A Data-driven Comparison of Plague Models

Senior Project

Luke Mattfeld

Eastern Washington University

Fall, 2020

Outline

The Plague

Which Plague?

- Names: The Black Death, Bubonic Plague, etc.

Figure: "The Triumph of Death" - Pieter Bruegel the Elder - 1562

The Plague

Which Plague?

- Names: The Black Death, Bubonic Plague, etc.

Figure: "The Triumph of Death" - Pieter Bruegel the Elder - 1562

The Plague

Which Plague?

- Names: The Black Death, Bubonic Plague, etc.
    - Bubonic plague

Figure: "The Triumph of Death" - Pieter Bruegel the Elder - 1562

The Plague

Which Plague?

- Names: The Black Death, Bubonic Plague, etc.
    - Bubonic plague
    - Pneumonic plague

Figure: "The Triumph of Death" - Pieter Bruegel the Elder - 1562

The Plague

Which Plague?

- Names: The Black Death, Bubonic Plague, etc.
  - Bubonic plague
  - Pneumonic plague
  - Septicemic plague

Figure: "The Triumph of Death" - Pieter Bruegel the Elder - 1562

The Plague

Which Plague?

- Names: The Black Death, Bubonic Plague, etc.
    - Bubonic plague
    - Pneumonic plague
    - Septicemic plague
- Bacteria behind it all: Yersinia pestis

Figure: "The Triumph of Death" - Pieter Bruegel the Elder - 1562

The Plague

How it Spread

The Plague

How it Spread

- Aspirated

The Plague

How it Spread

- Aspirated $\rightarrow$ Pneumonic model

The Plague

How it Spread

- Aspirated $\rightarrow$ Pneumonic model
- Rats to Fleas to Humans

The Plague

How it Spread

- Aspirated $\rightarrow$ Pneumonic model
- Rats to Fleas to Humans $\rightarrow$ Rat-Flea transmission (RFT) model

The Plague

How it Spread

- Aspirated $\rightarrow$ Pneumonic model
- Rats to Fleas to Humans $\rightarrow$ Rat-Flea transmission (RFT) model
- Other

The Plague

How it Spread

- Aspirated → Pneumonic model
- Rats to Fleas to Humans → Rat-Flea transmission (RFT) model
- Other → Human-Ectoparasite model (Dean et al.)

The Plague

How it Spread

- Aspirated $\rightarrow$ Pneumonic model
- Rats to Fleas to Humans $\rightarrow$ Rat-Flea transmission (RFT) model
- Other $\rightarrow$ Human-Ectoparasite model (Dean et al.)
- New RFT Model

The Plague

How it Spread

- Aspirated → Pneumonic model
- Rats to Fleas to Humans → Rat-Flea transmission (RFT) model
- Other → Human-Ectoparasite model (Dean et al.)
- New RFT Model → Lynch-Oster RFT model

The Plague

How it Spread

- Aspirated → Pneumonic model
- Rats to Fleas to Humans → Rat-Flea transmission (RFT) model
- Other → Human-Ectoparasite model (Dean et al.)
- New RFT Model → Lynch-Oster RFT model

Goal

The Plague

How it Spread

- Aspirated $\rightarrow$ Pneumonic model
- Rats to Fleas to Humans $\rightarrow$ Rat-Flea transmission (RFT) model
- Other $\rightarrow$ Human-Ectoparasite model (Dean et al.)
- New RFT Model $\rightarrow$ Lynch-Oster RFT model

Goal

- Use data on plague spread to compare proposed models

The Plague

How it Spread

- Aspirated → Pneumonic model
- Rats to Fleas to Humans → Rat-Flea transmission (RFT) model
- Other → Human-Ectoparasite model (Dean et al.)
- New RFT Model → Lynch-Oster RFT model

Goal

- Use data on plague spread to compare proposed models
- Get indication of spread type per data

Outline

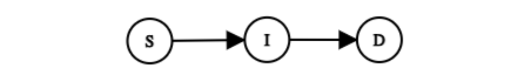
Pneumonic Model

Humans - SID

$$\frac{dS_h}{dt} = -\beta_p \frac{S_h I_h}{N_h}$$

$$\frac{dI_h}{dt} = \beta_p \frac{S_h I_h}{N_h} - \gamma_p I_h$$

$$\frac{dD_h}{dt} = \gamma_p I_h$$



Pneumonic Model

Humans - SID

$$\frac{dS_h}{dt} = -\beta_p \frac{S_h I_h}{N_h}$$

$$\frac{dI_h}{dt} = \beta_p \frac{S_h I_h}{N_h} - \gamma_p I_h$$

$$\frac{dD_h}{dt} = \gamma_p I_h$$

- SID model
- S_h - Susceptible, I_h - Infected, D_h - Deaths, N_h - Total Population
- β_p - Transmission rate, γ^{-1} - Infectious period

Keeling-Gilligan Rat Model

Fleas

$$\frac{dH}{dt} = r_f H \left(1 - \frac{H}{K_f}\right)$$

$$\frac{dF}{dt} = (1 - g_r)\gamma_r I_r H - d_f F$$

- H - Number of fleas per rat
- F - Number of infected fleas not on rats

Keeling-Gilligan Rat Model

Fleas

$$\frac{dH}{dt} = r_f H\left(1 - \frac{H}{K_f}\right)$$

$$\frac{dF}{dt} = (1 - g_r)\gamma_r I_r H - d_f F$$

- H - Number of fleas per rat
- F - Number of infected fleas not on rats

Keeling-Gilligan RFT Model

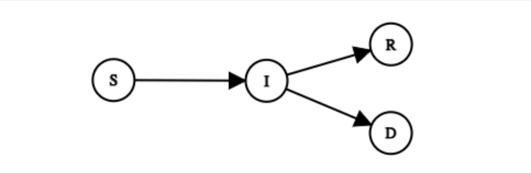
SIRD - Rats & Humans

$$\frac{dS_j}{dt} = -\beta_r \frac{S_j F}{N_j}\left[1 - e^{-aN_r}\right]$$

$$\frac{dI_j}{dt} = \beta_r \frac{S_j F}{N_j}\left[1 - e^{-aN_r}\right] - \gamma_j I_j$$

$$\frac{dR_j}{dt} = g_j \gamma_j I_j$$

$$\frac{dD_j}{dt} = (1 - g_j)\gamma_j I_j$$



Keeling-Gilligan RFT Model

SIRD - Rats & Humans

$$\frac{dS_j}{dt} = -\beta_r \frac{S_j F}{N_j}\left[1 - e^{-aN_r}\right]$$

$$\frac{dI_j}{dt} = \beta_r \frac{S_j F}{N_j}\left[1 - e^{-aN_r}\right] - \gamma_j I_j$$

$$\frac{dR_j}{dt} = g_j \gamma_j I_j$$

$$\frac{dD_j}{dt} = (1 - g_j)\gamma_j I_j$$

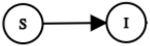
- SIRD model where j = Rats, Fleas
- S_j - Susceptible, I_j - Infected, R_j - Recovered, D_j - Dead, N_j - Total Population

Human-Ectoparasite Model

Parasites - SI

$$\frac{dS_L}{dt} = r_L S_L \left(1 - \frac{N_L}{K_L}\right) - \left[(\beta_{low} I_{low} + \beta_{high} I_{high}) \frac{S_L}{N_h}\right]$$

$$\frac{dI_L}{dt} = \left[(\beta_{low} I_{low} + \beta_{high} I_{high}) \frac{S_L}{N_h}\right] - \gamma_L I_L$$



Human-Ectoparasite Model

Parasites - SI

$$\frac{dS_L}{dt} = r_L S_L \left(1 - \frac{N_L}{K_L}\right) - \left[\left(\beta_{low} I_{low} + \beta_{high} I_{high}\right) \frac{S_L}{N_h}\right]$$

$$\frac{dI_L}{dt} = \left[\left(\beta_{low} I_{low} + \beta_{high} I_{high}\right) \frac{S_L}{N_h}\right] - \gamma_L I_L$$

- SI Model
- S_L - Susceptible, I_L - Infected
- γ_L^{-1} - Avg. infectious period, r_L - Intrinsic growth rate, K_L - Lice carrying capacity

Human-Ectoparasite Model

Humans - SIIRD

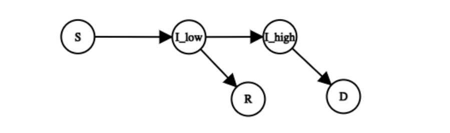
$$\frac{dS_h}{dt} = -\beta_L \frac{S_h I_L}{N_h}$$

$$\frac{dI_{low}}{dt} = \beta_L \frac{S_h I_L}{N_h} - \sigma_b I_{low}$$

$$\frac{dI_{high}}{dt} = (1 - g_h)\sigma_b I_{low} - \gamma_b I_{high}$$

$$\frac{dR_h}{dt} = g_h \sigma_b I_{low}$$

$$\frac{dD_h}{dt} = \gamma_b I_{high}$$



Human-Ectoparasite Model

Humans - SIIRD

$$\frac{dS_h}{dt} = -\beta_L \frac{S_h I_L}{N_h}$$

$$\frac{dI_{low}}{dt} = \beta_L \frac{S_h I_L}{N_h} - \sigma_b I_{low}$$

$$\frac{dI_{high}}{dt} = (1 - g_h)\sigma_b I_{low} - \gamma_b I_{high}$$

$$\frac{dR_h}{dt} = g_h \sigma_b I_{low}$$

$$\frac{dD_h}{dt} = \gamma_b I_{high}$$

- $SI_L I_h RD$ Model
- S_h - Susceptible, I_{low} - Infected (low level), I_{high} - Infected (high level), R_h - Recovered, D_h - Dead

Lynch-Oster RFT Model

Rats & Fleas - Logistic

$$\frac{dR_T}{dt} = (\frac{\beta_R}{K_R})R_T(K_R - R_T) - \delta R_c$$

$$\frac{dR_c}{dt} = \alpha\frac{F_c}{F_T}(R_T - R_c) - \frac{\beta_R}{K_R}(R_T)(R_c) - \delta R_c - \gamma R_c$$

$$\frac{dF_T}{dt} = (\frac{\beta_F}{K_F})F_T(K_F - F_T) - \rho F_T$$

$$\frac{dF_c}{dt} = \lambda\frac{R_c}{R_T}(F_T - F_c) - \rho F_c$$

- Logistic Model

Lynch-Oster RFT Model

Rats & Fleas - Logistic

$$\frac{dR_T}{dt} = (\frac{\beta_R}{K_R})R_T(K_R - R_T) - \delta R_c$$

$$\frac{dR_c}{dt} = \alpha\frac{F_c}{F_T}(R_T - R_c) - \frac{\beta_R}{K_R}(R_T)(R_c) - \delta R_c - \gamma R_c$$

$$\frac{dF_T}{dt} = (\frac{\beta_F}{K_F})F_T(K_F - F_T) - \rho F_T$$

$$\frac{dF_c}{dt} = \lambda\frac{R_c}{R_T}(F_T - F_c) - \rho F_c$$

- Logistic Model
- T - total, c - Infected
- β - Intrinsic birth rate, μ - Intrinsic death rate, γ - Rat recovery rate, $\{\rho, \delta\}$ - Plague death rate, $\{\lambda, \alpha\}$ - Plague infectivity, K - Carrying capacity

Lynch-Oster RFT Model

Rats & Fleas - Logistic

$$\frac{dR_T}{dt} = (\frac{\beta_R}{K_R})R_T(K_R - R_T) - \delta R_c$$

$$\frac{dR_c}{dt} = \alpha\frac{F_c}{F_T}(R_T - R_c) - \frac{\beta_R}{K_R}(R_T)(R_c) - \delta R_c - \gamma R_c$$

$$\frac{dF_T}{dt} = (\frac{\beta_F}{K_F})F_T(K_F - F_T) - \rho F_T$$

$$\frac{dF_c}{dt} = \lambda\frac{R_c}{R_T}(F_T - F_c) - \rho F_c$$

- Logistic Model
- T - total, c - Infected
- β - Intrinsic birth rate, μ - Intrinsic death rate, γ - Rat recovery rate, $\{\rho, \delta\}$ - Plague death rate, $\{\lambda, \alpha\}$ - Plague infectivity, K - Carrying capacity

Lynch-Oster Rat Model

Humans - SEIR

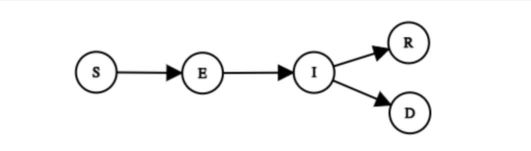
$$\frac{dS}{dt} = \beta(S + R_b) - \sigma S \frac{F_c}{F_T} - \mu S$$

$$\frac{dE}{dt} = \sigma S \frac{F_c}{F_T} - \nu E - \mu E$$

$$\frac{dI}{dt} = \nu E - \phi I - rI$$

$$\frac{dR_b}{dt} = rI - \mu R_b$$

$$\frac{dD}{dt} = \phi I - \mu N_h$$



Lynch-Oster Rat Model

Humans - SEIR

$$\frac{dS}{dt} = \beta(S + R_b) - \sigma S \frac{F_c}{F_T} - \mu S$$

$$\frac{dE}{dt} = \sigma S \frac{F_c}{F_T} - \nu E - \mu E$$

$$\frac{dI}{dt} = \nu E - \phi I - rI$$

$$\frac{dR_b}{dt} = rI - \mu R_b$$

$$\frac{dD}{dt} = \phi I - \mu N_h$$

- SEIR model
- S - Susceptible, E - Infected, I - Infected, R_b - Recovered, D - Deaths, N_h - Total Population
- β - Human birth rate, σ - chance of being infected by flea, μ - Intrinsic death rate, ν^{-1} - Incubation period, r^{-1} - Recovery rate

Outline

Method

How to Compare

Method

How to Compare

- Given data and a model, why not just run directly?

Method

How to Compare

- Given data and a model, why not just run directly?
- Problem: Unknown Parameters

Method

How to Compare

- Given data and a model, why not just run directly?
- Problem: Unknown Parameters
    - Lack of data

Method

How to Compare

- Given data and a model, why not just run directly?
- Problem: Unknown Parameters
  - Lack of data
  - No way to measure

Method

How to Compare

- Given data and a model, why not just run directly?
- Problem: Unknown Parameters
  - Lack of data
  - No way to measure
  - Nontrivial to compute

Method

How to Compare

- Given data and a model, why not just run directly?
- Problem: Unknown Parameters
  - Lack of data
  - No way to measure
  - Nontrivial to compute
  - Changes from case to case

Method

How to Compare

- Given data and a model, why not just run directly?
- Problem: Unknown Parameters
  - Lack of data
  - No way to measure
  - Nontrivial to compute
  - Changes from case to case
- Solution: Markov-Chain Monte-Carlo

Method

Monte Carlo Method

Method

Monte Carlo Method

- Have: Unknown distribution from behavior

Method

Monte Carlo Method

- Have: Unknown distribution from behavior
- Want: Distribution information

Method

Monte Carlo Method

- Have: Unknown distribution from behavior
- Want: Distribution information
  - 1. Simulate the underlying behavior

Method

Monte Carlo Method

- Have: Unknown distribution from behavior
- Want: Distribution information
  - 1. Simulate the underlying behavior
  - 2. Run lots of simulations

Method

Monte Carlo Method

- Have: Unknown distribution from behavior
- Want: Distribution information
  - 1. Simulate the underlying behavior
  - 2. Run lots of simulations
  - 3. Examine results as number of iterations N $\rightarrow$ Really Big

Method

Monte Carlo Method

- Have: Unknown distribution from behavior
- Want: Distribution information
  - 1. Simulate the underlying behavior
  - 2. Run lots of simulations
  - 3. Examine results as number of iterations N $\rightarrow$ Really Big
- Get: Pretty good distribution estimation

Method

Markov Chains

Method

Markov Chains

- Series (chain) of computed values

Method

Markov Chains

- Series (chain) of computed values
- Dependent on the previous state

Method

Markov Chains

- Series (chain) of computed values
- Dependent on the previous state
- Many uses

Method

Markov Chains

- Series (chain) of computed values
- Dependent on the previous state
- Many uses
    - Model a state machine

Method

Markov Chains

- Series (chain) of computed values
- Dependent on the previous state
- Many uses
  - Model a state machine
  - Train into predictive algorithms

Method

Markov Chains

- Series (chain) of computed values
- Dependent on the previous state
- Many uses
  - Model a state machine
  - Train into predictive algorithms
  - Explore parameter distribution space

Method

Markov Chains

- Series (chain) of computed values
- Dependent on the previous state
- Many uses
    - Model a state machine
    - Train into predictive algorithms
    - Explore parameter distribution space

Method

MCMC

Method

MCMC

- Monte Carlo simulation generated using Markov Chains

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data
  - $\mathcal{M}$ - model

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data
  - $\mathcal{M}$ - model
- Want to find: $P(\vec{\alpha}|\mathcal{D},\mathcal{M})$

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data
  - $\mathcal{M}$ - model
- Make use of Bayes Formula:

$$P(\vec{\alpha}|\mathcal{D},\mathcal{M}) = \frac{P(\mathcal{D}|\vec{\alpha},\mathcal{M})P(\vec{\alpha}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}$$

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data
  - $\mathcal{M}$ - model
- Make use of Bayes Formula:

$$P(\vec{\alpha}|\mathcal{D},\mathcal{M}) = \frac{P(\mathcal{D}|\vec{\alpha},\mathcal{M})P(\vec{\alpha}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}$$

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data
  - $\mathcal{M}$ - model
- Make use of Bayes Formula:

$$P(\vec{\alpha}|\mathcal{D},\mathcal{M}) = \frac{P(\mathcal{D}|\vec{\alpha},\mathcal{M})P(\vec{\alpha}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}$$

- Posterior - $P(\vec{\alpha}|\mathcal{D},\mathcal{M})$
- Likelihood - $P(\mathcal{D}|\vec{\alpha},\mathcal{M})$
- Prior - $P(\vec{\alpha}|\mathcal{M})$

$$P(\vec{\alpha}|\mathcal{D},\mathcal{M}) \propto P(\mathcal{D}|\vec{\alpha},\mathcal{M})P(\vec{\alpha}|\mathcal{M})$$

Method

MCMC

- Monte Carlo simulation generated using Markov Chains
- Bayesian Statistics
- Given:
  - $\vec{\alpha}$ - vector of unknown parameters
  - $\mathcal{D}$ - data
  - $\mathcal{M}$ - model
- Make use of Bayes Formula:

$$P(\vec{\alpha}|\mathcal{D},\mathcal{M}) = \frac{P(\mathcal{D}|\vec{\alpha},\mathcal{M})P(\vec{\alpha}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}$$

- Posterior - $P(\vec{\alpha}|\mathcal{D},\mathcal{M})$
- Likelihood - $P(\mathcal{D}|\vec{\alpha},\mathcal{M})$
- Prior - $P(\vec{\alpha}|\mathcal{M})$

$$P(\vec{\alpha}|\mathcal{D},\mathcal{M}) \propto P(\mathcal{D}|\vec{\alpha},\mathcal{M})P(\vec{\alpha}|\mathcal{M})$$

Method

MCMC Process

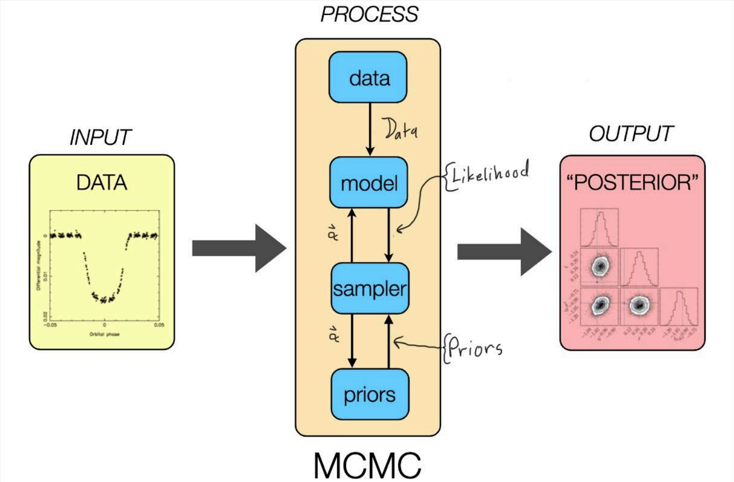
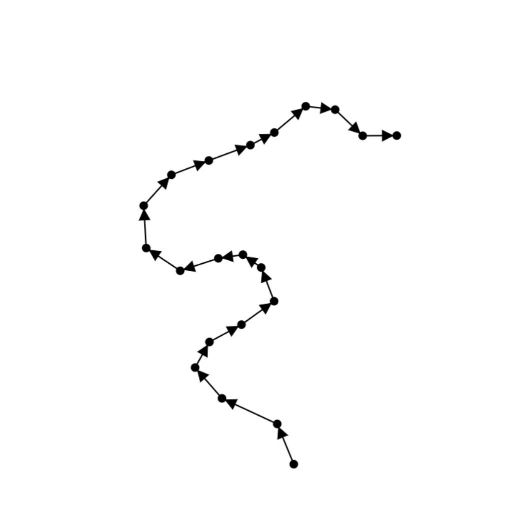


Figure: David Kipping - Sagan 2016 Presentation on MCMC

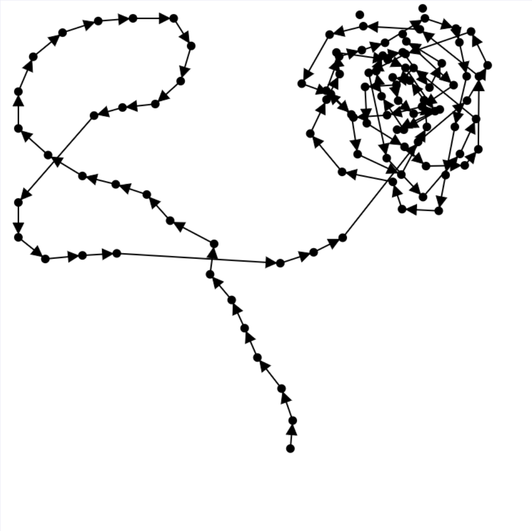
Method

MCMC - Metropolis



Method

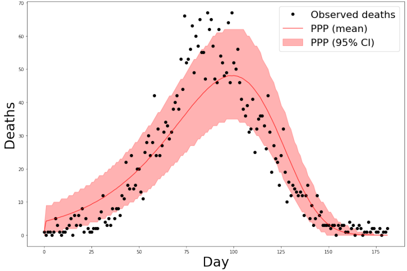
Markov Chains - Metropolis



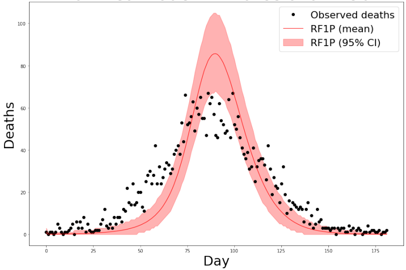
Outline

Barcelona - 1490

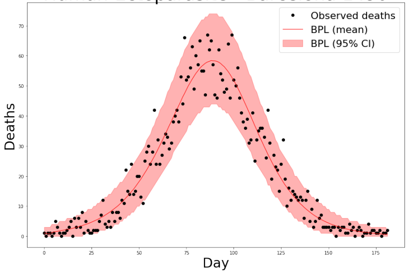
Pneumonic - Barcelona 1490



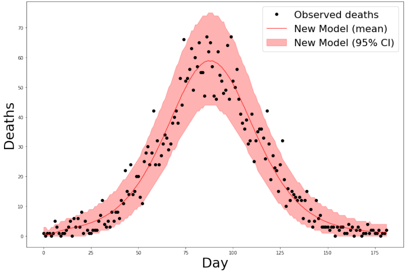
Rat-Flea Model 1 - Barcelona 1490



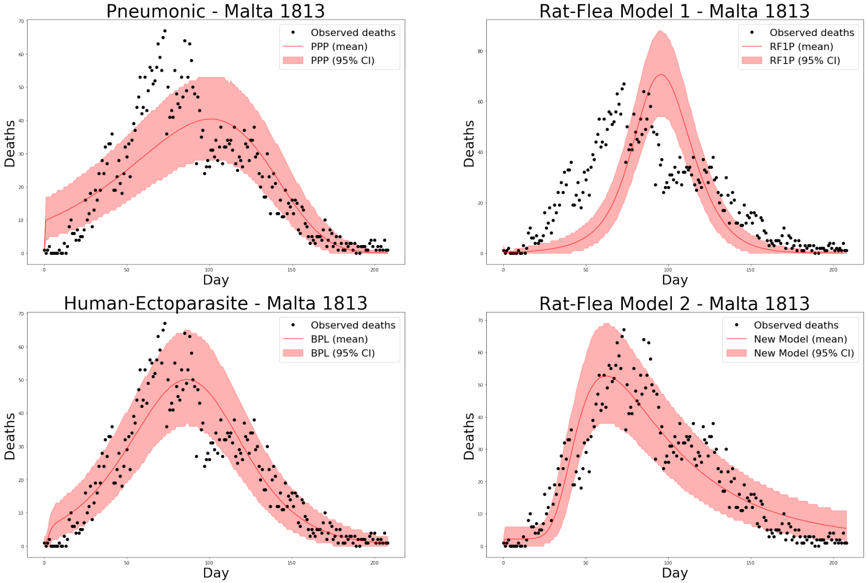
Human-Ectoparasite - Barcelona 1490



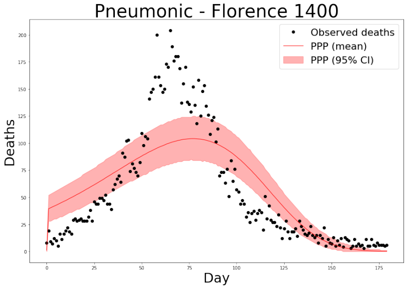
Rat-Flea Model 2 - Barcelona 1490

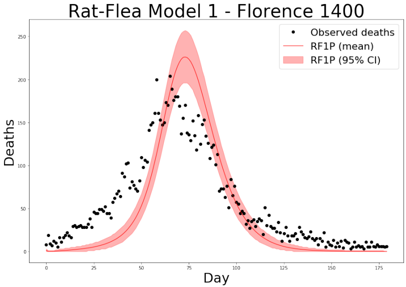


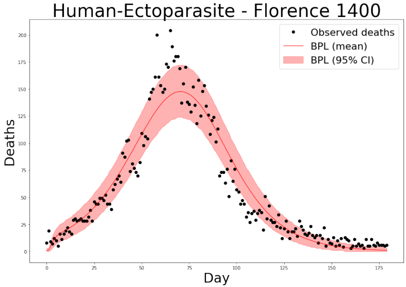
Malta - 1813

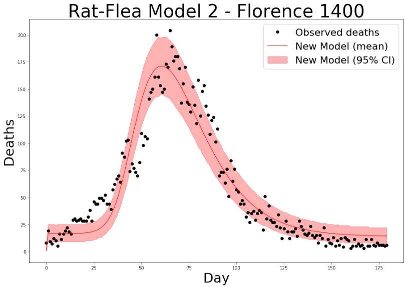

Pneumonic - Malta 1813
Observed deaths
PPP (mean)
PPP (95% CI)
Rat-Flea Model 1 - Malta 1813
Observed deaths
RF1P (mean)
RF1P (95% CI)
Human-Ectoparasite - Malta 1813
Observed deaths
BPL (mean)
BPL (95% CI)
Rat-Flea Model 2 - Malta 1813
Observed deaths
New Model (mean)
New Model (95% CI)
Deaths
Day

Florence - 1400









Metrics

BIC

Data set	Model	BIC
Barcelona	Human-Ecto	1945
	Rat-flea 2	2002
	Pneumonic	2411
	Rat-flea 1	3392
Malta	Human-Ecto	1945
	Rat-flea 2	2491
	Pneumonic	3806
	Rat-flea 1	8274
Florence	Rat-flea 2	2375
	Human-Ecto	6105
	Pneumonic	4660
	Rat-flea 1	2

Metrics

RMSE

Data set	Model	RMSE
Barcelona	Rat-flea 2	4.8
	Human-Ecto	4.9
	Pneumonic	8.1
	Rat-flea 1	10.6
Malta	Human-Ecto	7.4
	Rat-flea 2	7.8
	Pneumonic	10.0
	Rat-flea 1	17.6
Florence	Human-Ecto	15.6
	Rat-flea 2	16.9
	Pneumonic	31.3
	Rat-flea 1	32.7

Metrics

Conclusion

- Neither model outperforms significantly

Metrics

Conclusion

- Neither model outperforms significantly
- Longer testing, more data, better MCMC alg.

Metrics

Ranking

- Neither model outperforms significantly
- Longer testing, more data, better MCMC alg.
- A RFT Model is viable

Outline

Future Work

Ideas

Future Work

Ideas

- Update libraries:

Future Work

Ideas

- Update libraries:
  - Pymc3

Future Work

Ideas

- Update libraries:
    - Pymc3
    - NUTS algorithm

Future Work

Ideas

- Update libraries:
  - Pymc3
  - NUTS algorithm
- Create a framework:
  - Generalize comparison process

Future Work

Ideas

- Update libraries:
  - Pymc3
  - NUTS algorithm
- Create a framework:
  - Generalize comparison process
  - Use on other historical data
  - Use on new outbreaks (COVID)

Q & A

Questions?