Evidence-based controls for epidemics using spatio-temporal stochastic models in a Bayesian framework

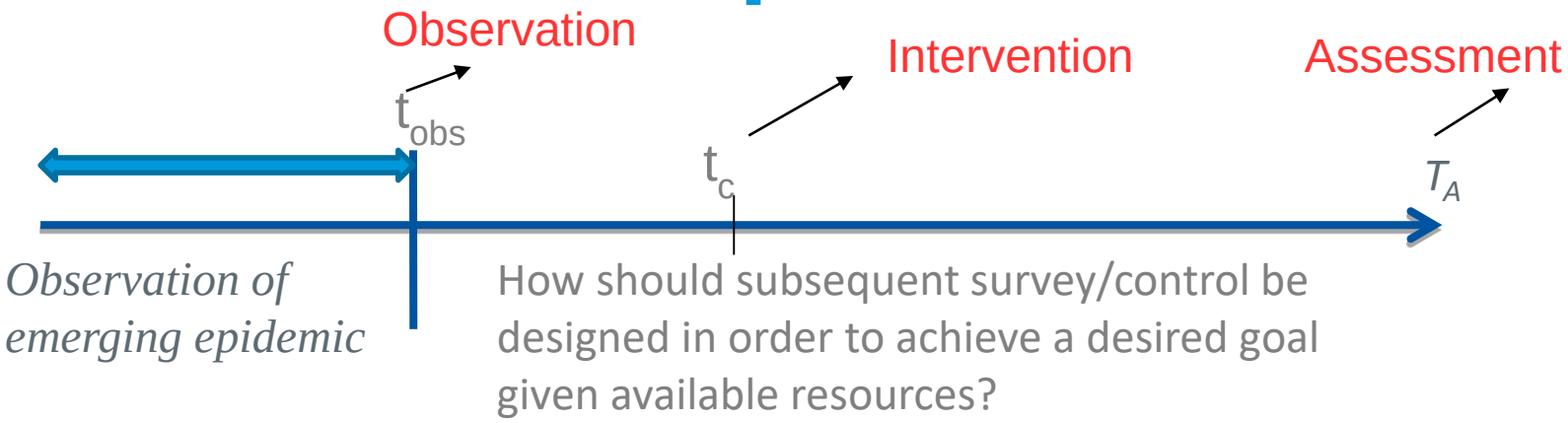
Hola Adrakey

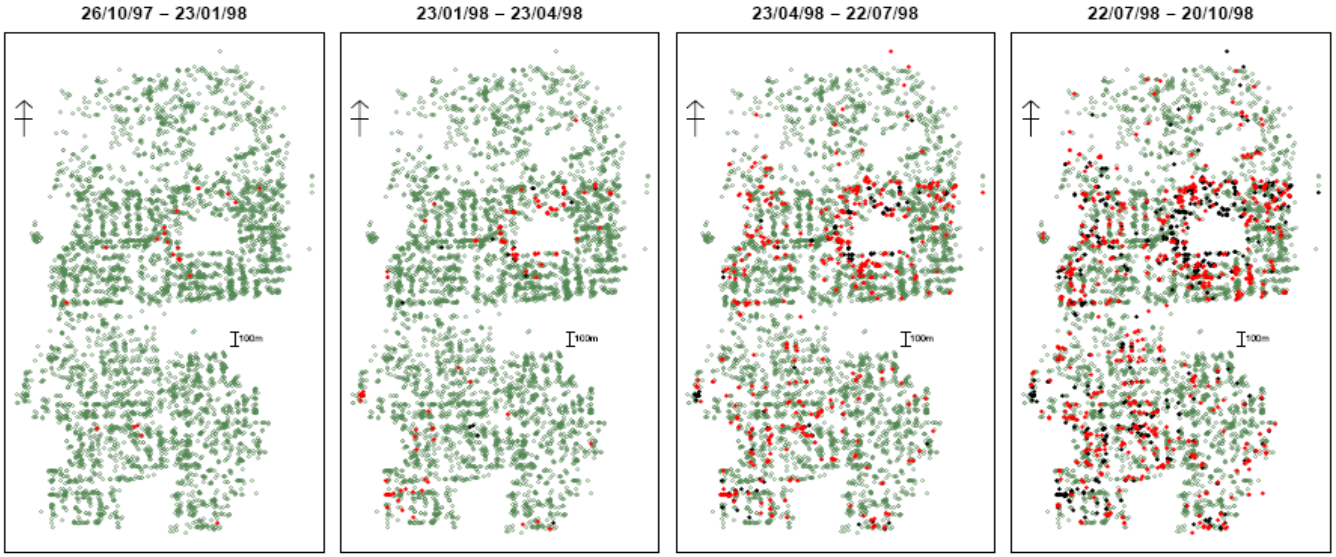
JOINT WORK WITH: George Streftaris, Nik Cuniffe, Tim Gottwald, Chris Gilligan And Gavin Gibson

Main messages

- 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?

Generic problem

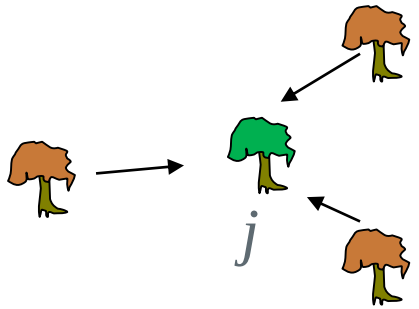




***Citrus canker epidemic**: Dade County, Miami, Florida*

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: $\theta = (\varepsilon, \beta, \alpha, \theta_I)$

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)).

Model fitting in Bayesian framework

- For 'complete' data $\boldsymbol{x}$ (e.g. times and nature of all transitions) $\pi(\boldsymbol{x} \mid \theta)$ tractable
- Given censored/filtered/noisy data $\boldsymbol{y}$, $\pi(\boldsymbol{y} \mid \theta)$ typically intractable
- Use data augmentation and sample from

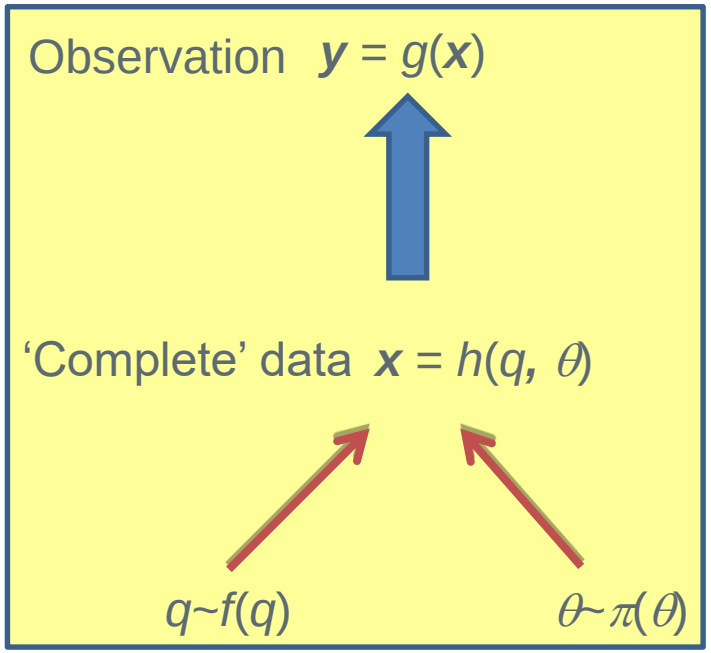
$$\pi(\theta, \boldsymbol{x} \mid \boldsymbol{y}) \propto \pi(\theta)\pi(\boldsymbol{x}, \boldsymbol{y} \mid \theta)$$

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)

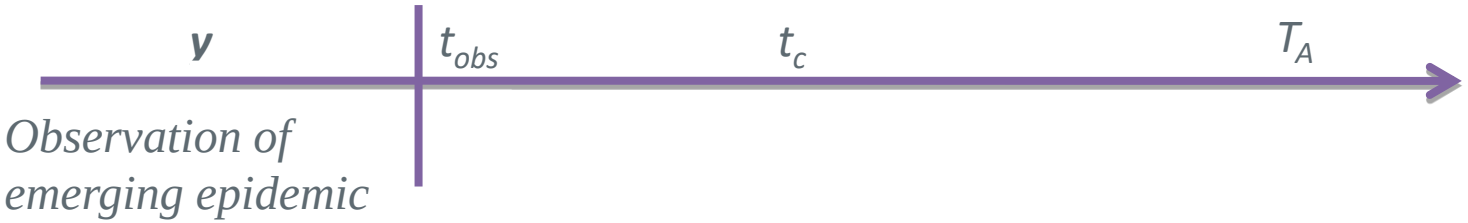
Assigns threshold q_i to each individual. If $R_i(t)$ denotes infectious challenge to i at time t, infection time x_i occurs when integrated challenge reaches threshold

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

- Epidemic dynamics specified (for SI with cryptic) by q (vector of Sellke Thresholds) and θ, i.e. ($\boldsymbol{x} = h(q, \theta)$).
- For controls $\boldsymbol{d}$, based on removal of infected individuals, it follows that $\boldsymbol{x} = h^*(q, \theta, \boldsymbol{d})$.
- Gives a means of coupling epidemic trajectories under different control strategies.

Control strategies

- Based on removal of hosts found to be infected at control time t_c.
- 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 $t > t_{obs}$

Candidate measures - (x_j denotes infection time of j)

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

Simulated epidemic 1

- Host population of size 1000 uniformly located over square region
- Simulate epidemic from SI model with:

$\alpha = 0.08, \beta = 7.10^{-6}, \epsilon = 5.10^{-5}$ and

$$K(d, \alpha) = \frac{1}{2\pi d\alpha} \exp(-d/\alpha)$$

- Observed data ***y***, snapshots of symptomatic sets at t =130, 160, 190,….,460
- Control applied at t_c=460, 470
- Performance measure – reduction in posterior expectation of number of infections up to time t_A=500 (relative to uncontrolled epidemic)

Estimating expected reduction

- Use random sample from $\pi(\theta, \boldsymbol{x}(t) \mid \boldsymbol{y})$ to generate sample of size m from $\pi(\theta, Q \mid \boldsymbol{y})$.
- Let $u(x(T))$ denote the number of infections by time T for trajectory $x(T)$. Let d denote control strategy.

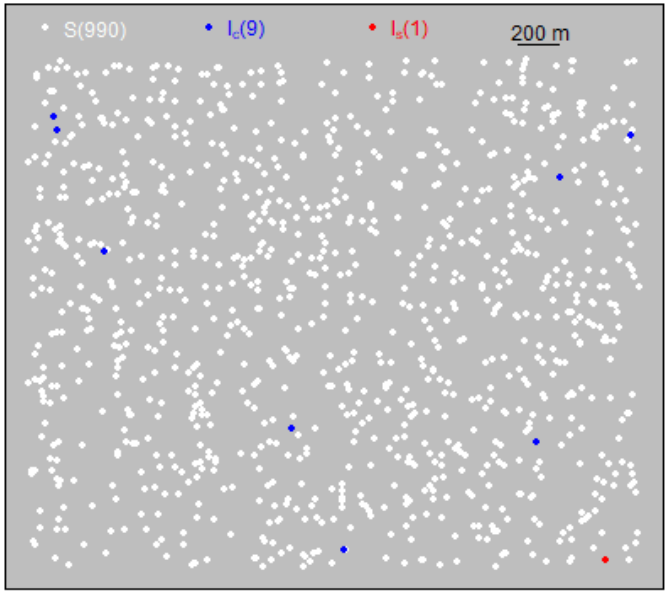
$$x(T) = h(\theta, Q), \qquad x_d(T) = h^*(\theta, Q, d)$$

$$\text{EER}(d) = \frac{1}{n}\Sigma_i \{u(h^*(\theta_i, Q_i, d)) - u(h(\theta_i, Q_i))\}$$

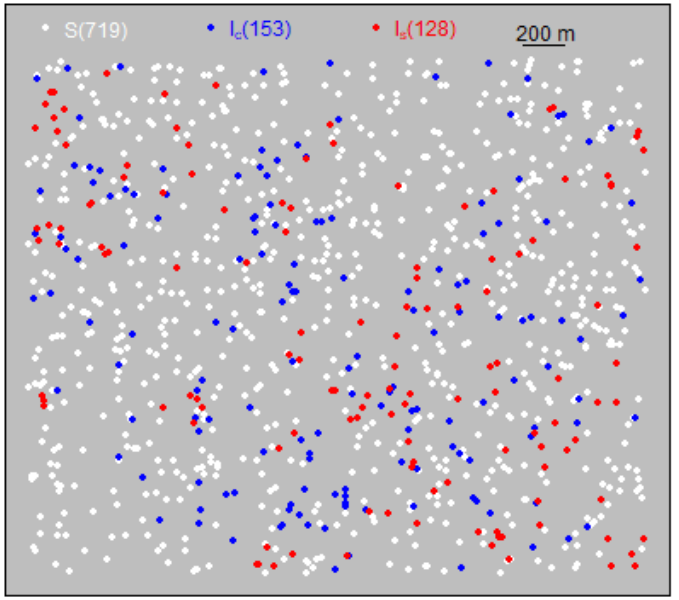
- Here we take m = 1000 draws of (θ_i, Q_i) using these as a test-bed of ‘pre-epidemics’ on which to compare controls.

Results – snapshots of system state

t = 130 (days)

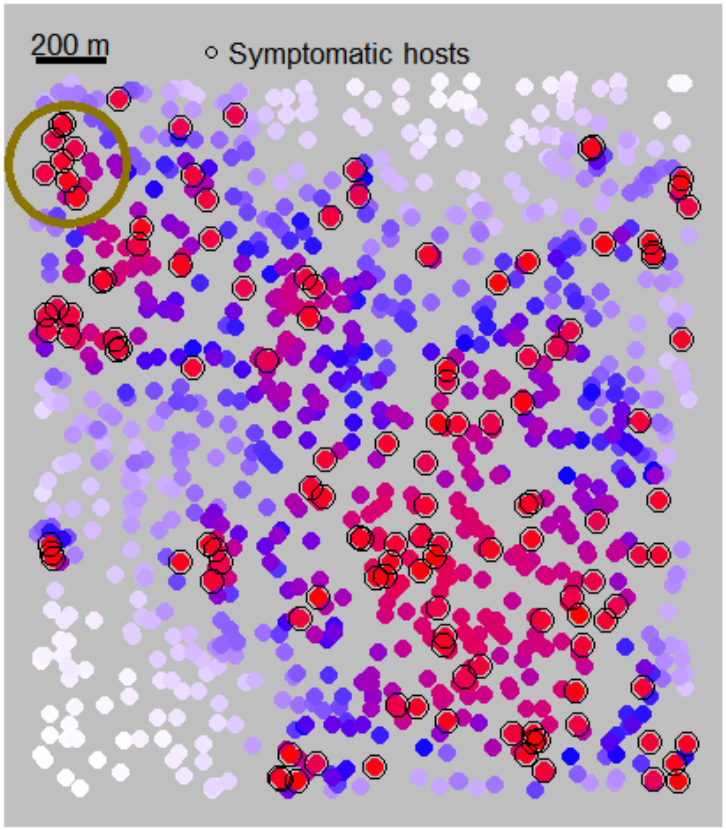


t = 460 (days)

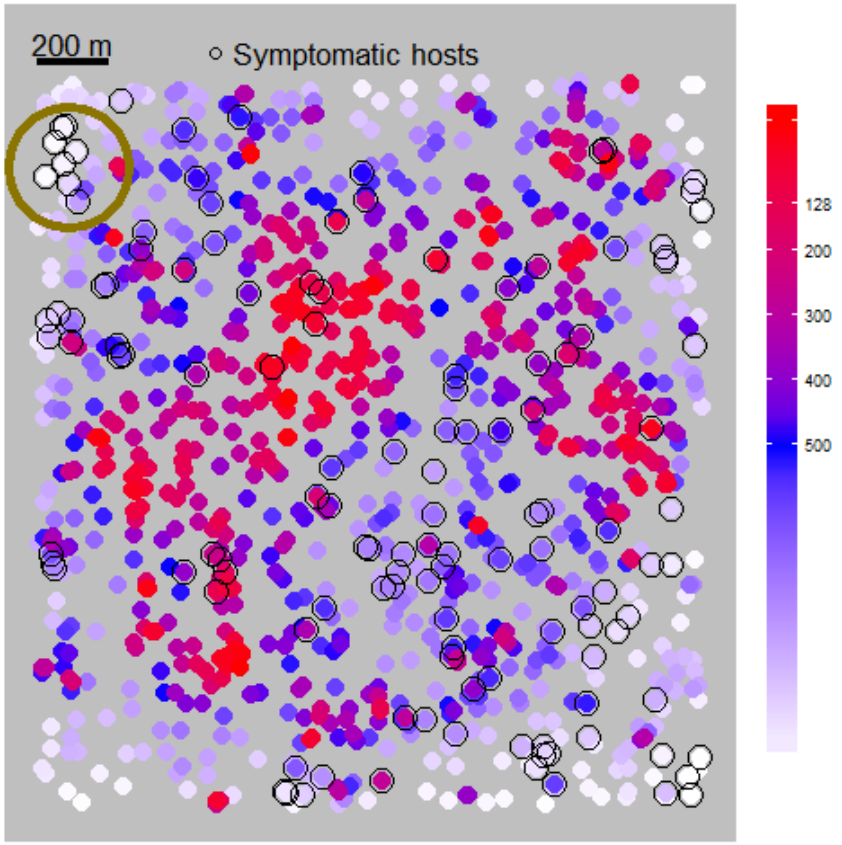


Maps of risk, hazard threat

risk

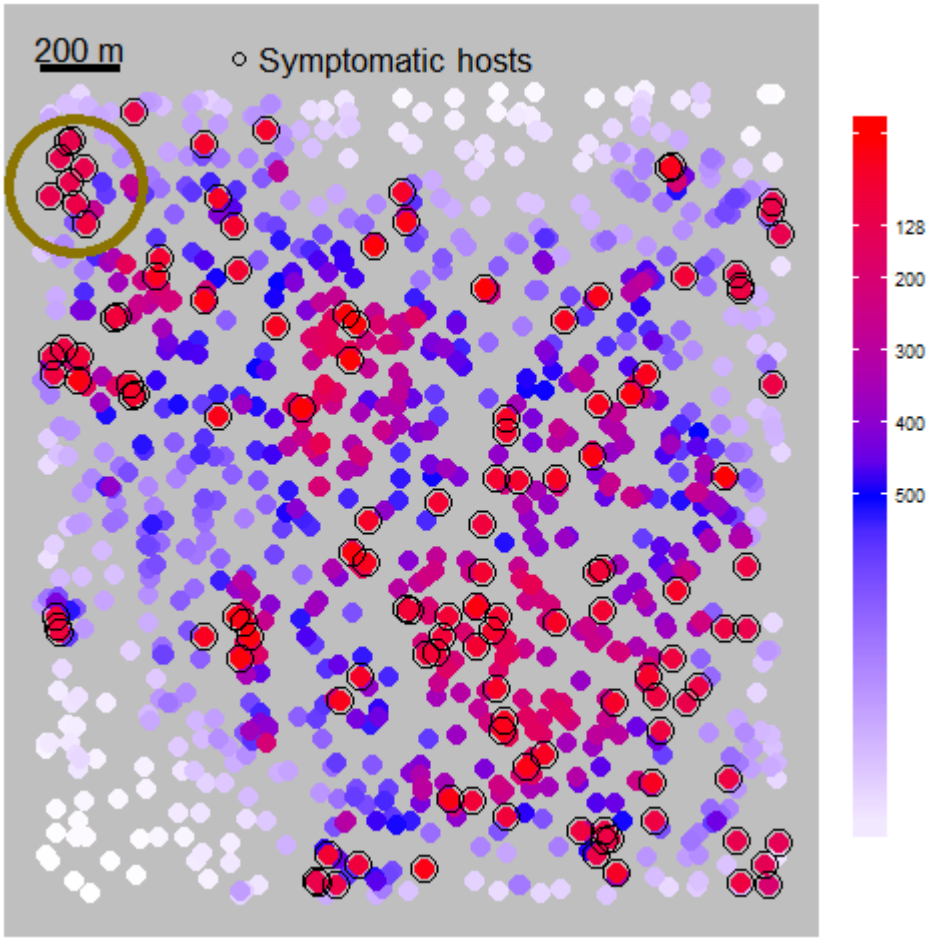


hazard



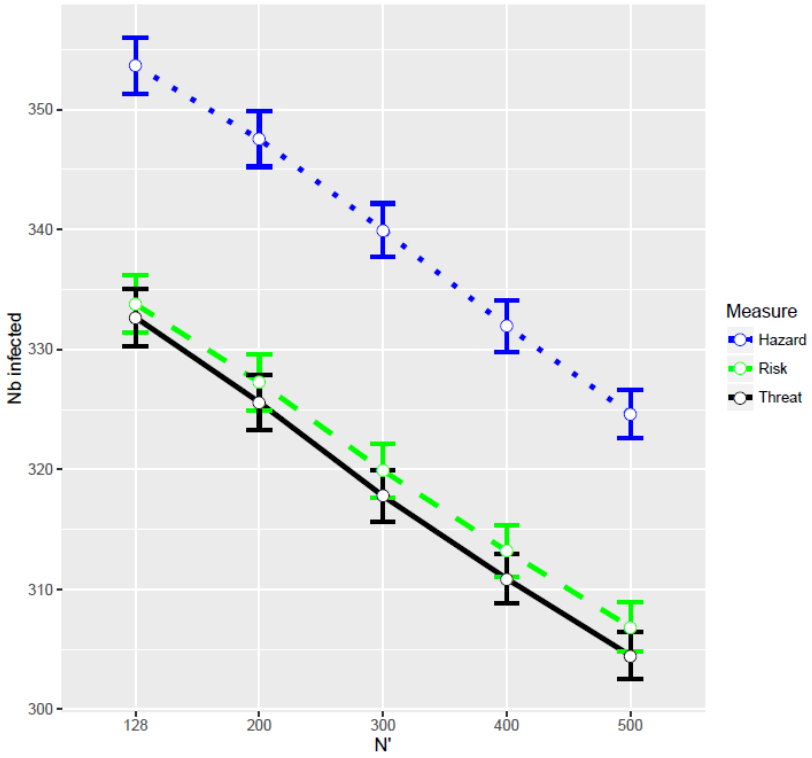
- The hazard values are greatest in regions of low infection while the risk
- measure is greatest for symptomatic individuals.

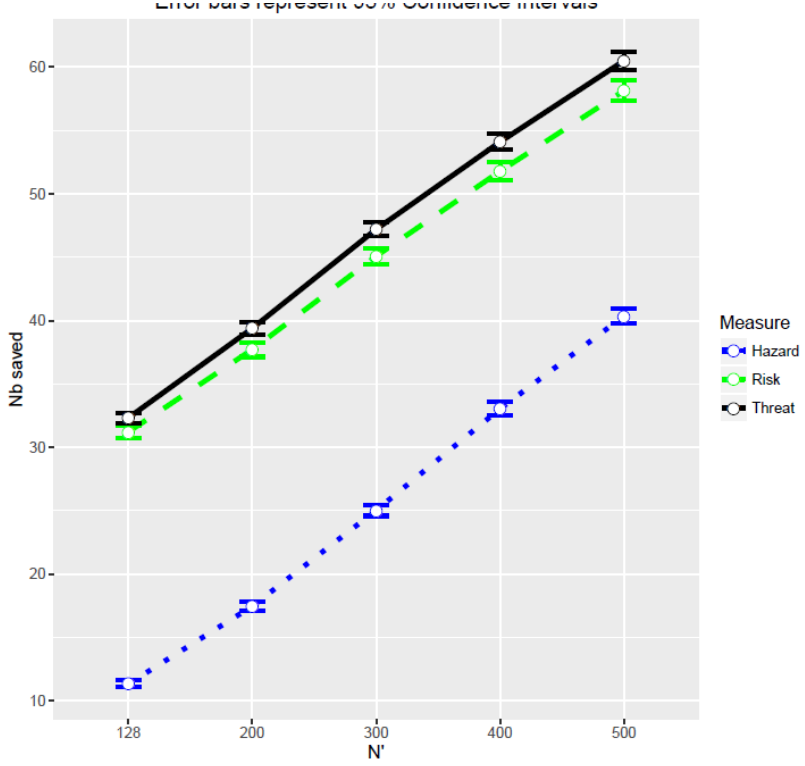
threat



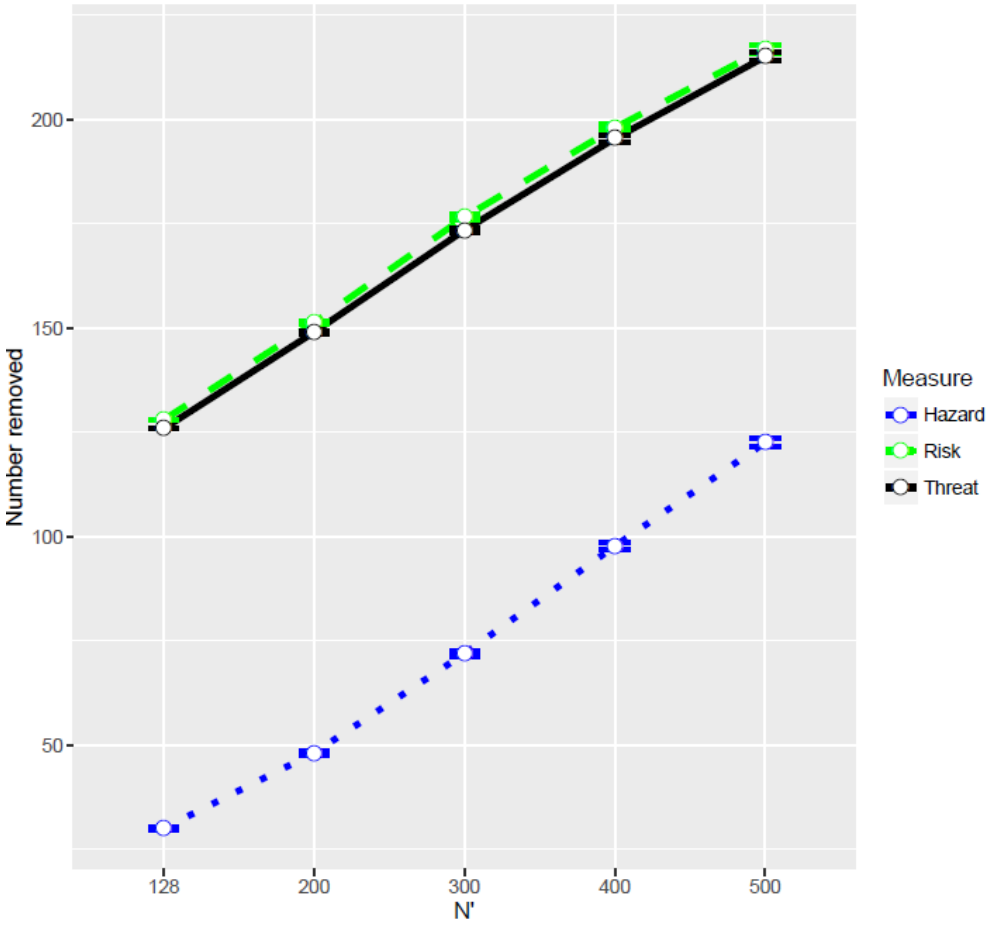
- The dependence of the threat measure on the positions of likely susceptible individuals in relation to an infected host can be discerned
- The infected hosts (circled) in the top left corner of the population naturally exhibit high values of the risk.
- The corresponding threat measure is comparatively lower for these hosts, as a high proportion of their immediate neighbours are already infected.

EER(t_A) and expected infections



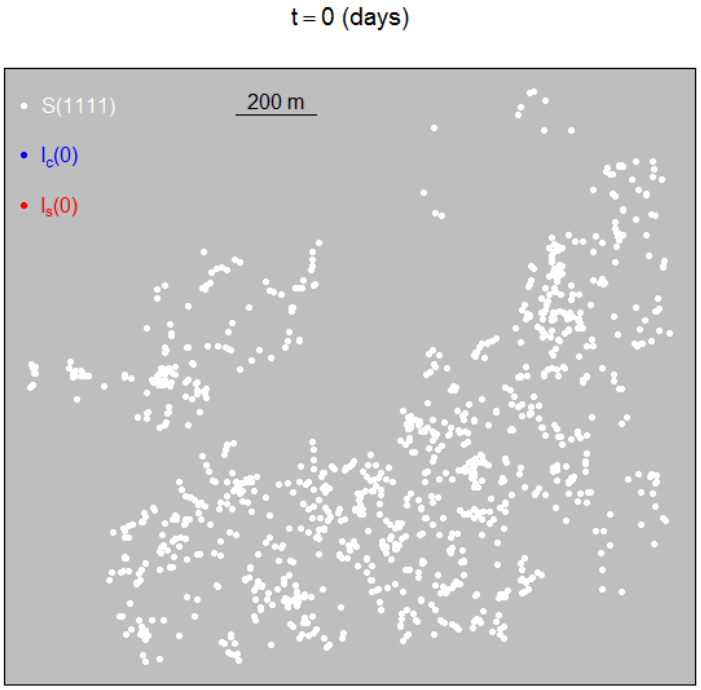


Expected number of removals



- Marginally improved control provided by threat map at expense of marginally fewer removals.
- Risk and threat generally comparable – suggesting risk of infection is main determinant in threat map.
- Intuitive given uniform distribution of hosts (?) and relatively homogenous appearance of the epidemic.

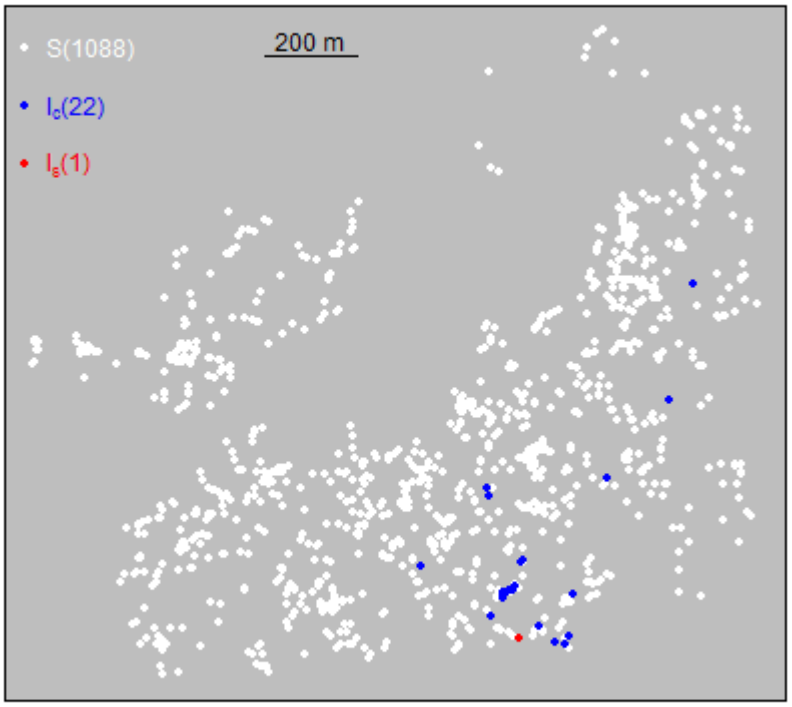
Clustered host populations



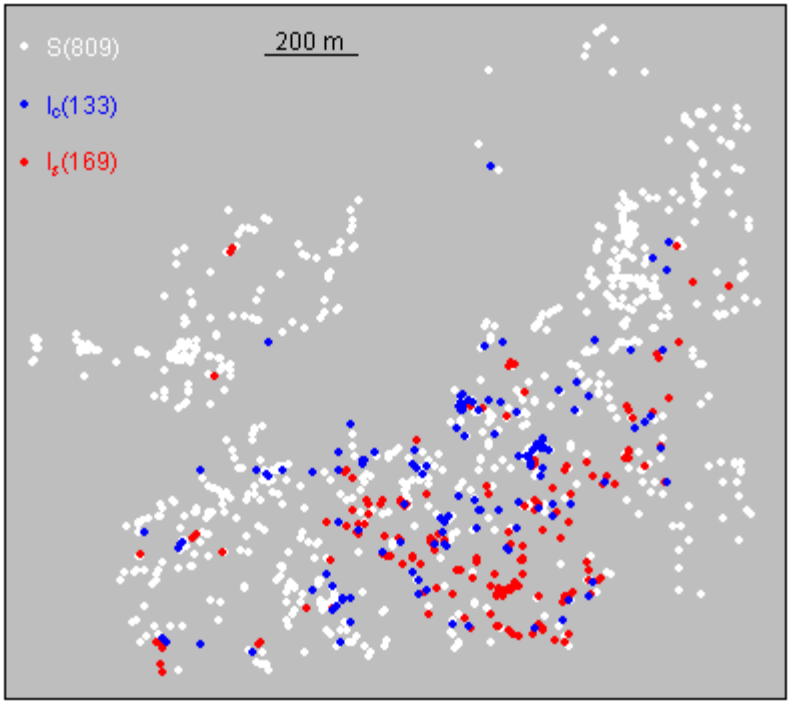
- Citrus locations from Broward county
- 1111 trees spatially distributed
- Citrus canker epidemic on this population analysed by Neri et al (2014)
- Canker typically controlled using ring-culling strategies (not yet considered in this framework but amenable to it)
- Simulate epidemics of 2 types:
  - exponential kernel with primary
  - exponential kernel no primary

Snapshots with primary

t = 130 (days)

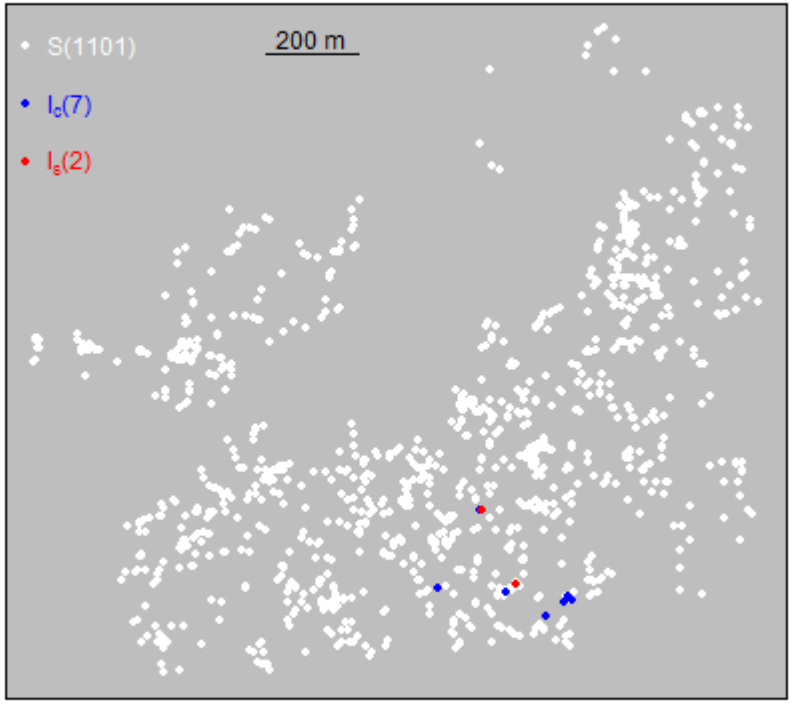


t = 460 (days)

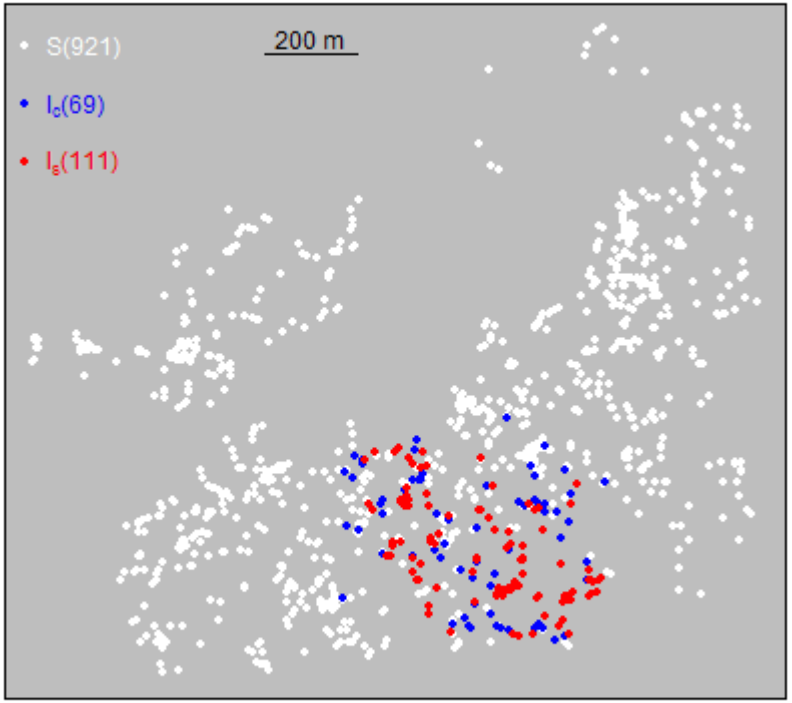


Snapshots without primary

t = 130 (days)

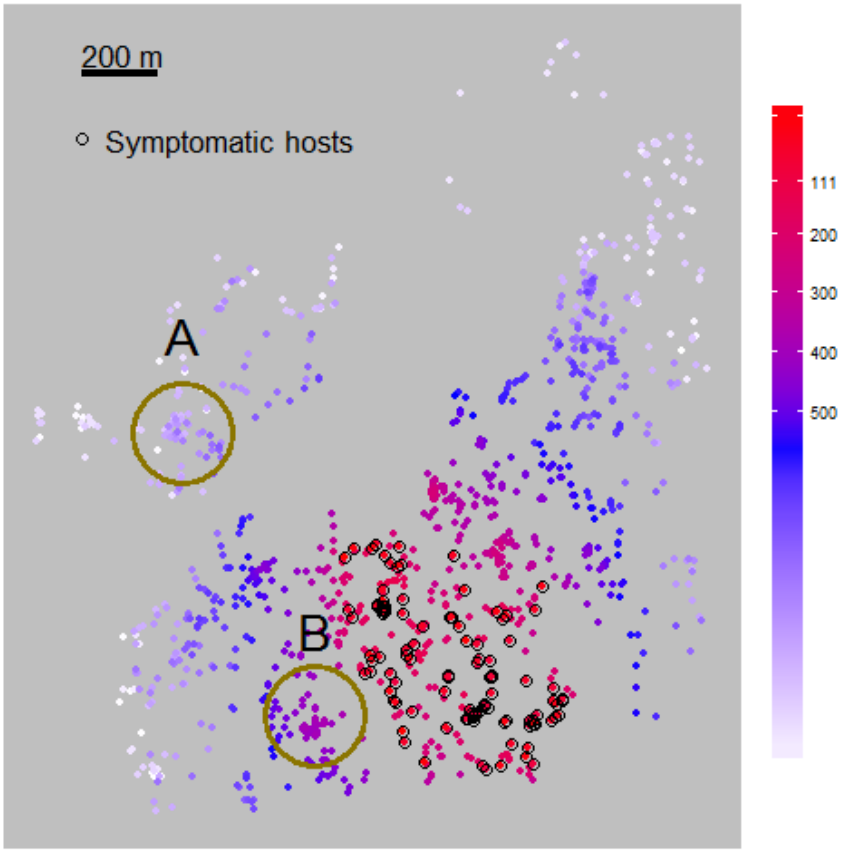


t = 460 (days)

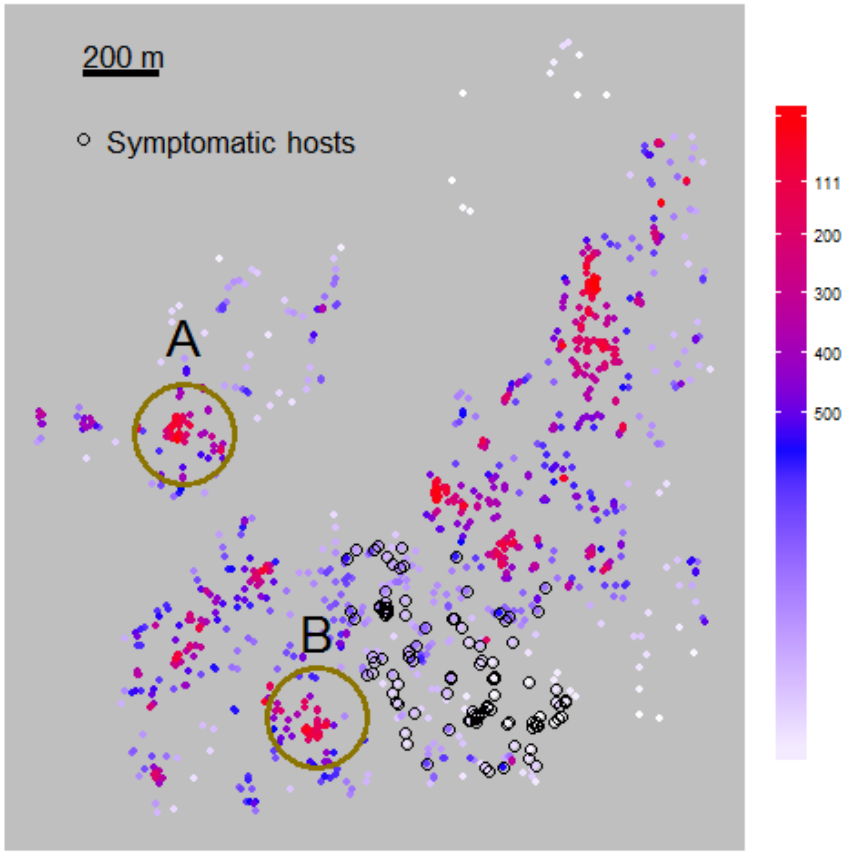


Maps, no primary

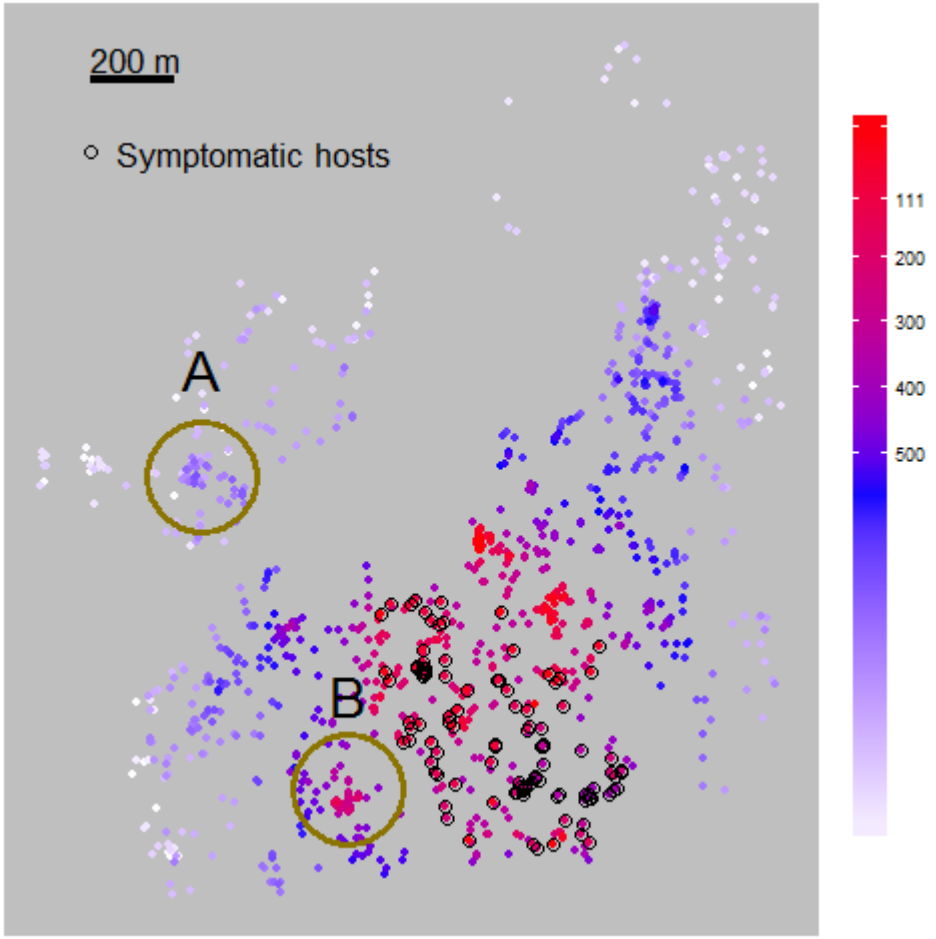
Risk



Hazard

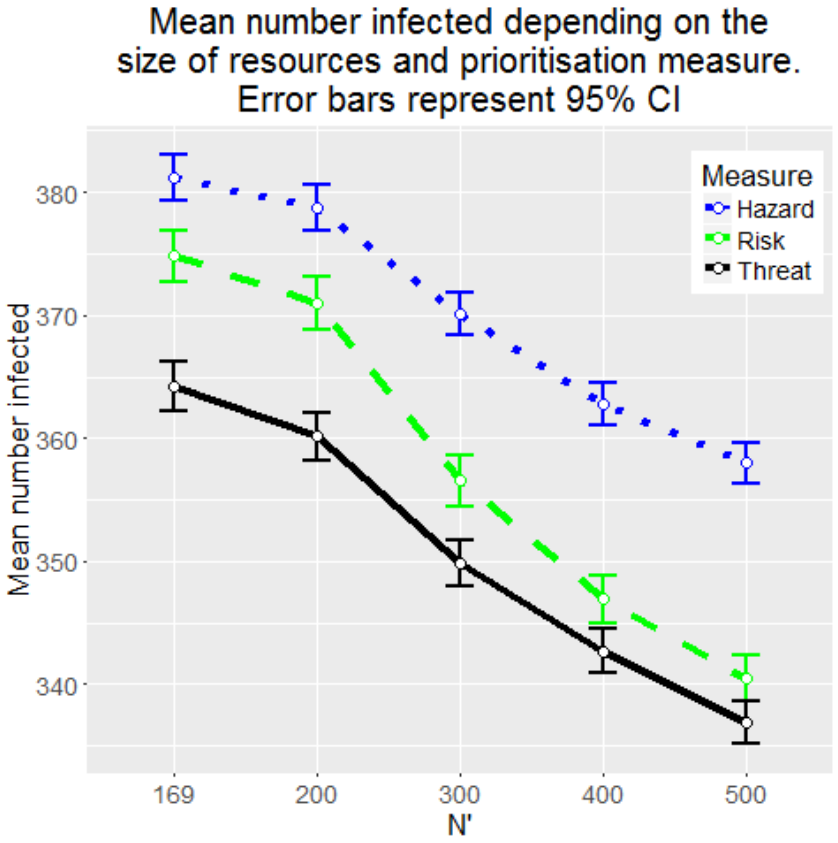


Threat

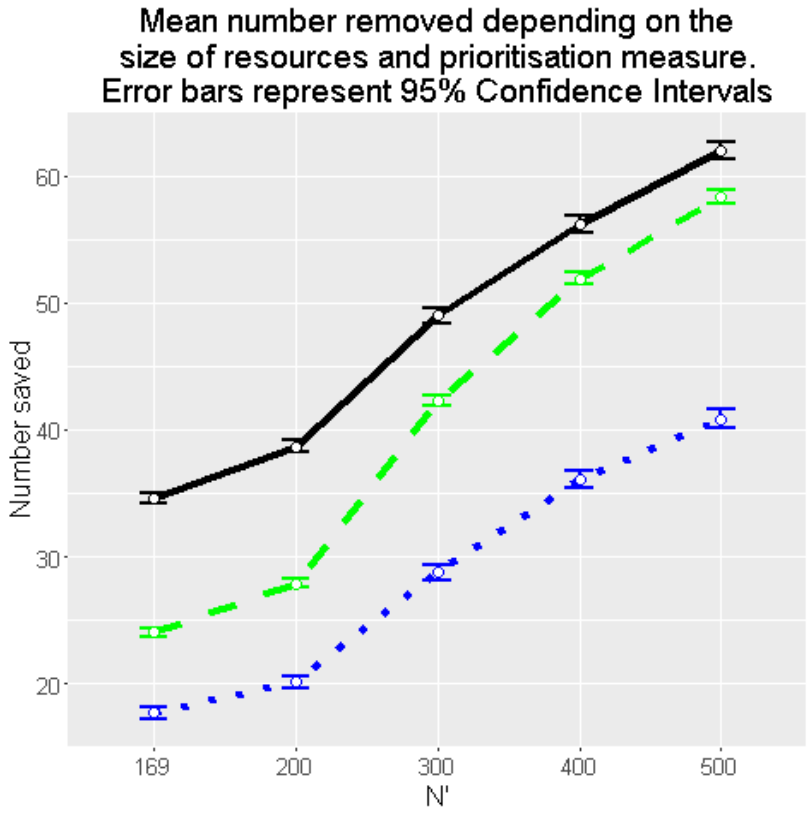


- The 111 symptomatic hosts detected during the survey are indicated by the black circles.
- A cluster with intermediate risk (B) leads to high threat due to the high hazard.
- while one with very low risk (A) ends up with relatively low threat even though the hazard is high.

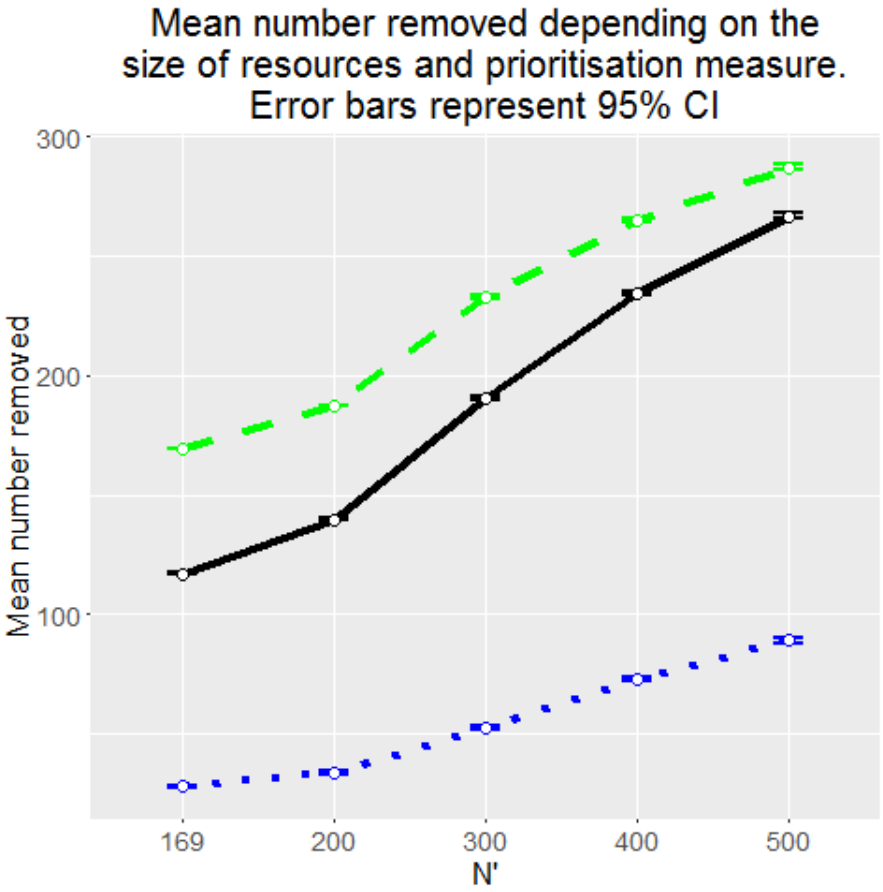
EER(t_A) and expected infections, Case (I)



d

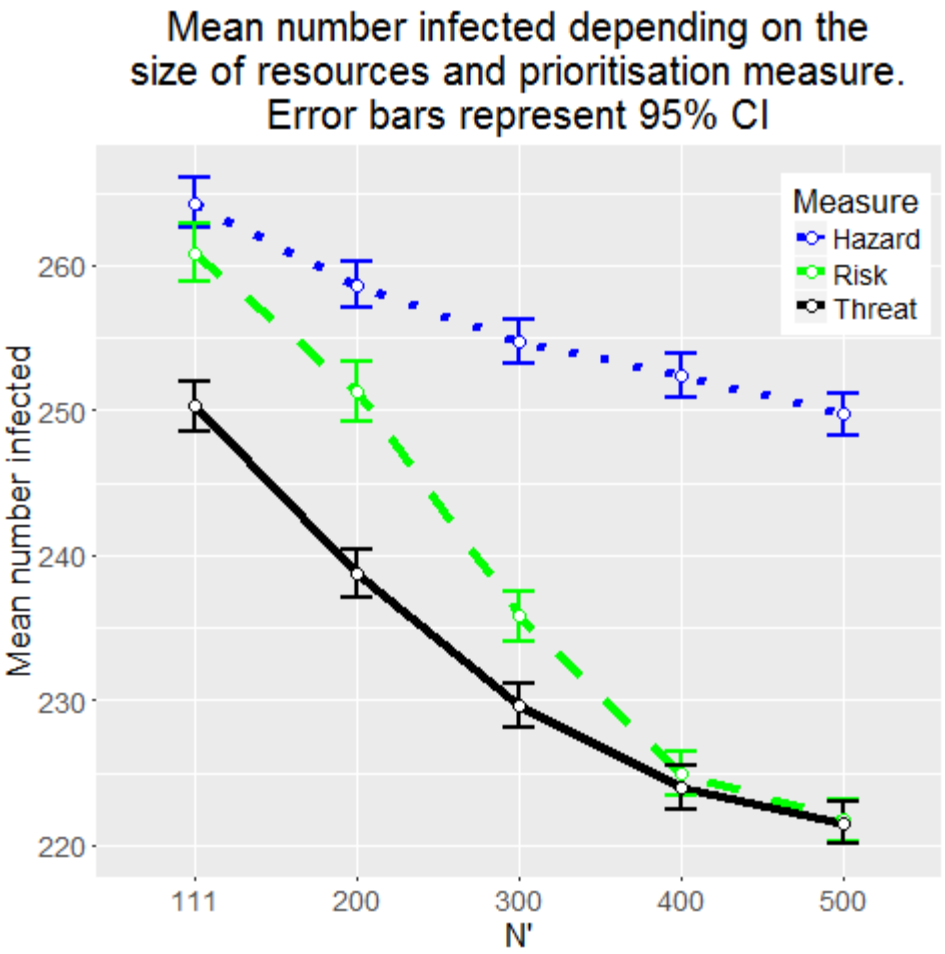


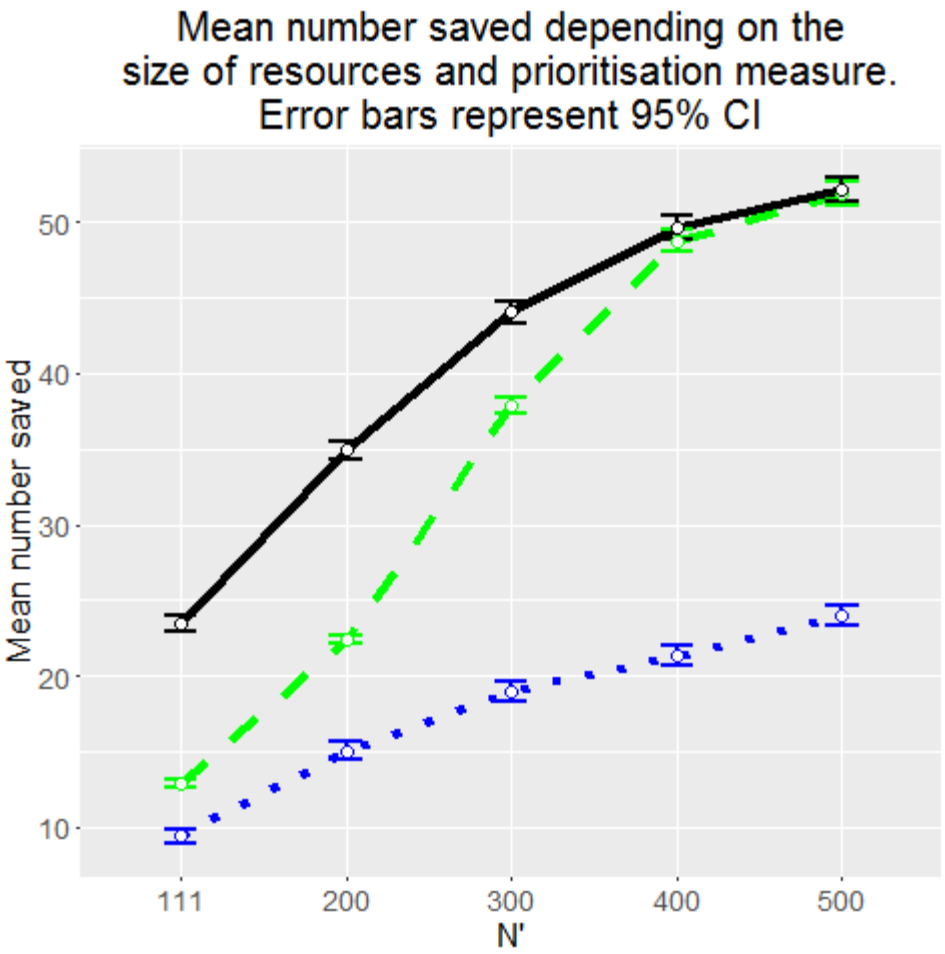
Expected number of removals, Case (I)



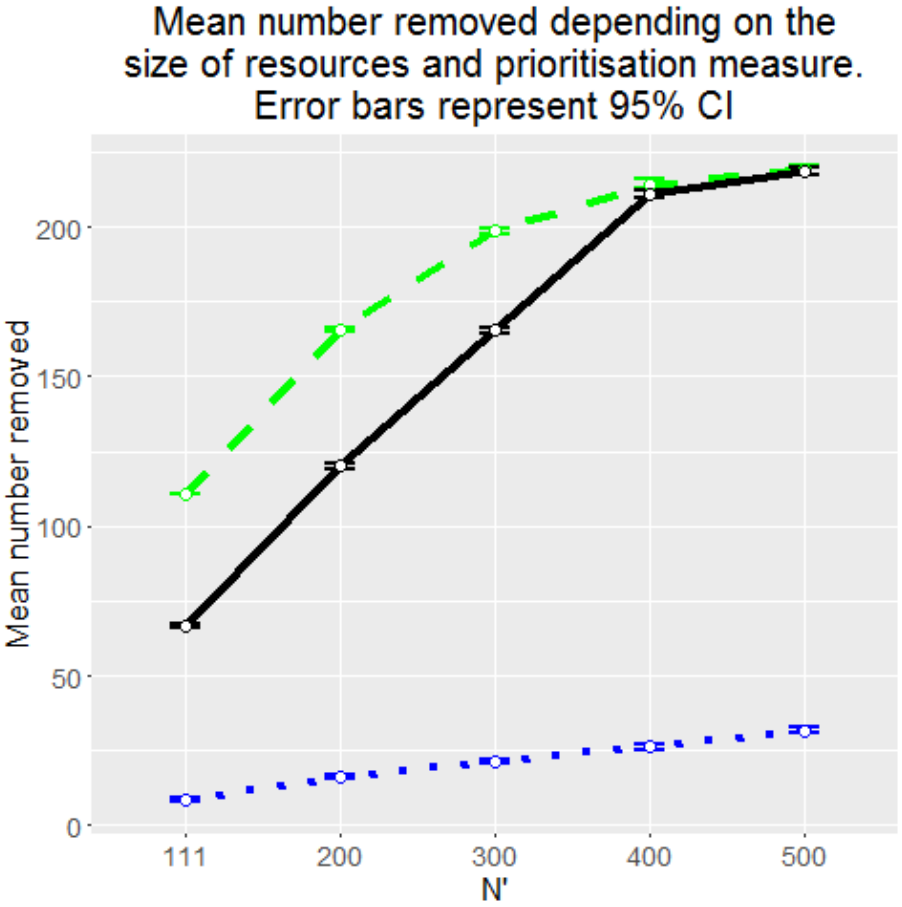
- Difference in performance between the risk and threat measure than was observed for the uniformly distributed population. prioritisation based
- Prioritisation based on the threat map is the most cost-effective control strategy in reducing the impact of the epidemics.
- With scarce resources (lower values of N') the difference between results for the threat and risk measure decreasing as N' increases.
- The change in the discrepancy between threat and risk maps with increasing N' is most pronounced in Case (II),

EER(t_A) and expected infections, Case (II)





Expected number of removals, Case (II)



- For small values of N' the risk map's performance improves little on that of the hazard map but converges to that of the threat map as N' approaches its maximal value.
- Less removal with the threat compared to the risk and hazard with higher host saved.

Summing up

- Data augmentation valuable for designing control strategies.
- Removal of hosts based on the threat map is the most effective strategy to reduce the impact of an epidemic – even though fewer hosts are targeted for removal.
- Latent processes (Sellke thresholds) can be used to couple epidemics and subsequently reduce the variability in the difference of control strategies.
- Sample size needed for the estimation is reduced compared to an independent sampling.
- The approach is parallisable .


UNIVERSITY OF
CAMBRIDGE


HERIOT
WATT
UNIVERSITY

Thank you


USDA
United States
Department of
Agriculture