

Catalyst: agents of change

Integration of compartment and agent-based models for epidemiology

Shannon K. Gallagher

Carnegie Mellon University, Department of Statistics & Data Science

Tuesday, July 9, 2019

- William F. Eddy, Advisor
- Joel Greenhouse
- Howard Seltman
- Samuel L. Ventura

Motivation: measles outbreak in Hagelloch, Germany 1861

- 188 susceptible children
- 56 households
- Initial infection date: 10/30/1861
- Final infection date: 01/27/1862
- 12 deaths



Hagelloch, Germany

What can we learn from this outbreak?

- Was this outbreak of measles comparable to others?
- Does the spatial/class structure contribute anything?
- How easy is it to become infected?

We want to make models to answer these questions

- Specify **how** and **why** a disease moves through a population

We want to make models to answer these questions

- Specify **how** and **why** a disease moves through a population
- "A Thousand and One Epidemic Models" (Hethcote 1994)
 - epidemiological compartment structure, i.e. **states** (susceptible, infectious, recovered, ...)
 - incidence and distribution of waiting times
 - demographic structure
 - epidemiological-demographic interactions
 - homogeneous interaction:
 - $P(A_{m,t} = i | A_{m,t-1} = j) = P(A_{n,t} = i | A_{n,t-1} = j)$ for individuals A_m and A_n

We want to make models to answer these questions

- Specify **how** and **why** a disease moves through a population
- "A Thousand and One Epidemic Models" (Hethcote 1994)
 - epidemiological compartment structure, i.e. **states** (susceptible, infectious, recovered, ...)
 - incidence and distribution of waiting times
 - demographic structure
 - epidemiological-demographic interactions
 - homogeneous interaction:
 - $P(A_{m,t} = i | A_{m,t-1} = j) = P(A_{n,t} = i | A_{n,t-1} = j)$ for individuals A_m and A_n
- Estimate parameters

We want to make models to answer these questions

- Specify **how** and **why** a disease moves through a population
- "A Thousand and One Epidemic Models" (Hethcote 1994)
 - epidemiological compartment structure, i.e. **states** (susceptible, infectious, recovered, ...)
 - incidence and distribution of waiting times
 - demographic structure
 - epidemiological-demographic interactions
 - homogeneous interaction:
 - $P(A_{m,t} = i | A_{m,t-1} = j) = P(A_{n,t} = i | A_{n,t-1} = j)$ for individuals A_m and A_n
- Estimate parameters
- Understand the disease

We want to make models to answer these questions

- Specify **how** and **why** a disease moves through a population
- "A Thousand and One Epidemic Models" (Hethcote 1994)
 - epidemiological compartment structure, i.e. **states** (susceptible, infectious, recovered, ...)
 - incidence and distribution of waiting times
 - demographic structure
 - epidemiological-demographic interactions
 - homogeneous interaction:
 - $P(A_{m,t} = i | A_{m,t-1} = j) = P(A_{n,t} = i | A_{n,t-1} = j)$ for individuals A_m and A_n
- Estimate parameters
- Understand the disease
- Change the next outbreak

We want to make models to answer these questions

- Specify **how** and **why** a disease moves through a population
- "A Thousand and One Epidemic Models" (Hethcote 1994)
 - epidemiological compartment structure, i.e. **states** (susceptible, infectious, recovered, ...)
 - incidence and distribution of waiting times
 - demographic structure
 - epidemiological-demographic interactions
 - homogeneous interaction:
 - $P(A_{m,t} = i | A_{m,t-1} = j) = P(A_{n,t} = i | A_{n,t-1} = j)$ for individuals A_m and A_n
- Estimate parameters
- Understand the disease
- Change the next outbreak

Which model(s) do we choose?

Compartment (CM) and agent-based (AM) models

Given disease-level states (e.g. Susceptible, Infectious)...

- **CM**– (Stochastic) equations which specify how individuals move through a disease
 - e.g. $S(t) = S(t-1) - \beta \cdot S(t-1)$
 - Base unit is **# of individuals in a state**
 - **homogeneous** interactions among individuals in different states

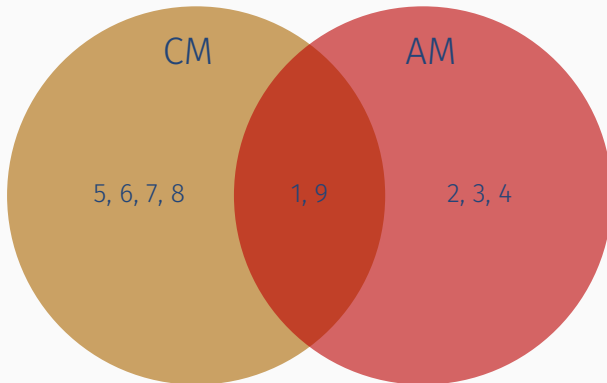
Compartment (CM) and agent-based (AM) models

Given disease-level states (e.g. Susceptible, Infectious)...

- **CM**– (Stochastic) equations which specify how individuals move through a disease
 - e.g. $S(t) = S(t - 1) - \beta \cdot S(t - 1)$
 - Base unit is **# of individuals in a state**
 - **homogeneous** interactions among individuals in different states
- **AM**– (Stochastic) simulations which specify how agents move through a disease
 - e.g. $A_{t,n} = I$ if condition $E|F|G$
 - Base unit is an **agent** (an individual)
 - **heterogeneous** interactions among individuals in different states

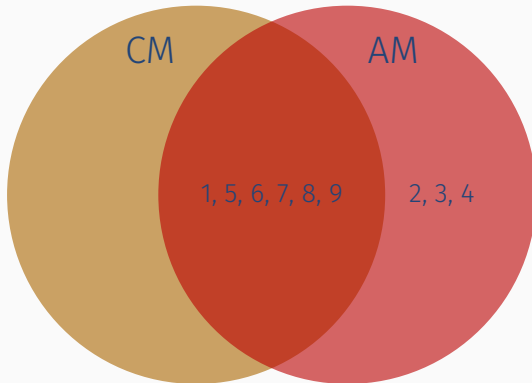
We compare CMs to AMs

Quality	CM	AM
1. Interpretable	✓	✓
2. Accessible		✓
3. Modular		✓
4. Individual info		✓
5. Fast computer run time	✓	
6. Low computer memory	✓	
7. Theory	✓	
8. Parameter estimation	✓	
9. Statistical software	✓	✓



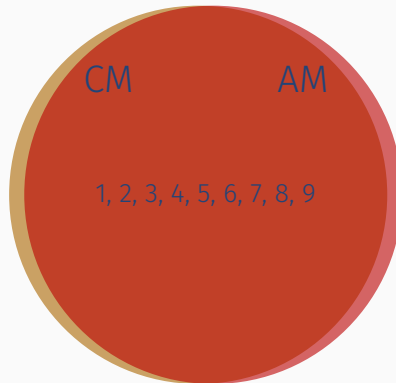
We compare CMs to AMs

Quality	CM	AM
1. Interpretable	✓	✓
2. Accessible	✓	✓
3. Modular	✓	✓
4. Individual info	✓	✓
5. Fast computer run time	✓	
6. Low computer memory	✓	
7. Theory	✓	
8. Parameter estimation	✓	
9. Statistical software	✓	✓



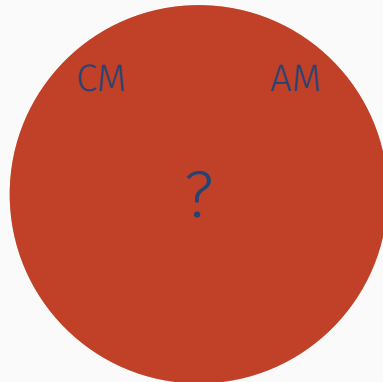
We compare CMs to AMs

Quality	CM	AM
1. Interpretable	✓	✓
2. Accessible	✓	✓
3. Modular	✓	✓
4. Individual info	✓	✓
5. Fast computer run time	✓	✓
6. Low computer memory	✓	✓
7. Theory	✓	✓
8. Parameter estimation	✓	✓
9. Statistical software	✓	✓



We compare CMs to AMs

Quality	CM	AM
1. Interpretable	✓	✓
2. Accessible	✓	✓
3. Modular	✓	✓
4. Individual info	✓	✓
5. Fast computer run time	✓	✓
6. Low computer memory	✓	✓
7. Theory	✓	✓
8. Parameter estimation	✓	✓
9. Statistical software	✓	✓



Are these classes statistically the same?

Dissertation Goals

1. Statistically relate CMs and AMs
2. Model selection methodology for CM-AM pairs
3. Apply methods to applications

Relating CMs and AMs

Epidemiological states

- $S(t)$ - # Susceptible individuals at t
- $I(t)$ - # Infectious individuals at t
- $R(t)$ - # Recovered individuals at t

Kermack and McKendrick CM (1927) - Deterministic transitions

Epidemiological states

- $S(t)$ - # Susceptible individuals at t
- $I(t)$ - # Infectious individuals at t
- $R(t)$ - # Recovered individuals at t

Demographics and interactions

- β - rate of infection
- γ - rate of recovery
- N - fixed population size
- $S(0), I(0), R(0)$ known

Kermack and McKendrick CM (1927) - Deterministic transitions

Epidemiological states

- $S(t)$ - # Susceptible individuals at t
- $I(t)$ - # Infectious individuals at t
- $R(t)$ - # Recovered individuals at t

Demographics and interactions

- β - rate of infection
- γ - rate of recovery
- N - fixed population size
- $S(0), I(0), R(0)$ known

Incidence and distributions

$$\left\{ \begin{array}{l} \frac{\Delta S}{\Delta t} = -S \times \beta \frac{I}{N} \\ \frac{\Delta I}{\Delta t} = S \times \beta \frac{I}{N} - I \times \gamma \\ \frac{\Delta R}{\Delta t} = I \times \gamma \end{array} \right.$$

Stochastic CM based on K&M (Gallagher and Eddy, 2017a)

For $t = 1, \dots, T$, S_0, I_0, R_0 known

$$S_t | S_{t-1}, I_{t-1} = S_{t-1} - \text{Binomial} \left(S_{t-1}, \beta \frac{I(t-1)}{N} \right)$$

$$R_t | S_{t-1}, I_{t-1} = R_{t-1} + \text{Binomial} (I(t-1), \gamma)$$

- Model is unbiased w.r.t original K&M equations
- We show recursive/closed form of variance – uncertainty \uparrow as $t \uparrow$
- We estimate $\hat{\beta}, \hat{\gamma}$ using likelihood or sum of squares
 - See (Gallagher et al., in prep. 2019a)

We can also make an AM for this scenario (Gallagher and Eddy, 2017a)

For an agent $A_{t,n}$, $n = 1, 2, \dots, N$, with A_0 known, the agent update is given by for $t = 1, \dots, T$

$$A_{t,n}|A_{t-1} = \begin{cases} 1 + \text{Bernoulli}\left(\frac{\beta X_{t-1,2}}{N}\right) & \text{if } A_{t-1,n} = 1 \\ 2 + \text{Bernoulli}(\gamma) & \text{if } A_{t-1,n} = 2 \\ 3 & \text{if } A_{t-1,n} = 3 \end{cases}$$

$A_{t,n} \in \{1, 2, 3\}$ where $1 \rightarrow S$, $2 \rightarrow I$, and $3 \rightarrow R$

Let $X_{t,k} = \sum_{n=1}^N \mathcal{I}\{A_{t,n} = k\}$ be the # of agents in state k at time t .

SIR Notation: $(X_{t,1} \rightarrow S_t^{AM}, X_{t,2} \rightarrow I_t^{AM}, X_{t,3} \rightarrow R_t^{AM})$

For a non-random SIR-CM, we can write “equiv.” CM-AM pairs

Ex. K&M SIR

CM

$$S_t | S_{t-1}, I_{t-1} = S_{t-1} - \text{Binomial} \left(S_{t-1}, \beta \frac{I_{t-1}}{N} \right)$$

$$R_t | S_{t-1}, I_{t-1} = R_{t-1} + \text{Binomial} (I_{t-1}, \gamma)$$

AM

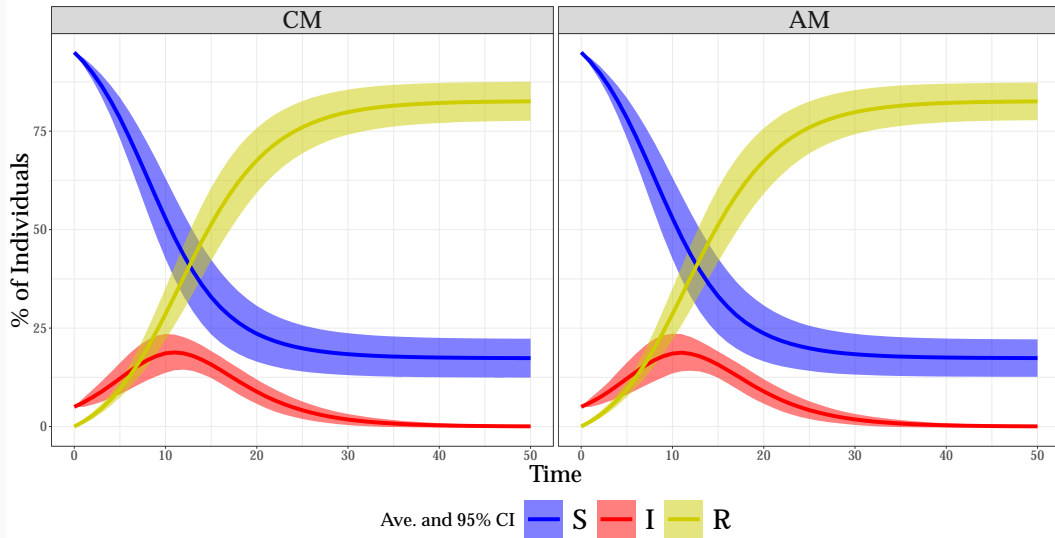
$$A_{t,n} | A_{t-1} = \begin{cases} 1 + \text{Bernoulli} \left(\frac{\beta X_{t-1,2}}{N} \right) & \text{if } A_{t-1,n} = 1 \\ 2 + \text{Bernoulli} (\gamma) & \text{if } A_{t-1,n} = 2 \\ 3 & \text{if } A_{t-1,n} = 3 \end{cases}$$

Take away: $(S_t, I_t, R_t)^{(CM)} \stackrel{d}{=} (S_t, I_t, R_t)^{(AM)}$ for all $t = 0, \dots, T$

'Proof' by picture

Simulations of SIR

$N = 1000$, $\beta = 0.50$, $\gamma = 0.25$, $S_0 = 950$, $I_0 = 50$



We can write “equivalent” CM-AM pairs (Gallagher and Eddy, 2017b)

Theorem 1: Given deterministic transition matrix $D(t)$ of size $K \times K$, there exists a stochastic CM-AM pair such that $X^{CM} \stackrel{d}{=} X^{AM}$ and the models are unbiased w.r.t $D(t)$

- K is the number of states
- $D_{ij}(t)$ is the **non-negative** # of individuals moving from state i to j from time t to $t + 1$
- Row sums are total number individuals moving out of state i
- Column sums are total number of individuals moving into state j
- $D(t) - D^T(t)$ gives back the original difference equations

We can write “equivalent” CM-AM pairs (Gallagher and Eddy, 2017b)

Theorem 1: Given deterministic transition matrix $D(t)$ of size $K \times K$, there exists a stochastic CM-AM pair such that $X^{CM} \stackrel{d}{=} X^{AM}$ and the models are unbiased w.r.t $D(t)$

- K is the number of states
- $D_{ij}(t)$ is the **non-negative** # of individuals moving from state i to j from time t to $t + 1$
- Row sums are total number individuals moving out of state i
- Column sums are total number of individuals moving into state j
- $D(t) - D^T(t)$ gives back the original difference equations

Ex. SIR $D(t)$:

$$D(t) = \begin{pmatrix} S(t) - \beta S(t) \frac{I(t)}{N} & \beta S(t) \frac{I(t)}{N} & 0 \\ 0 & I(t) - I(t)\gamma & I(t)\gamma \\ 0 & 0 & R(t) \end{pmatrix}$$

We can write “equivalent” CM-AM pairs (Gallagher and Eddy, 2017b)

Theorem 1: Given deterministic transition matrix $D(t)$ of size $K \times K$, there exists a stochastic CM-AM pair such that $X^{CM} \stackrel{d}{=} X^{AM}$ and the models are unbiased w.r.t $D(t)$

- K is the number of states
- $D_{ij}(t)$ is the **non-negative** # of individuals moving from state i to j from time t to $t + 1$
- Row sums are total number individuals moving out of state i
- Column sums are total number of individuals moving into state j
- $D(t) - D^T(t)$ gives back the original difference equations

Given D , there exists a stochastic, equivalent CM-AM pair

We do not need D to have CM-AM pairs

Theorem 2 (Gallagher and Eddy, 2019a).

Any CM with M states has an equivalent AM pair with M states and homogeneous agent interactions, in terms of numbers of individuals in each state at each time

- i.e. $\chi^{CM} \stackrel{d}{=} \chi^{AM}$

We do not need D to have CM-AM pairs

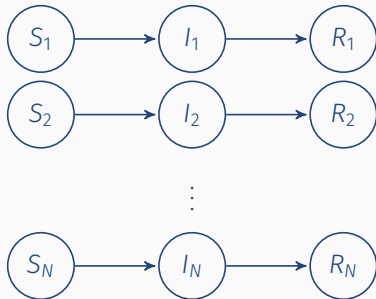
Result 3.

Each AM has an equivalent CM pair **provided** we adjust the number of total states K^* , in terms of numbers of individuals in each state at each time.

- M – number of disease level states. (e.g. SIR $\implies M=3$)
- N – number of individuals/agents
- $K^* = MN$



vs.



We do not need D to have CM-AM pairs

- **Theorem 2** (Gallagher and Eddy, 2019a). Any CM with M states has an equivalent AM pair with M states and homogeneous agent interactions, in terms of numbers of individuals in each state at each time.
- **Result 3**. Each AM has an equivalent CM pair **provided** we adjust the number of total states K^* , in terms of numbers of individuals in each state at each time. **provided** we adjust the number of total states K^*
 - M – number of disease level states. (e.g. SIR $\implies M=3$)
 - N – number of individuals/agents
 - $K^* = MN$

We do not need D to have CM-AM pairs

- **Theorem 2** (Gallagher and Eddy, 2019a). Any CM with M states has an equivalent AM pair with M states and homogeneous agent interactions, in terms of numbers of individuals in each state at each time.
- **Result 3**. Each AM has an equivalent CM pair **provided** we adjust the number of total states K^* , in terms of numbers of individuals in each state at each time. **provided** we adjust the number of total states K^*
 - M – number of disease level states. (e.g. SIR $\implies M=3$)
 - N – number of individuals/agents
 - $K^* = MN$
- **Implication:** $M \leq K^* \leq MN$

We do not need D to have CM-AM pairs

- **Theorem 2** (Gallagher and Eddy, 2019a). Any CM with M states has an equivalent AM pair with M states and homogeneous agent interactions, in terms of numbers of individuals in each state at each time.
- **Result 3**. Each AM has an equivalent CM pair **provided** we adjust the number of total states K^* , in terms of numbers of individuals in each state at each time. **provided** we adjust the number of total states K^*
 - M – number of disease level states. (e.g. SIR $\implies M=3$)
 - N – number of individuals/agents
 - $K^* = MN$
- **Implication:** $M \leq K^* \leq MN$

Takeaway: Finding an equivalent CM-AM pair means finding K^* , the number of states **needed** to model the outbreak

As a result of our CM-AM pairs, we suggest the following modeling workflow:

As a result of our CM-AM pairs, we suggest the following modeling workflow:

1. Estimate K^* in the CM-view
 - Select best model with K^* total states
 - Estimate parameters for model
2. Analyze scenarios in AM-view

Estimate K^* for SIR

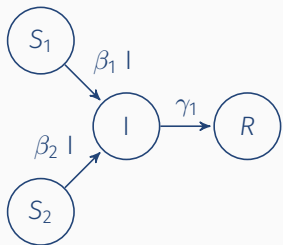
1. Visualize the SIR in a linear-regression framework
2. View all three states simultaneously with a ternary plot
3. Quantify how far “similar” agent-interaction structures are in terms of summary statistics

We turn ternary plots into a model diagnostic

- $S_t + I_t + R_t \equiv N \implies (S_t, I_t, R_t)$ lay in a constrained plane
- Used in Safan 2006 to visualize theoretical equilibria
- We extend the plot to include observations, estimates, confidence regions, and time

Example: Classic SIR - Two S groups

Picture of model with Binomial draws:



Data:

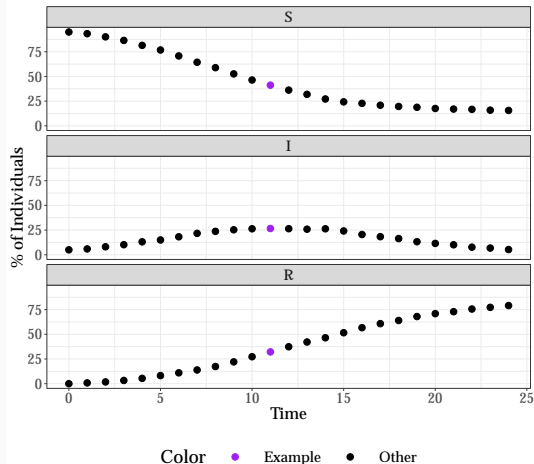
- (s_t, i_t, r_t) for $t = 1, \dots, 100$
- $\beta_1 = 0.8, \beta_2 = 0.3, \gamma = 0.2$

Estimate:

- $E[(S_t, I_t, R_t)]$ for $t = 1, \dots, 100$
- $\hat{\beta} = 0.5, \hat{\gamma} = 0.2$

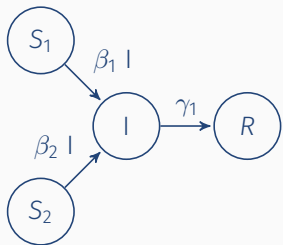
Simulation observations

$\beta_1 = 0.80; \beta_2 = 0.30; \gamma = 0.20$



Example: Classic SIR - Two S groups

Picture of model with Binomial draws:



Data:

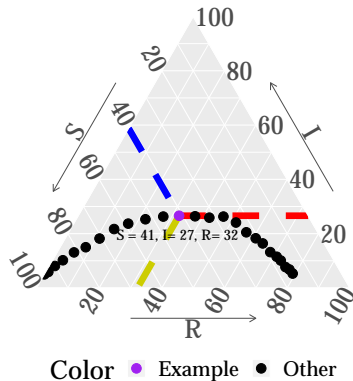
- (s_t, i_t, r_t) for $t = 1, \dots, 100$
- $\beta_1 = 0.8, \beta_2 = 0.3, \gamma = 0.2$

Estimate:

- $E[(S_t, I_t, R_t)]$ for $t = 1, \dots, 100$
- $\hat{\beta} = 0.5, \hat{\gamma} = 0.2$

Simulation observations

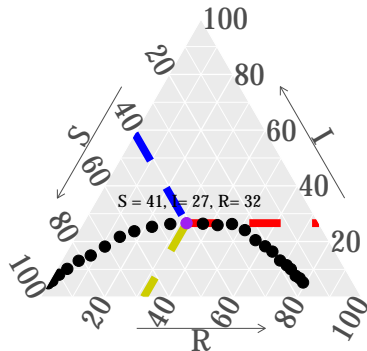
$\beta_1 = 0.80; \beta_2 = 0.30; \gamma = 0.20$



Example: Classic SIR - Two S groups

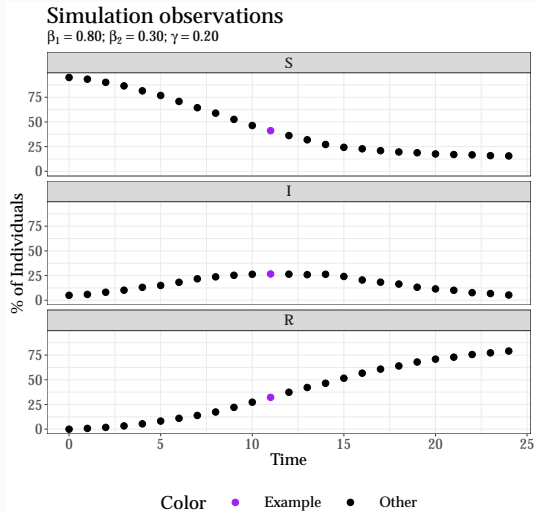
Simulation observations

$$\beta_1 = 0.80; \beta_2 = 0.30; \gamma = 0.20$$

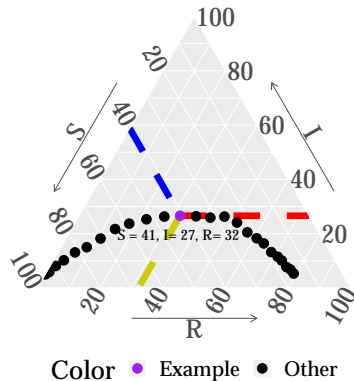


Color ● Example ● Other

Example: Classic SIR - Two S groups



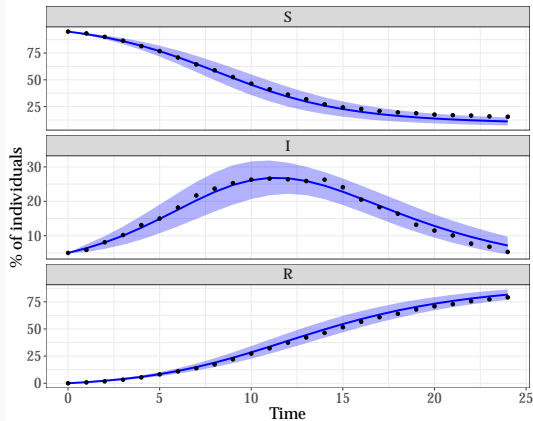
Simulation observations
 $\beta_1 = 0.80$; $\beta_2 = 0.30$; $\gamma = 0.20$



Example: Classic SIR - Two S groups

Simulation observations

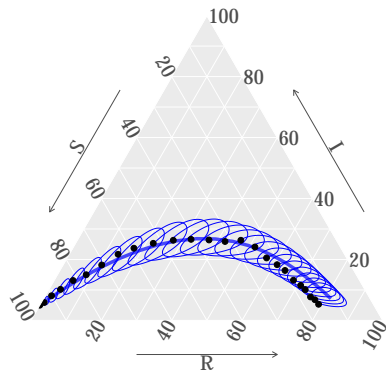
$\beta_1 = 0.80$; $\beta_2 = 0.30$; $\gamma = 0.20$



Model Avg. — Estimate

Obs. and ests. with 95% CR

$\beta_1 = 0.8$; $\beta_2 = 0.3$; $\gamma = 0.2$; $\hat{\beta} = 0.5$, $\hat{\gamma} = 0.2$

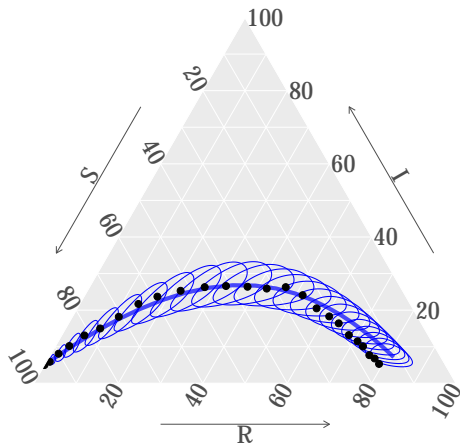


Data \blacktriangle Est. \circ Obs. Type \square Obs. — Est.

Example: Classic SIR - Two Susceptible groups

Obs. and ests. with 95% CR

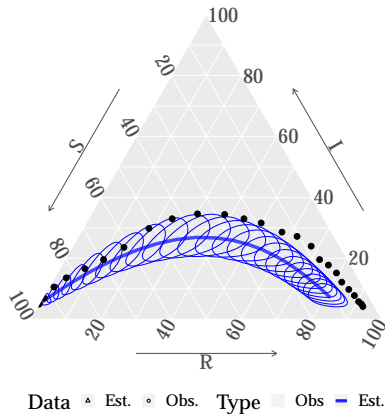
$$\beta_1 = 0.8; \beta_2 = 0.3, \gamma = 0.2; \hat{\beta} = 0.5, \hat{\gamma} = 0.2$$



Example: Classic SIR - Two Susceptible groups

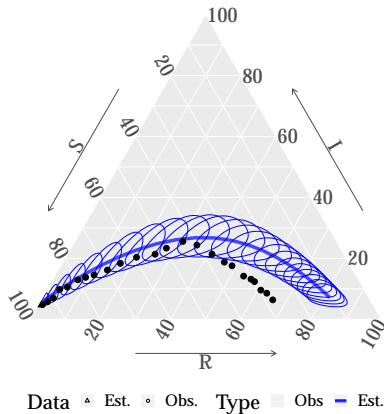
Group 1 obs. and ests. with 95% CR

$\beta_1 = 0.8; \beta_2 = 0.3, \gamma = 0.2; \hat{\beta} = 0.5, \hat{\gamma} = 0.2$



Group 2 obs. and ests. with 95% CR

$\beta_1 = 0.8; \beta_2 = 0.3, \gamma = 0.2; \hat{\beta} = 0.5, \hat{\gamma} = 0.2$



Example: Classic SIR - Two Susceptible groups

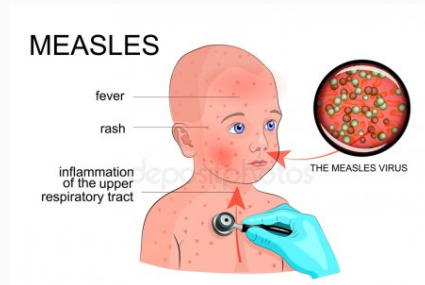
Takeaway:

- With all aspects in the ternary plot, we can assess our model
- Although this is restricted to SIR disease-level states, we can still look at groups of individuals within these states
- There is a possible extension to visualizing the SEIR model in 3D
 - E - “Exposed” state - already infected but not yet infectious

Measles!

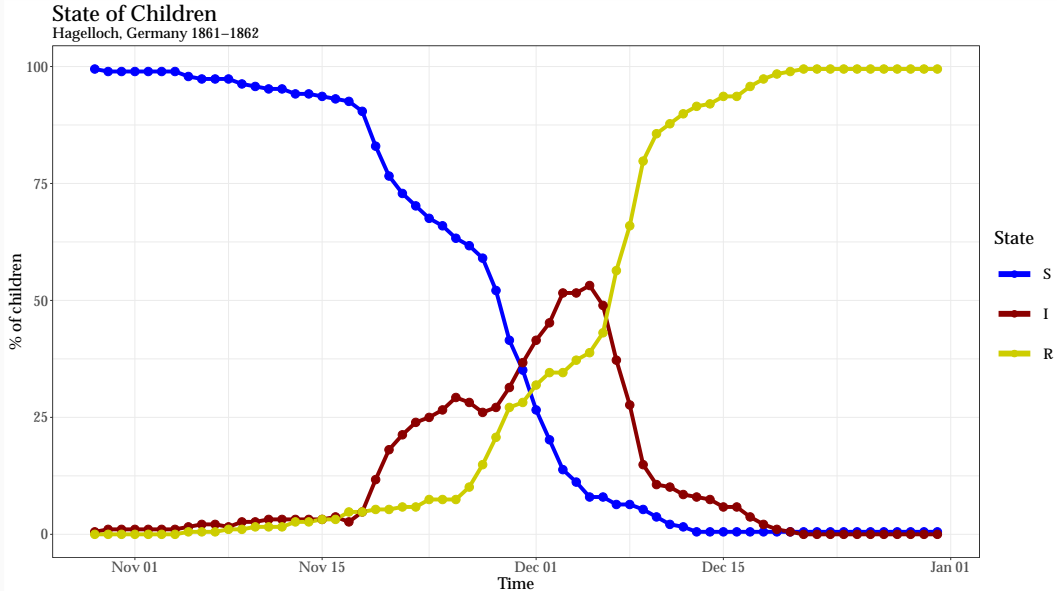
About measles

- Highly infectious childhood disease ($\mathcal{R}_0 = 19$) (Anderson & May, 1992)
 - Influenza $\mathcal{R}_0 \approx 1.2$
- Prodromes – initial symptoms
 - high fever, cough, runny nose, red, watery eyes
 - 2-3 days after, tiny white spots in mouth
- Measles rash and high fever: 3-5 days after symptoms begin
- 2-3 days after rash, child recovers
- CDC reports person is infectious ± 4 days after rash appearance
- Lifelong immunity after infection



ID	Household	Class	Age	Sex	I	R	Infector
1	61	1st	7	F	22	29	45
2	61	1st	6	F	23	32	45
3	62	pre-K	4	F	29	37	NA
4	63	2nd	13	M	27	32	180
5	63	1st	8	F	22	31	45

Measles: SIR curve of outbreak



Measles: questions about Hagelloch

- What is K^* , the minimal number of states?
- What is the associated CM-AM pair?
- What is \mathcal{R}_0 ?
- What would have happened...
 - if we generally reduce the infectivity?
 - if we isolate infectious individuals?
 - if we shut down the school?

Measles: reducing the infectivity

Scenario 1

1. We have estimate(s) of $\hat{\beta}$, the infection parameter
2. Assume we can reduce infectivity to $\rho \cdot \hat{\beta}$
3. How would outbreak have changed?

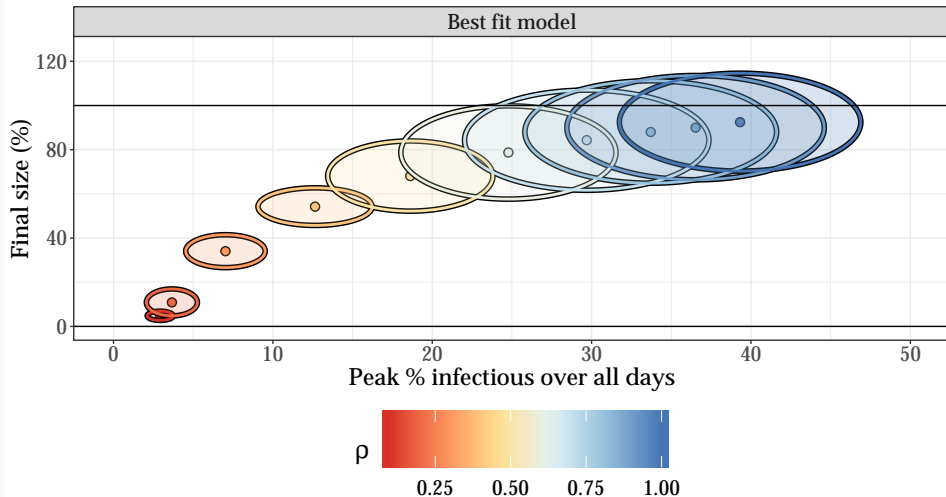
Analysis 1

1. Initialize our CM-AM pair with estimates
2. Vary ρ in our simulations
3. Analyze resulting outbreaks

Measles: reducing the infectivity results

Hagelloch AM simulation results

From $t=0$ onward



Measles: case study summary (see Chs. 6-7)

- What is K^* , the minimal number of states?
 - $K^* = 6$
- What is \mathcal{R}_0 ?
 - Between 4-5.
- What is the associated CM-AM pair?
 - $S^2 I^2 R^2$ with groups before and after $t = 25$
- What would have happened...
 - if we reduced the infectivity of the disease?
 - Want to reduce $\hat{\beta}$ by about half
 - if we isolated infectious individuals?
 - Reduce size of epidemic, even if isolated 8 days after initial infection
 - if we shut down the school?
 - Inconclusive results due to assumptions of model

Conclusions

We improve disease inference through CM-AM pairs

1. Statistically relate CMs and AMs
2. Develop methodology for model selection
3. Apply methodology to measles case study

We improve disease inference through CM-AM pairs

1. Statistically relate CMs and AMs ✓
 - CM-AM pairs with D , a pre-defined transition matrix
 - General CM-AM pairs
 - Importance of K^* , minimum number of states
2. Develop methodology for model selection
3. Apply methodology to measles case study

We improve disease inference through CM-AM pairs

1. Statistically relate CMs and AMs ✓

- CM-AM pairs with D , a pre-defined transition matrix
- General CM-AM pairs
- Importance of K^* , minimum number of states

2. Develop methodology for model selection ✓

- log-linear plot
- ternary plot
- quantifying differences for similar agent-interaction structures

3. Apply methodology to measles case study

We improve disease inference through CM-AM pairs

1. Statistically relate CMs and AMs ✓
 - CM-AM pairs with D , a pre-defined transition matrix
 - General CM-AM pairs
 - Importance of K^* , minimum number of states
2. Develop methodology for model selection ✓
 - log-linear plot
 - ternary plot
 - quantifying differences for similar agent-interaction structures
3. Apply methodology to measles case study ✓
 - Model selection for observed data
 - $\mathcal{R}_0 = 4\text{-}5$
 - Examination of hypothetical scenarios

- **Ebola case study**
 - Western District, Sierra Leone 2014-2015
 - 8,000+ cases
 - Population of ~ 1.4 million
 - SPEW synthetic agents (Gallagher et al. 2018)
 - Focus on sensitivity to initial conditions
 - Importance of N , the effective population size
- **catalyst** - R package with code for all my dissertation work
- SPEW and associated R package **spew** (Richardson et al., 2018)
- “Nine ways to estimate \mathcal{R}_0 in the SIR model” (Gallagher et al. In prep.)
- “Opening up the court (surface) in tennis” (Gallagher et al. In revision)

- What is the bias-variance trade-off for choosing a smaller (larger) K^* ?
- How can we incorporate N as a random variable?
- Explore implementation of different vaccine trials in CM-AM pairs

Hagelloch, Germany 1861 & Pittsburgh, PA 2019



What links the two?

Hagelloch, Germany 1861 & Pittsburgh, PA 2019



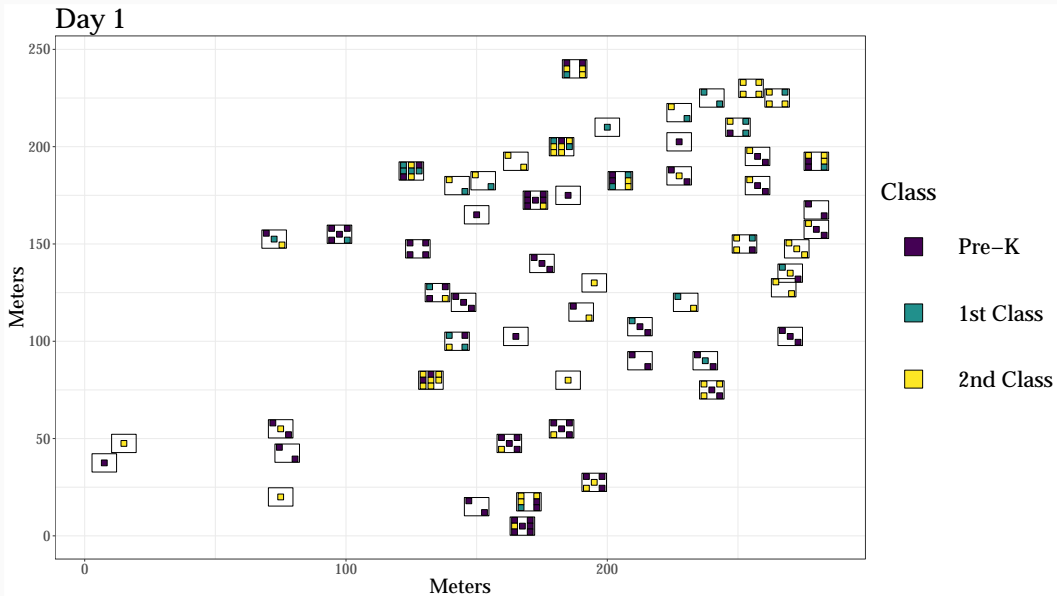
What links the two?

Measles

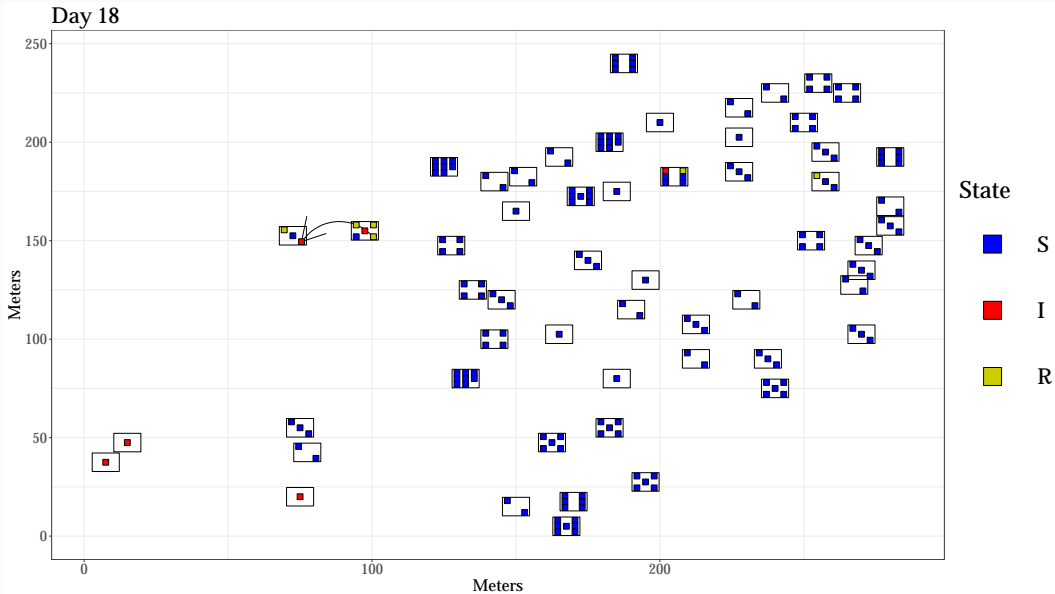
Thank you.

Questions?

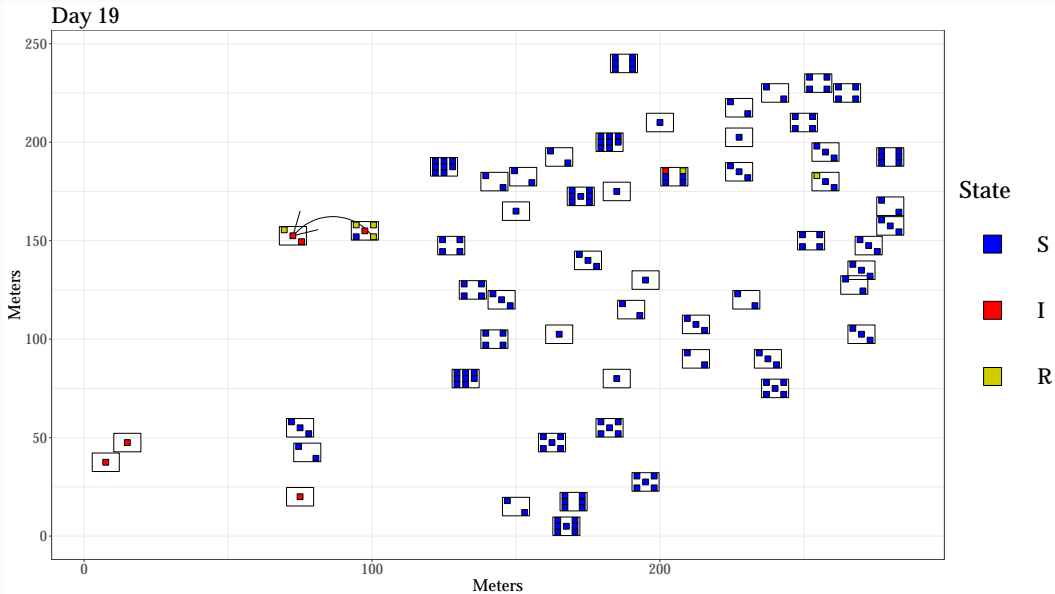
Hagelloch physical locations



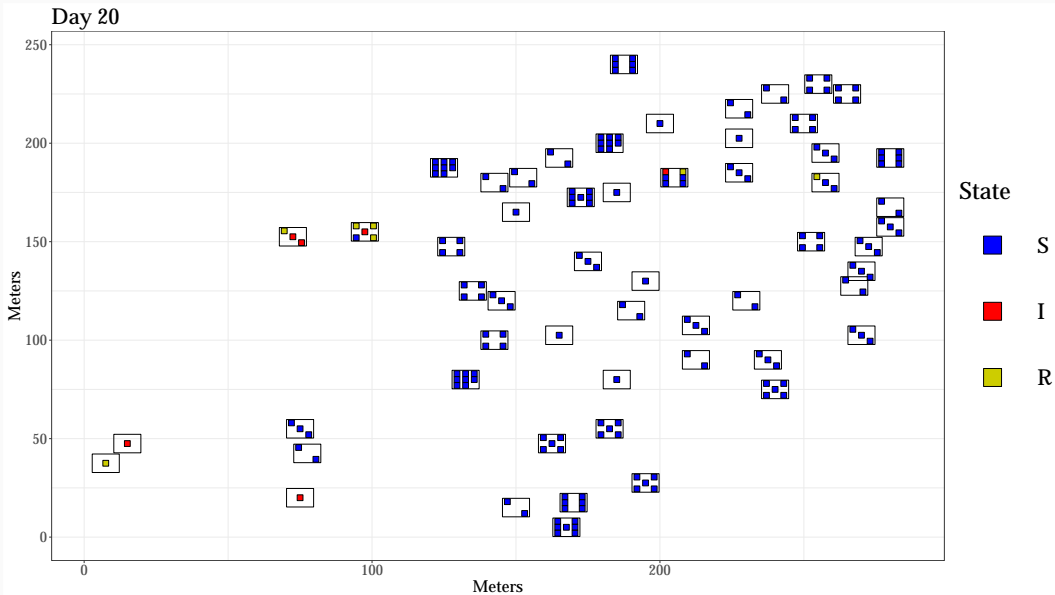
Infection map of Hagelloch, Germany 1861



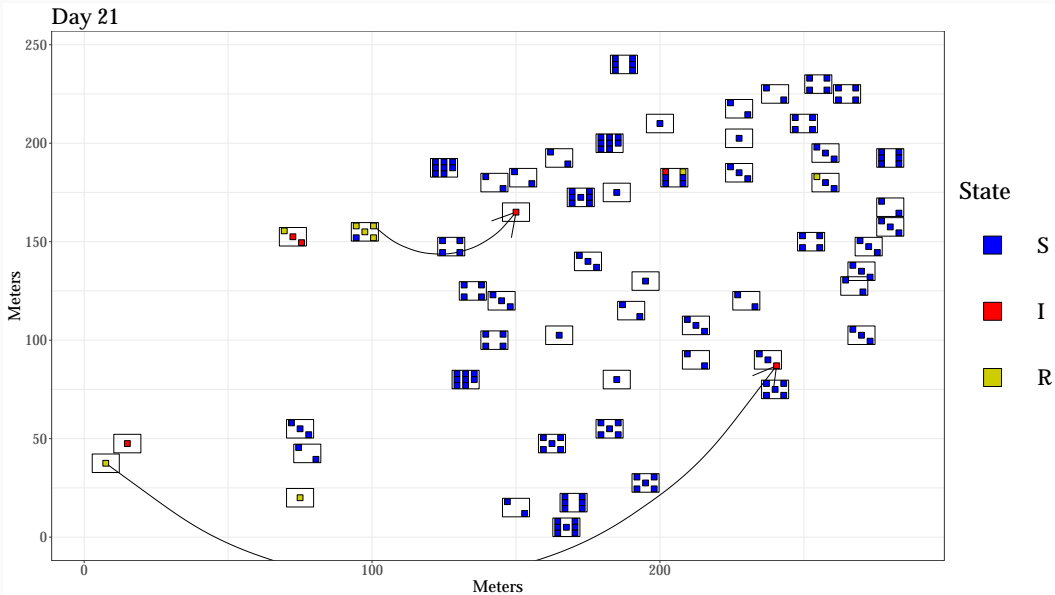
Infection map of Hagelloch, Germany 1861



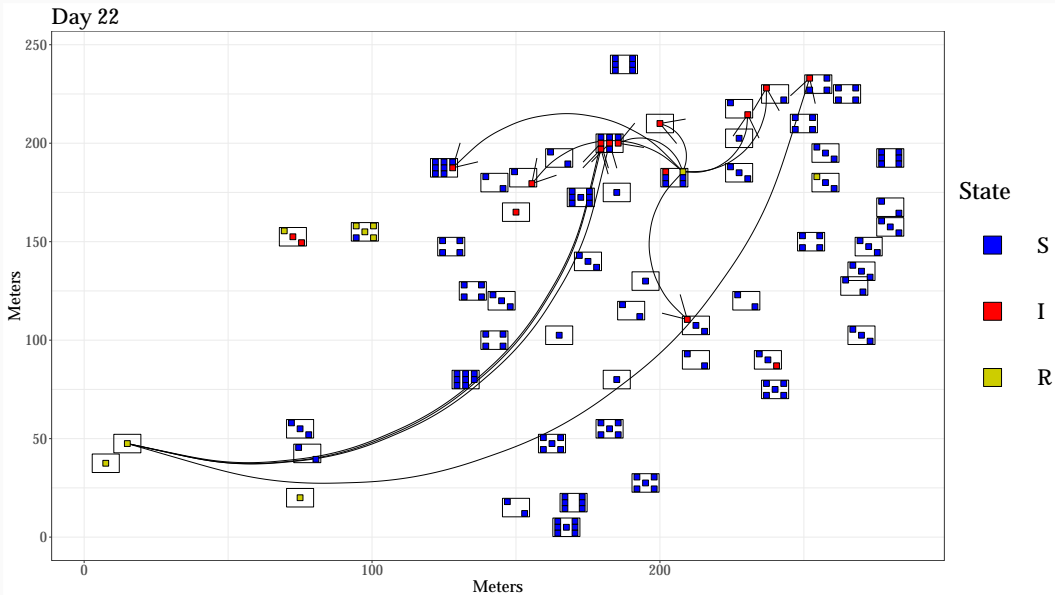
Infection map of Hagelloch, Germany 1861



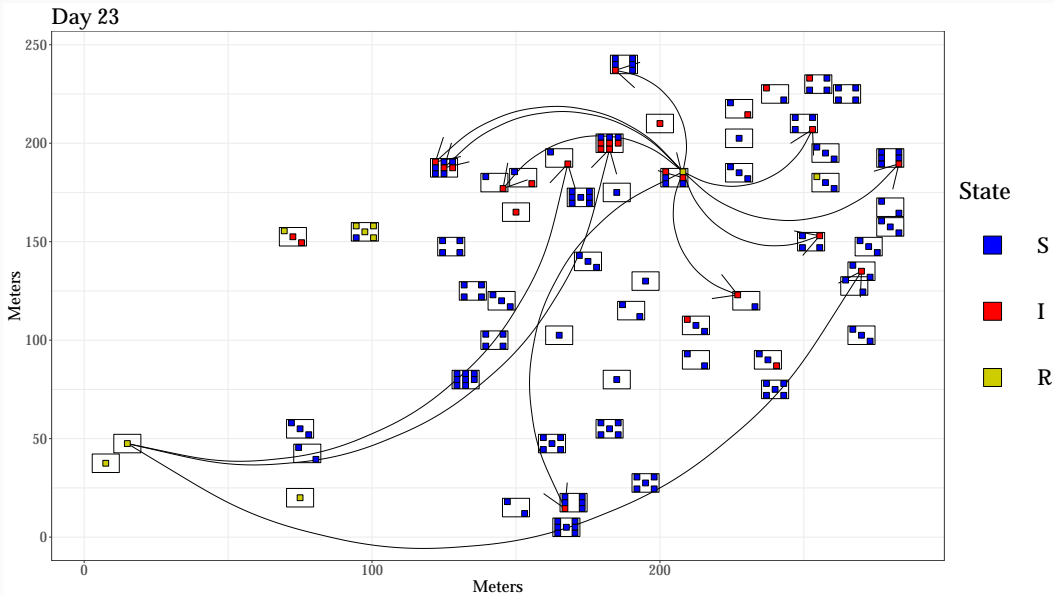
Infection map of Hagelloch, Germany 1861



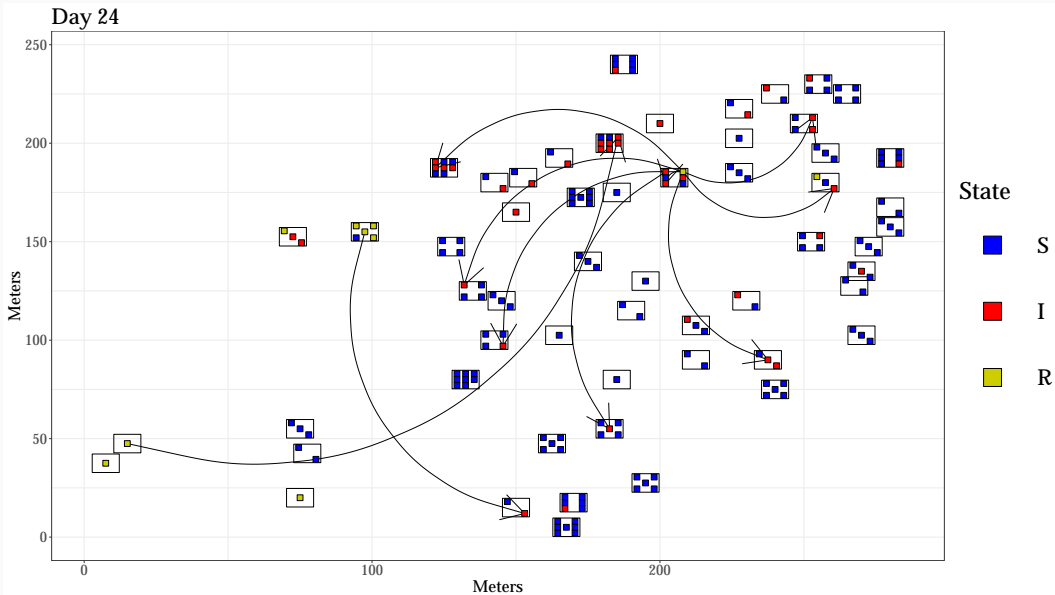
Infection map of Hagelloch, Germany 1861



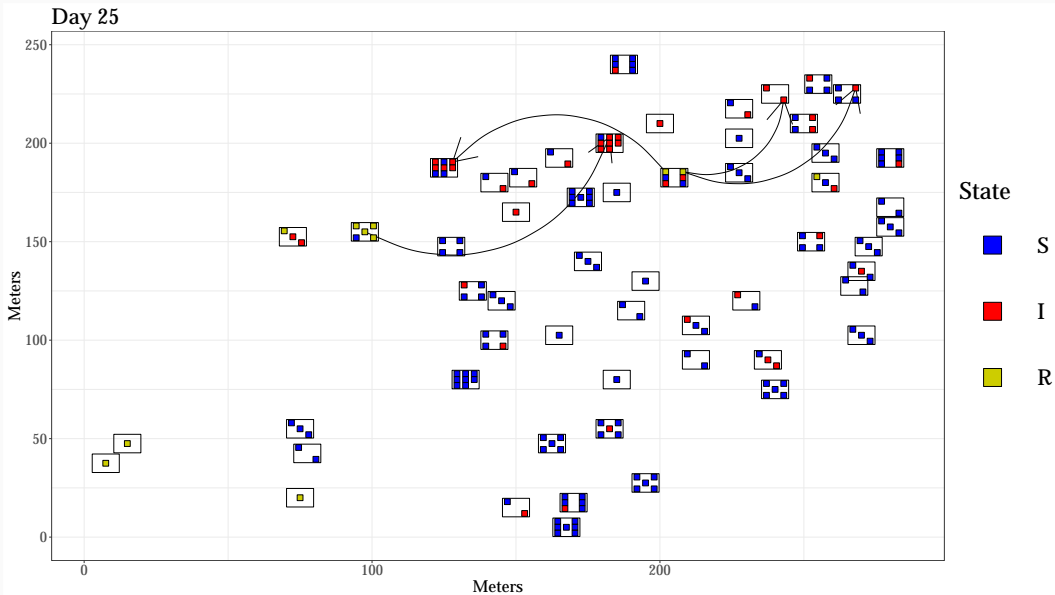
Infection map of Hagelloch, Germany 1861



Infection map of Hagelloch, Germany 1861



Infection map of Hagelloch, Germany 1861



Scenario 2

1. We have best fit estimates for our selected model
2. Isolate infectious individuals after delay period d
3. How would outbreak have changed?

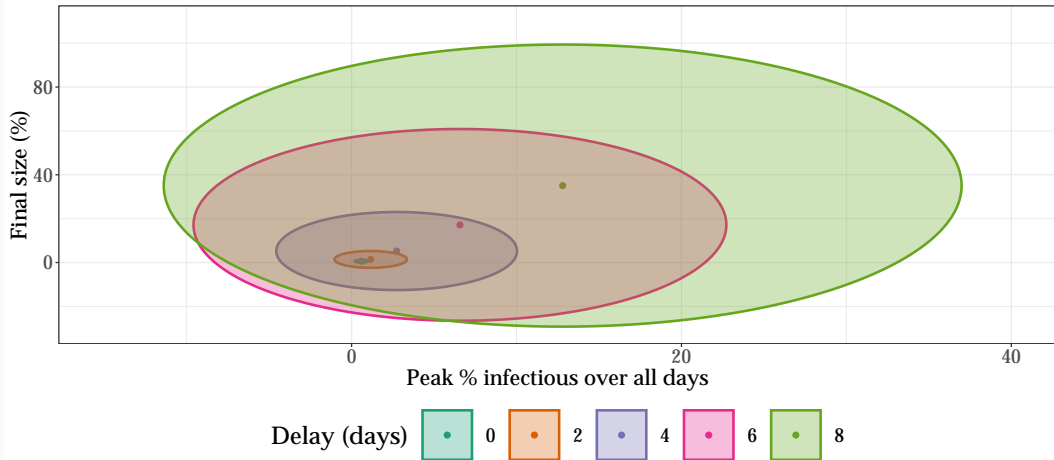
Analysis 2

1. Initialize our CM-AM pair with estimates
2. Vary d in our simulations
3. Analyze resulting outbreaks

Measles: isolation

Isolation results of Hagelloch

Isolation to household



Scenario 3

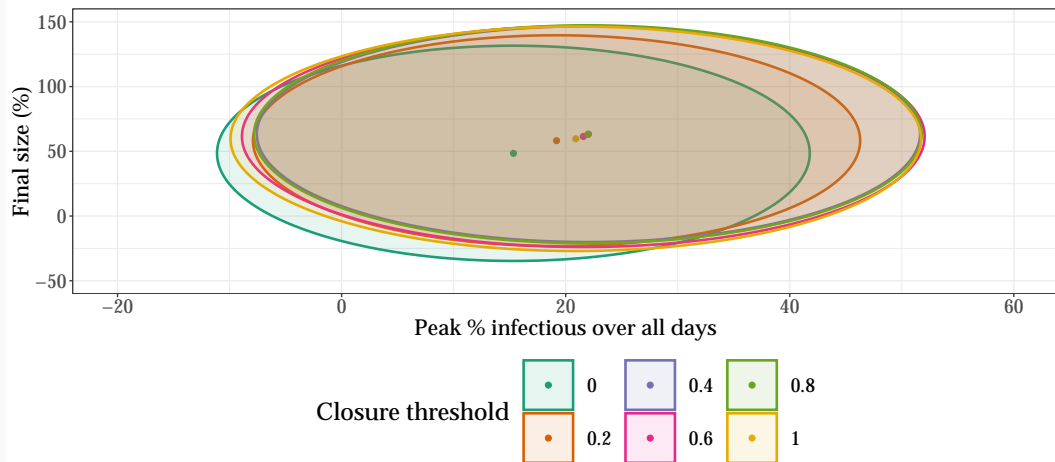
1. We have best fit estimates for our selected model
2. Close school after closure threshold C_s is met
3. How would outbreak have changed?

Analysis 3

1. Initialize our CM-AM pair with estimates
2. Vary C_s in our simulations
3. Analyze resulting outbreaks

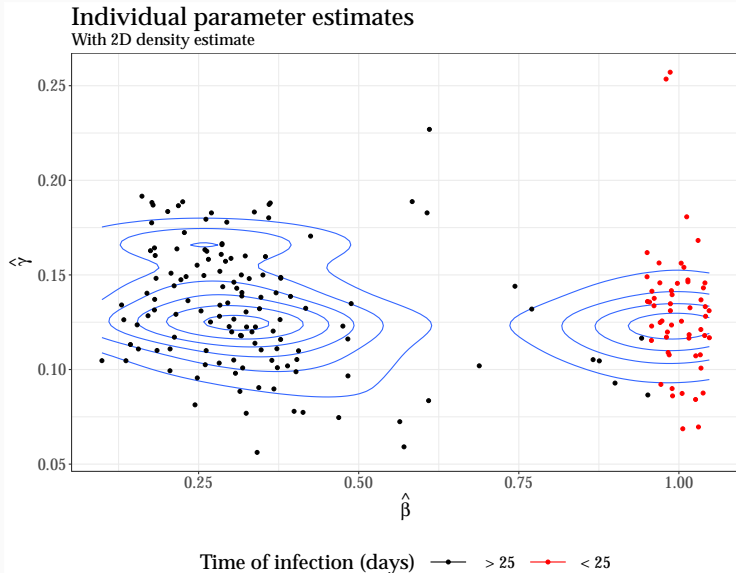
School closure results of Hagelloch children

School closure results



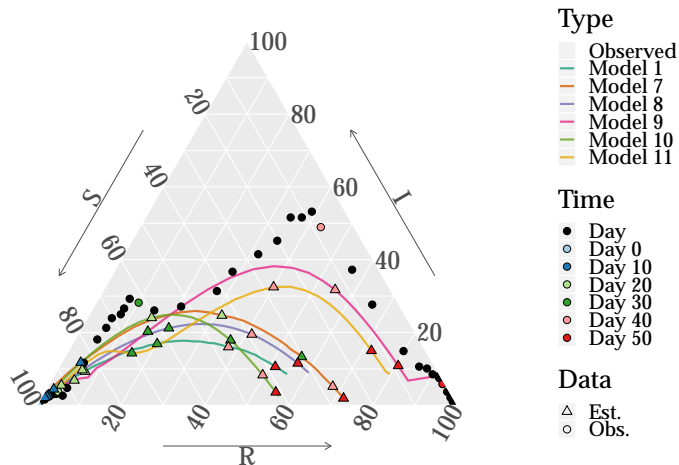
Measles: model selection

- Used EDA, basic clustering to find potential groups
- Fit SIR models with maximum log-likelihood
- Used MSE, AIC, and novel plots to assess fit
- Selected two models with two groups of agents



Observed data and fitted models

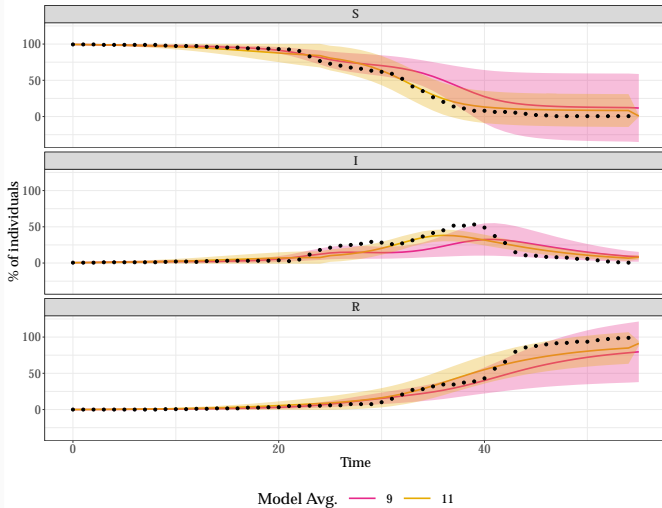
- Used EDA, basic clustering to find potential groups
- Fit SIR models with maximum log-likelihood
- Used MSE, AIC, and novel plots to assess fit
- Selected two models with two groups of agents



Measles: model selection

- Used EDA, basic clustering to find potential groups
- Fit SIR models with maximum log-likelihood
- Used MSE, AIC, and novel plots to assess fit
- Selected two models with two groups of agents

Hagelloch estimates and 95% CI



Measles: the reproduction number \mathcal{R}_0

Previous estimates:

- $\hat{\mathcal{R}}_0 = 17\text{-}19$ (Anderson and May 1992)
- $\hat{\mathcal{R}}_0 = 6\text{-}7$ (Getz 2016)

Our estimate(s):

Model	# Groups	Interaction	Partition	$\mathcal{R}_0^{(1)}$	95% CI	$\mathcal{R}_0^{(2)}$	95% CI
1	1	Homogeneous		4.94	[4.68, 5.21]		
2	2	Homogeneous	$tl > 25$	4.17	[3.89, 4.57]	2.49	[2.37, 2.62]
3	2	Heterogeneous	$tl > 25$	3.13	[2.84, 3.41]	4.35	[4.23, 4.48]

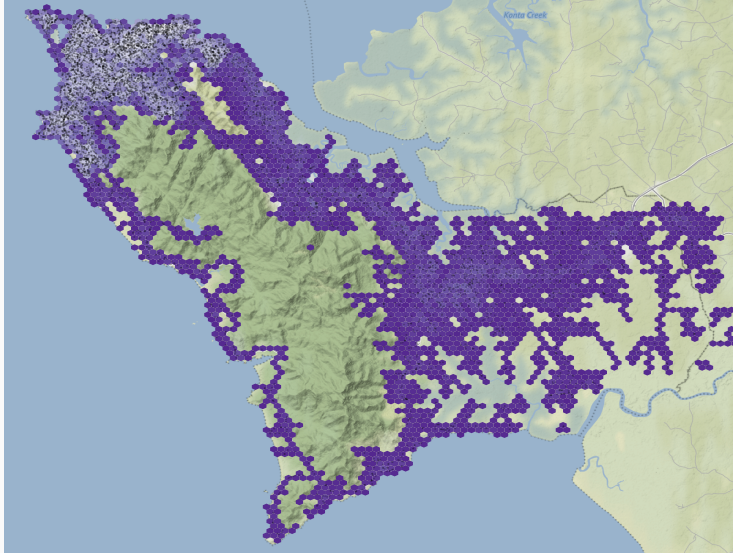
Takeaways:

- $\mathcal{R}_0 \approx 4-5$
- Difference in infectivity before and after day $t = 25$
- Which group is more infectious depends on our assumptions of interaction

Ebola

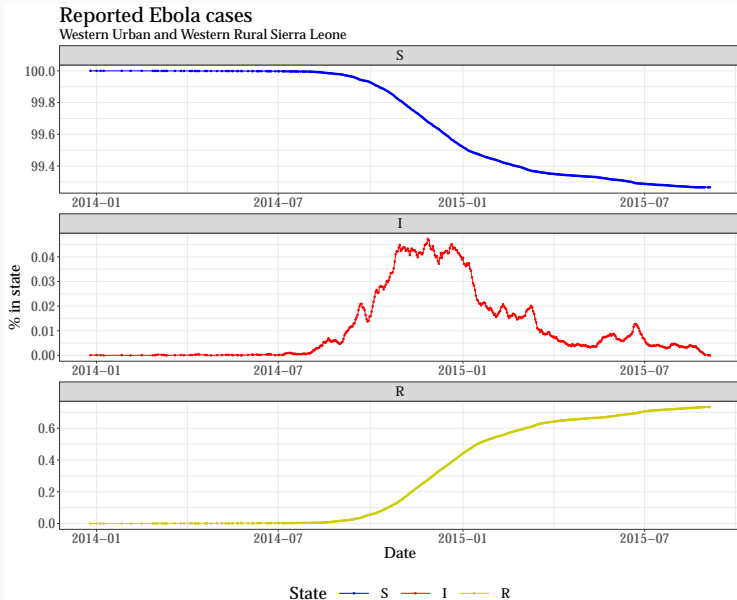
Ebola in Western District, Sierra Leone

Imputed 2014–2015 Ebola infection locations
Western District, Sierra Leone



Ebola

Ebola in Western District, Sierra Leone



Ebola in Western District, Sierra Leone

2014–2015 Ebola outbreak and best fit SIR model (by MSE)

$\beta = 0.16$; $\gamma = 0.12$; $N = 18758$

