u>.

Control strategies

- Based on removal of hosts found to be infected at control time t_c (by perfect diagnostic test if not obviously infected).
- N' hosts can be targeted (resource constraint)
- Impact assessed at time T_A *e.g. via* number of infections occurring by T_A.



Which hosts *j* to target?

Based on $E(G_M(\boldsymbol{x}(t), j) \mid \boldsymbol{y})$ at some time $t = t_M \geq t_{obs}$

Measure calculated on host

Host index

Epidemic trajectory up to t.

Candidate measures - (x_j denotes infection time of j)

$G_R(\boldsymbol{x}(t), j) = I_{\{x_j < t\}}$ - 'Risk'

$G_H(\boldsymbol{x}(t), j) = \Sigma_{x_i > t,\, i \neq j}\, \beta K(d_{ij}, \alpha)$ - 'Hazard'

$G_T(\boldsymbol{x}(t), j) = G_R(\boldsymbol{x}(t), j) \times G_H(\boldsymbol{x}(t), j)$ - 'Threat'

Estimating expected reduction

- Use random sample from $\pi(\theta, \boldsymbol{x}(t) \mid \boldsymbol{y})$ to generate sample of size m from $\pi(\theta, \boldsymbol{q} \mid \boldsymbol{y})$.

- Let $u(x(T))$ denote the number of infections by time T for trajectory $x(T)$. Let $\boldsymbol{d}$ denote control strategy.

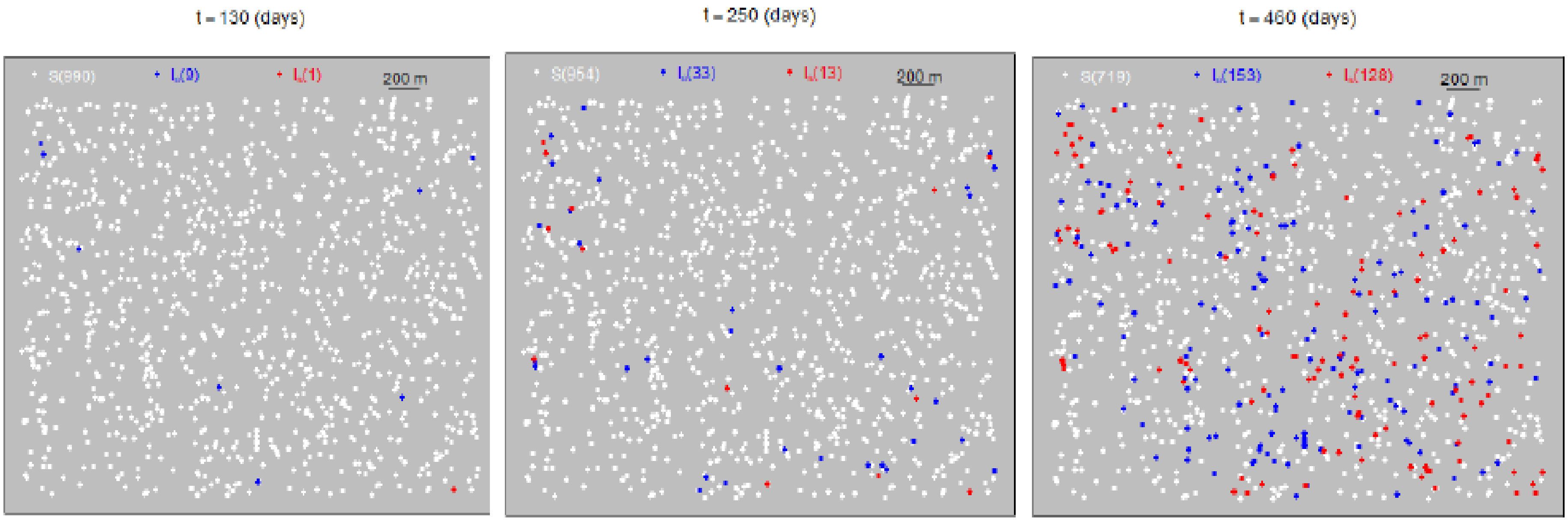
$$x(T) = h(\theta, \boldsymbol{q}), \qquad x_d(T) = h^*(\theta, \boldsymbol{q}, \boldsymbol{d})$$

$$\text{EER}(d) = -\Sigma_i \{u(h^*(\theta_i, \boldsymbol{q}_i, d)) - u(h(\theta_i, \boldsymbol{q}_i))\}/m$$

- Here we take $m = 1000$ draws (θ_i, q_i) from $\pi(\theta, \boldsymbol{q} \mid \boldsymbol{y})$, using these as a test-bed of 'pre-epidemics' on which to compare controls.

Simulation: non-clustered

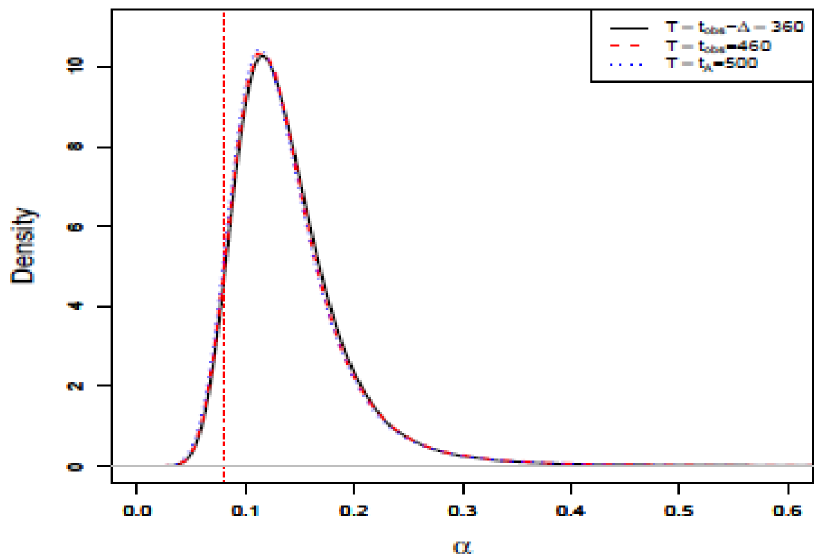
Simulated epidemic: Uniformly distributed population, primary + exponential kernel



Parameters Δ = 100, α = 0.08, β = 7 x 10^{-6}, ε = 5 x 10^{-5}

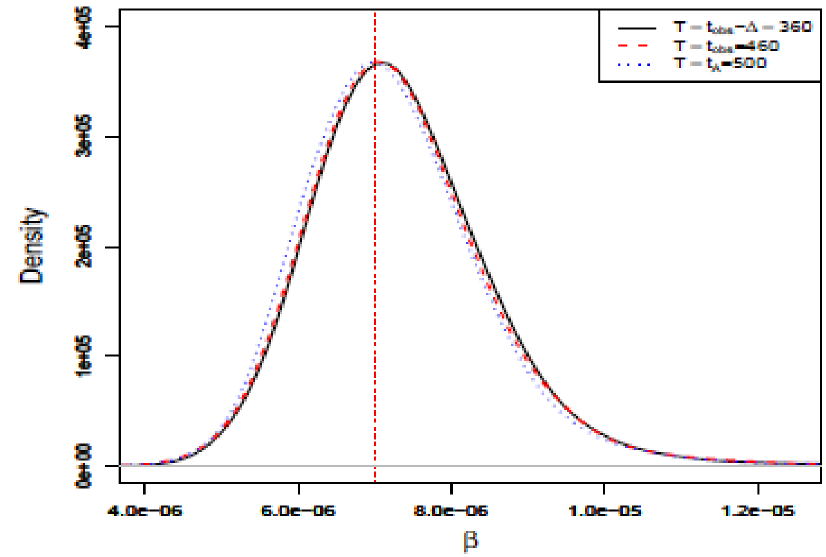
Estimated by Neri *et al* (2014)

Consider application of control applied at t = 460, with impact assessed at t = 500.



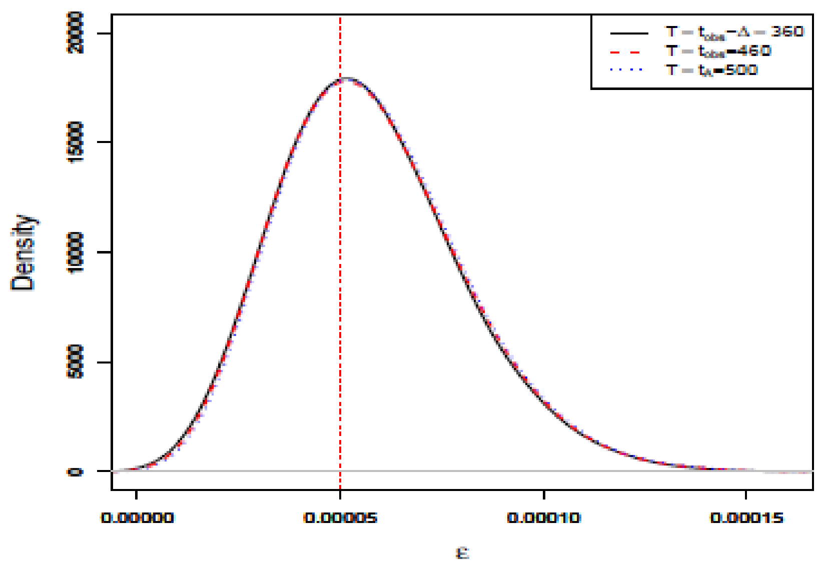
T = tobs−Δ = 360
T = tobs=460
T = tA=500
Density
0
2
4
6
8
10
0.0
0.1
0.2
0.3
0.4
0.5
0.6
α

(a)



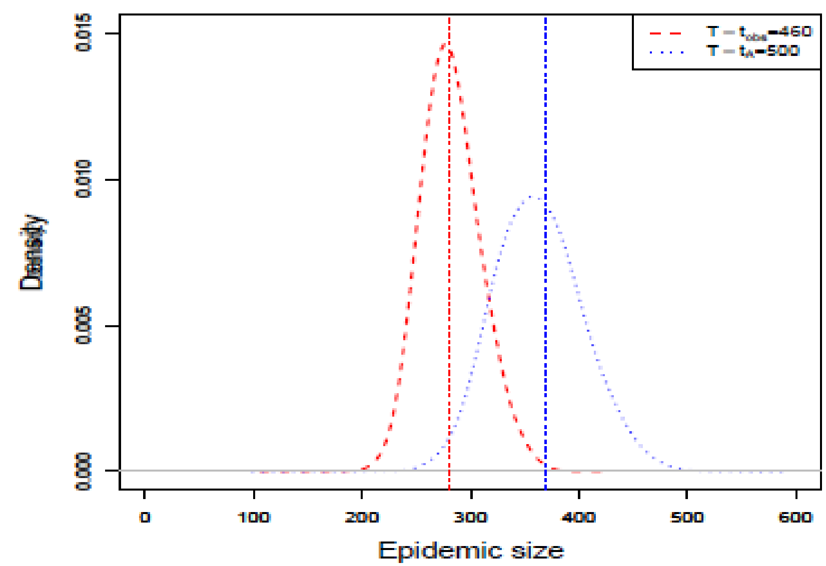
T = tobs−Δ = 360
T = tobs=460
T = tA=500
Density
0e+00
1e+05
2e+05
3e+05
4e+05
4.0e−06
6.0e−06
8.0e−06
1.0e−05
1.2e−05
β

(b)



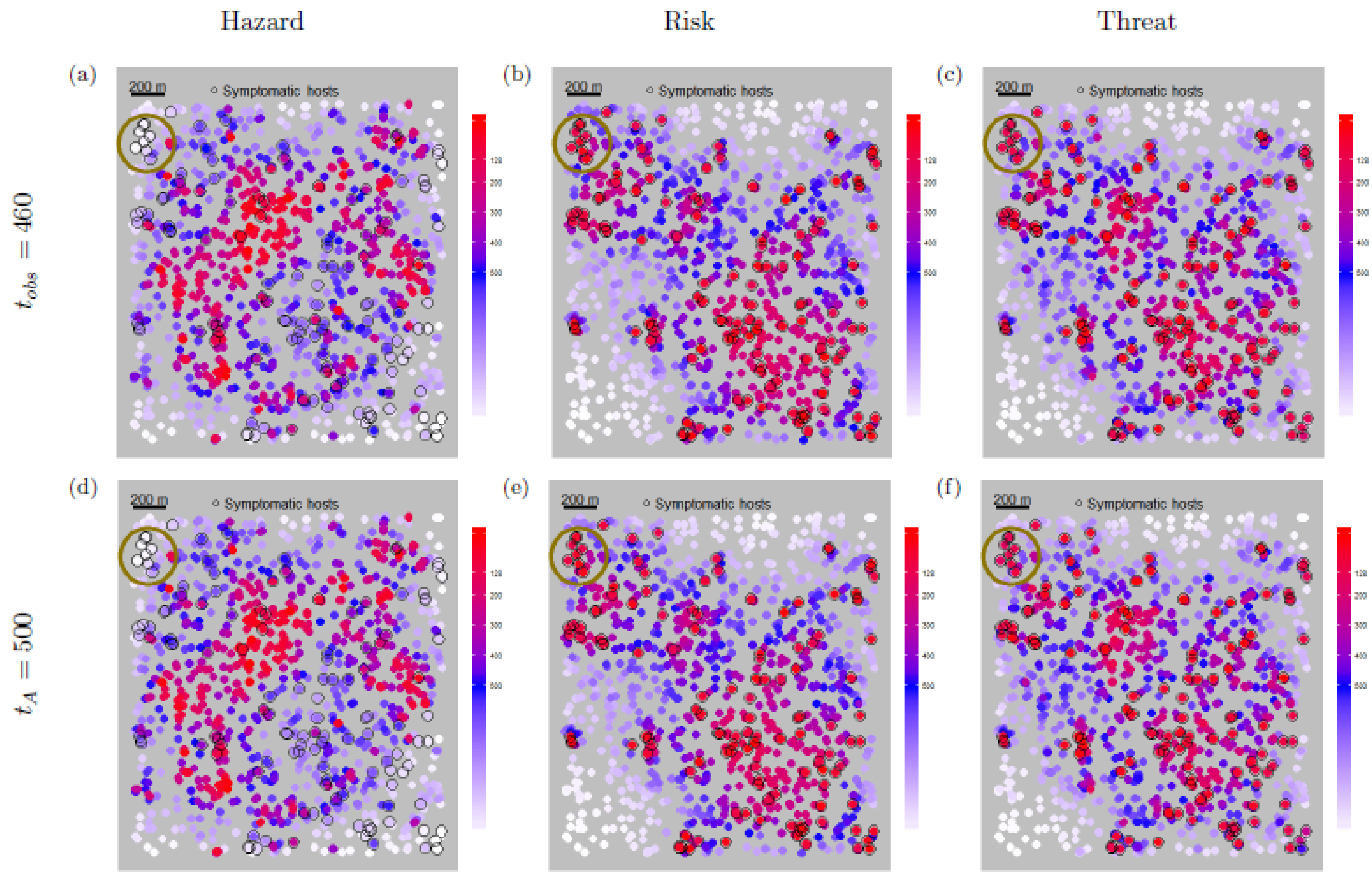
T = tobs−Δ = 360
T = tobs=460
T = tA=500
Density
0
5000
10000
15000
20000
0.00000
0.00005
0.00010
0.00015
ε

(c)

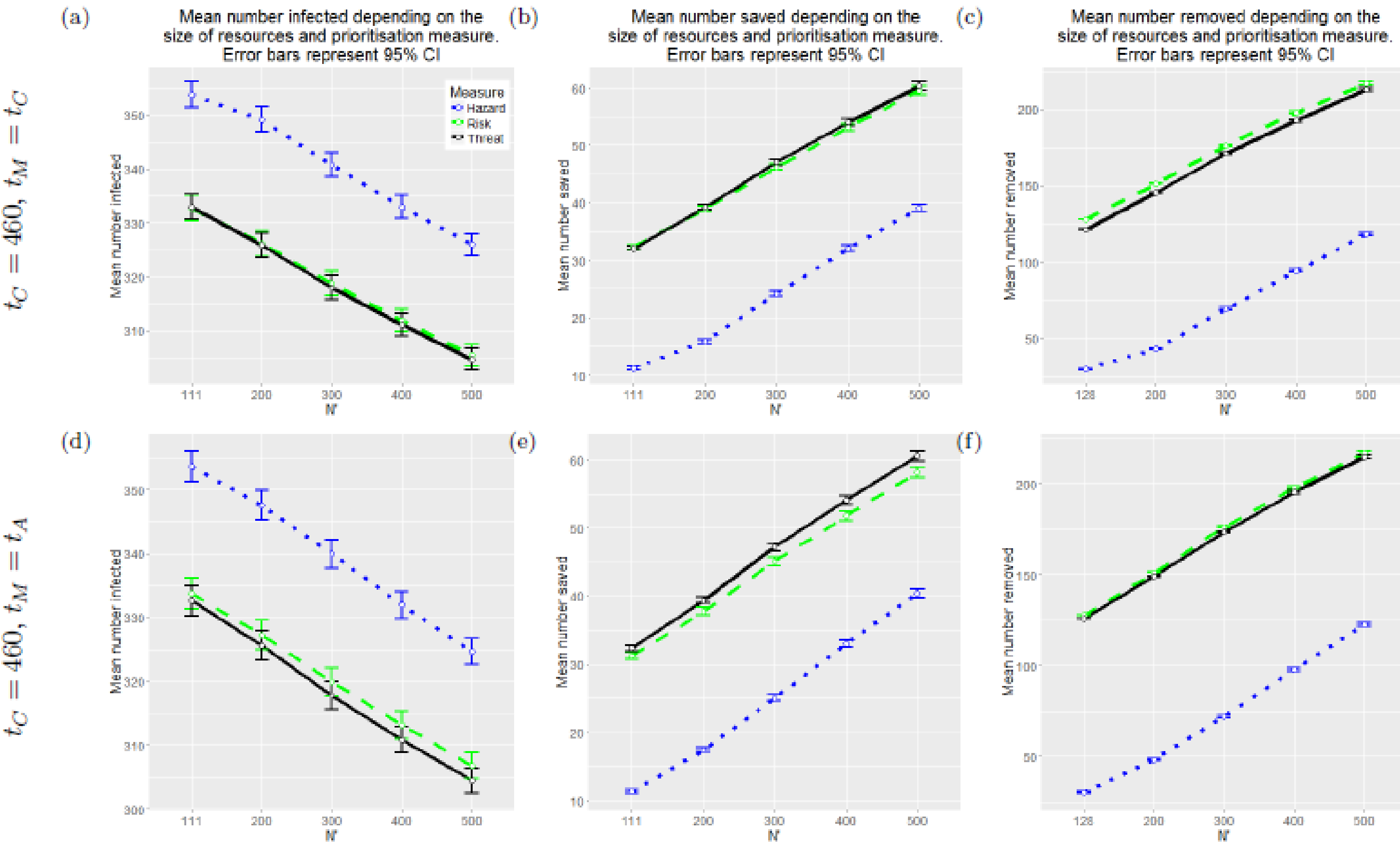


T = tobs=460
T = tA=500
Density
0.000
0.005
0.010
0.015
0
100
200
300
400
500
600
Epidemic size

(d)



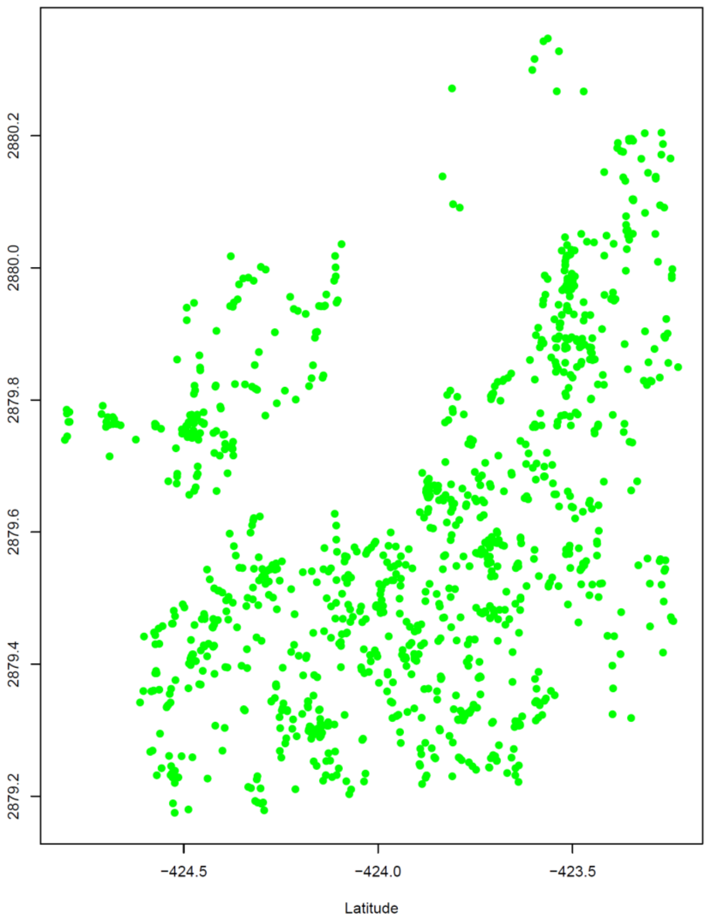
Hazard, Risk and Threat maps



Threat map marginally more effective that risk map regardless of when measures are estimated.

Clustered host populations

Citrus location in Broward county (Florida)



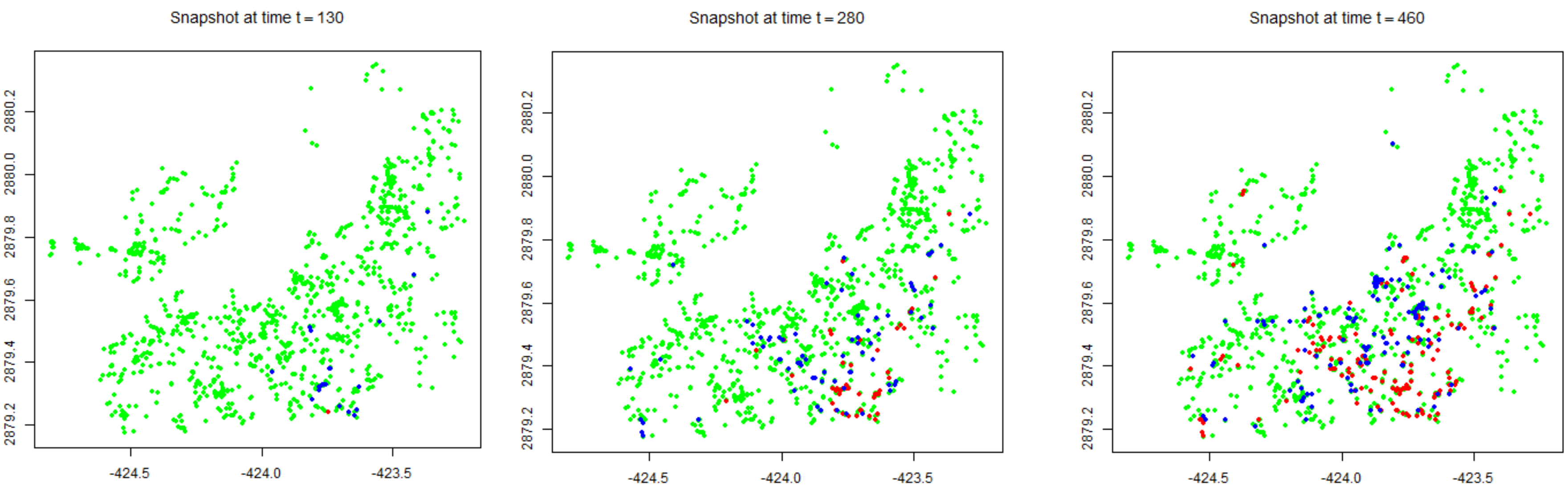
- Citrus locations from Dade county
- 1111 trees spatially distributed
- Citrus canker epidemic on this population analysed by Neri et al (2014)
- Canker typically controlled using ring-culling strategies (not yet considered in this framework but amenable to it)
- Simulate epidemics of 2 types:

  - exponential kernel with primary

  - exponential kernel no primary

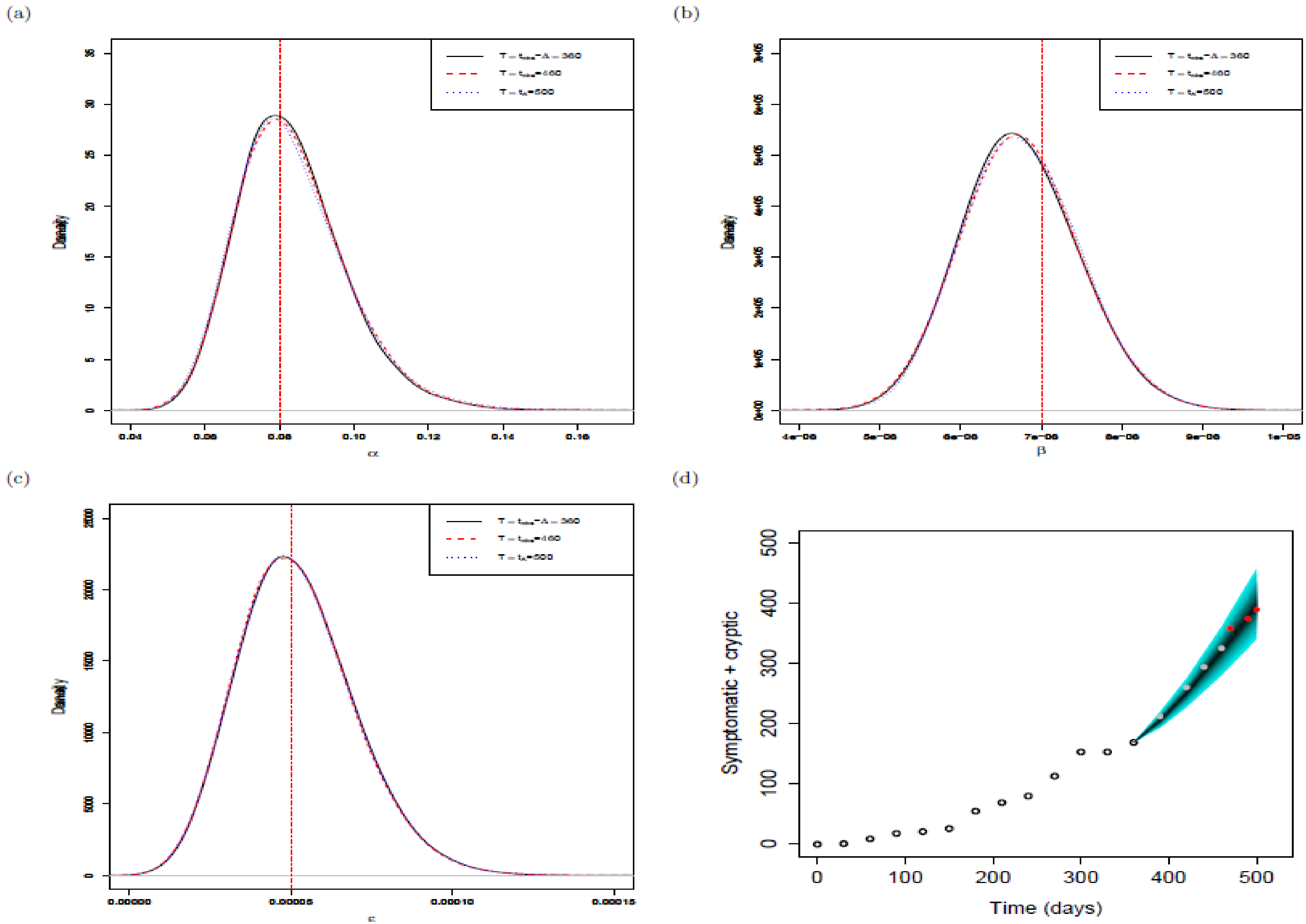
Case	α	β	ϵ	t_{obs}	Infections observed	Cryptic	T
(I)	0.08	7.10^{-6}	0.00005	460	169	133	500
(II)	0.08	8.10^{-6}	0	460	111	124	500

Simulated epidemic (with primary – Case I)



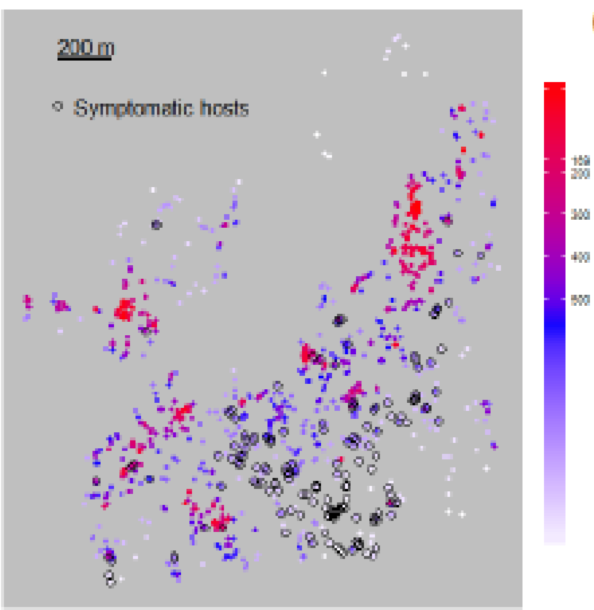
Parameters $\Delta = 100$, $\alpha = 0.08$, $\beta = 7 \times 10^{-6}$, $\varepsilon = 5 \times 10^{-5}$

Estimated by Neri *et al* (2014)

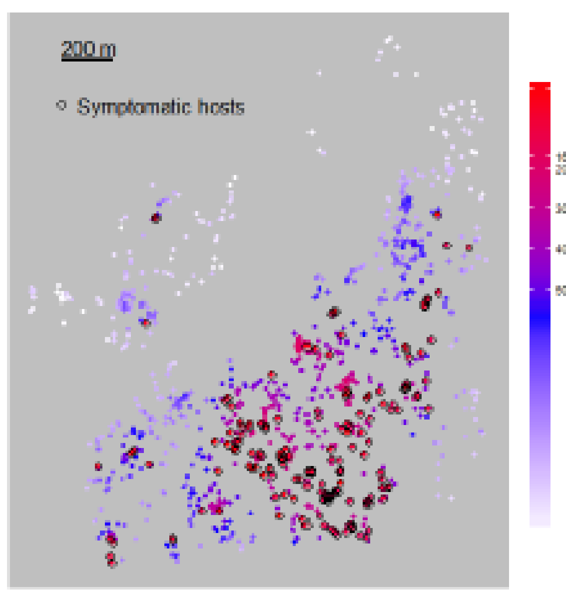

(a)
Density
α
(b)
Density
β
(c)
Density
ε
(d)
Symptomatic + cryptic
Time (days)
0
100
200
300
400
500

Hazard, risk and threat, $t_M = 460$

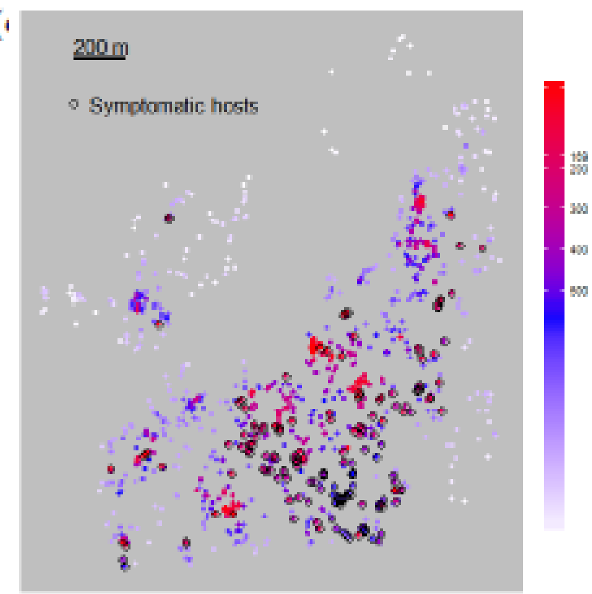
Hazard



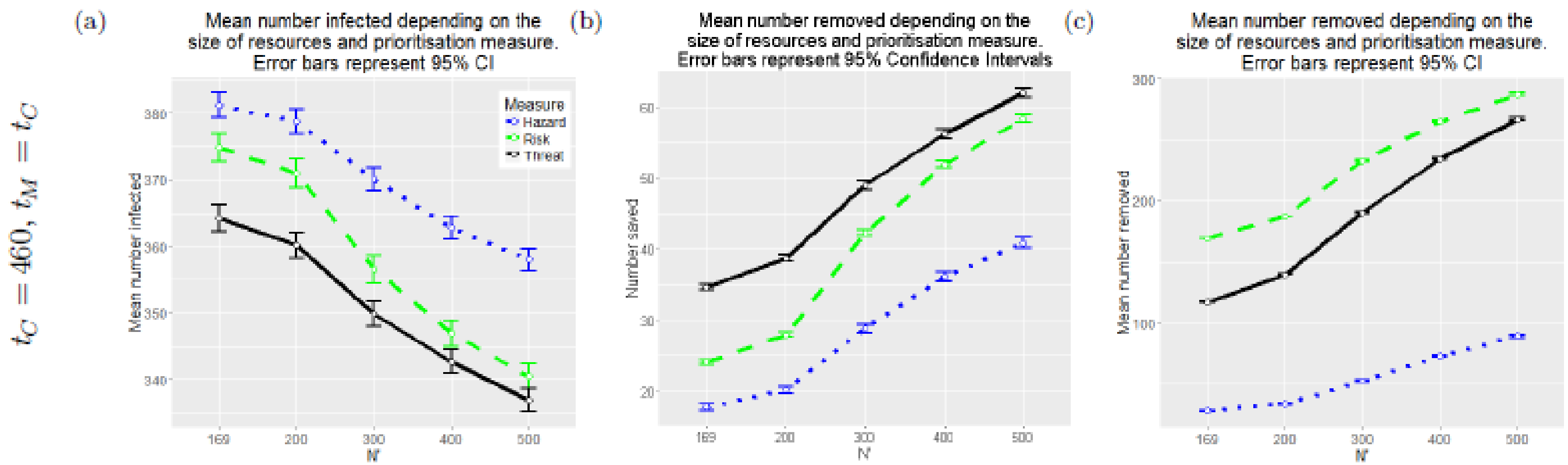
Risk



Threat

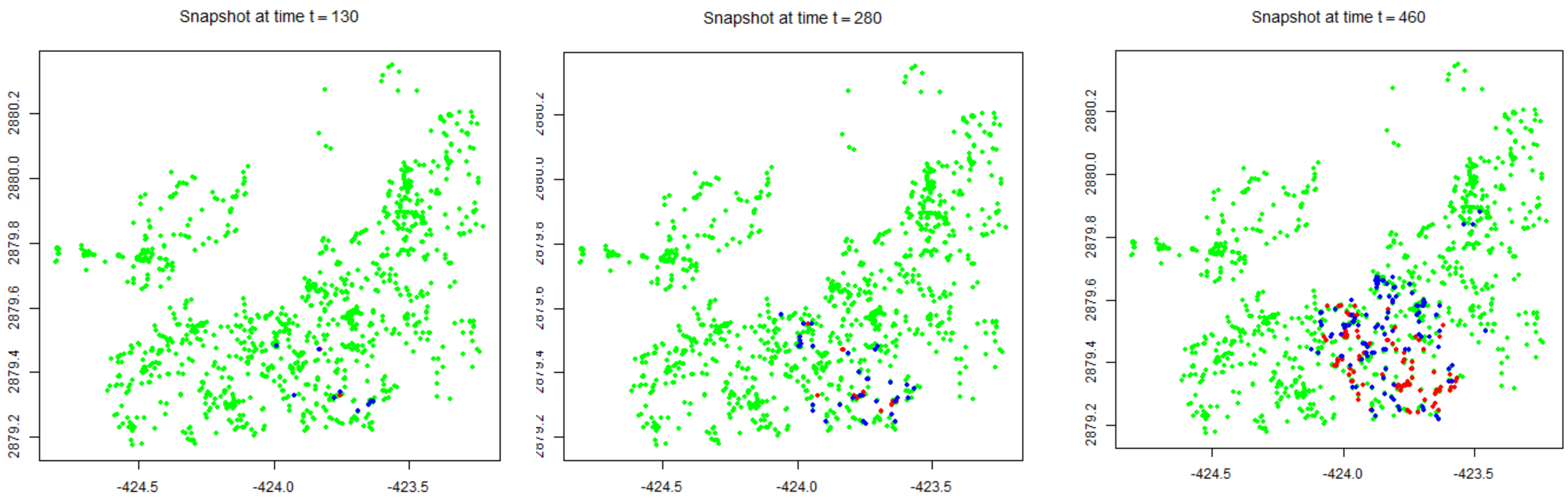


Case I: $t_M = t_C = 460$



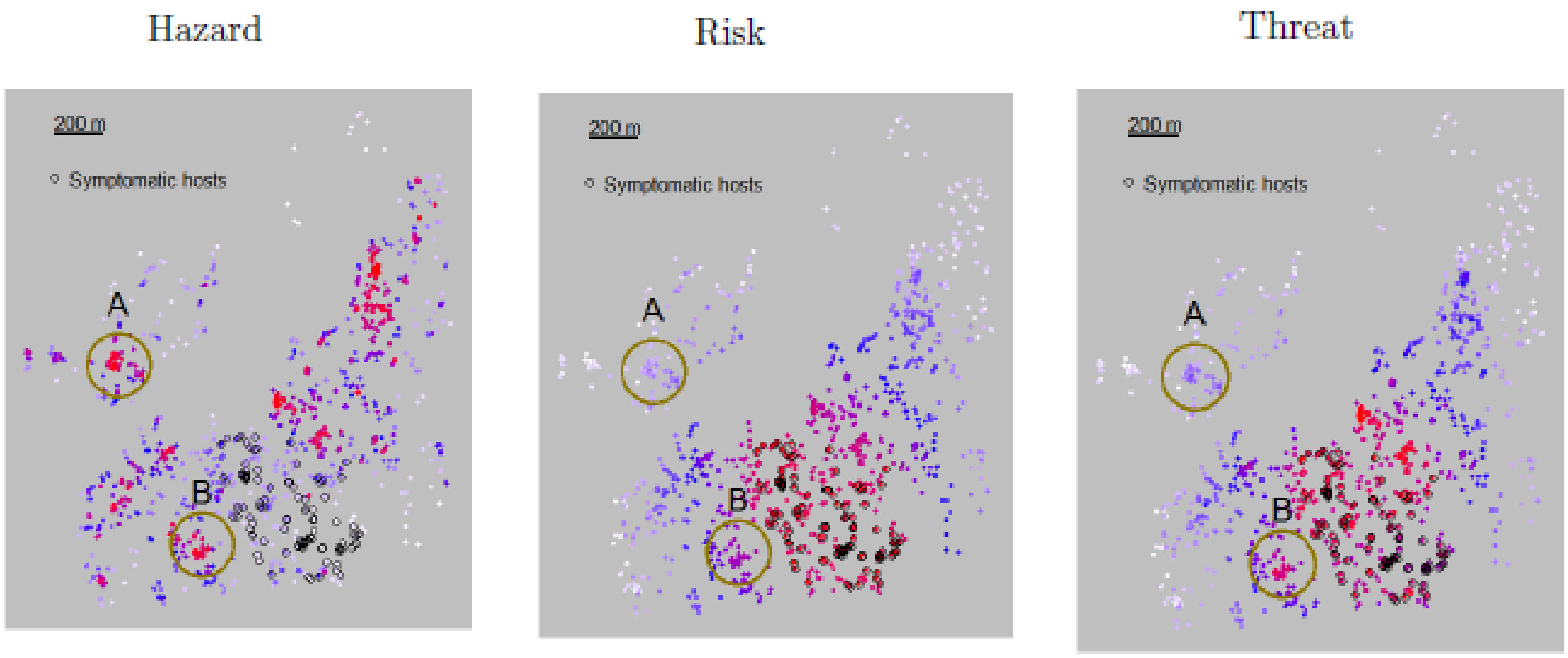
- Threat map gives largest expected reduction
- Effect largest for smaller N' – when resources are scarce
- Hazard map generally poorest performance

Simulated epidemic (no primary, Case 2)

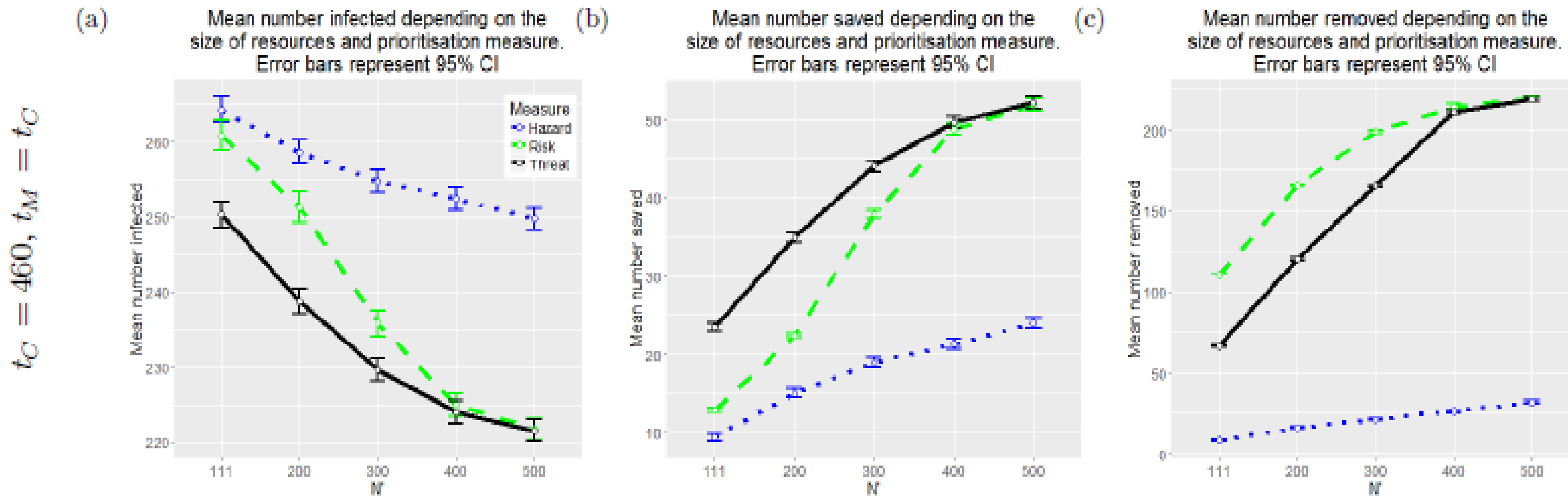


Parameters: $\alpha = 0.08$, $\beta = 8 \times 10^{-6}$

Risk, hazard and threat maps at $t_M = 460$

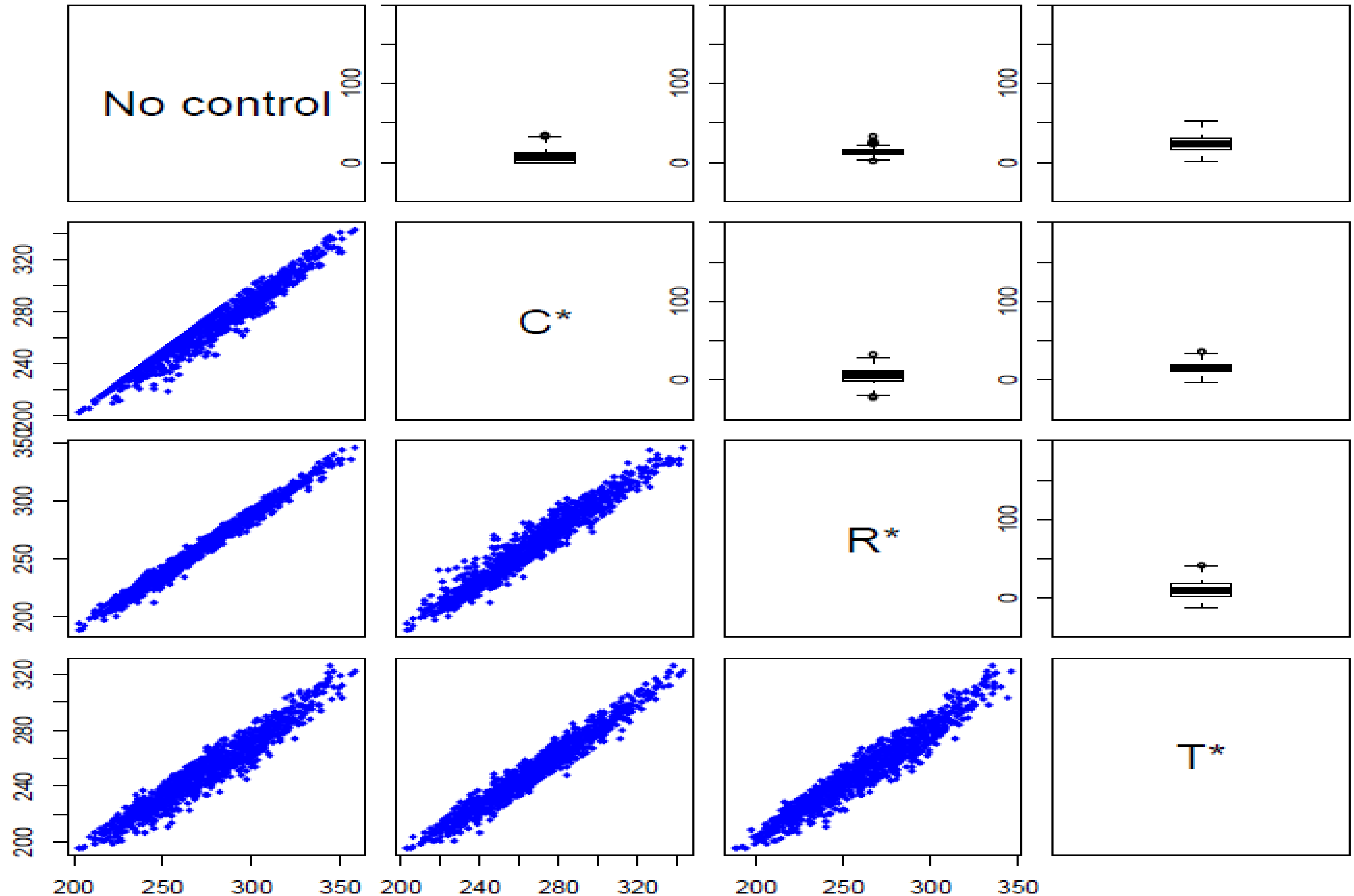


Case II: $t_M = t_C = 460$



- Threat map gives largest expected reduction
- Effect largest for smaller N' – when resources are scarce
- Reflects differences in hosts appearing at the 'top of the order'

Induced correlation under coupling, epidemic size for N’ = 111



Conclusions

- Bayesian framework provides flexible means to design controls using different prioritisation measures
- A strategy which ‘locates’ the most infections may not be optimal
- In structured populations strategies that prioritise searches using the threat measure may better identify potential ‘super-spreaders’(?)
- Coupling of epidemics leads to variance reduction, potentially removing need to embed simulation in optimisation routines
- Focus on small population of pre-epidemics makes approach inherently parallelisable

Next steps

- More complex constrained design spaces
- Applications to ‘ring-culling’ strategies
- Incorporation of uncertainty in diagnostic tests
- Generalisation of cost functions