Accounting for Immunodynamics in Epidemiological Models

Anton Camacho & Bernard Cazelles

UMR 7625, UPMC- -ENS

Paris, France

Accounting for Immunodynamics in Epidemiological Models

- Introduction
- Classical Epidemiological Models
- Stochastic Epidemiological Models
- Validity of the *van Kampen* Approximation
- Model Inference: Likelihood-based Inference
- Explaining rapid reinfections in multiple-wave influenza outbreaks
- Accounting for immunodynamics in epidemiological model of human flu

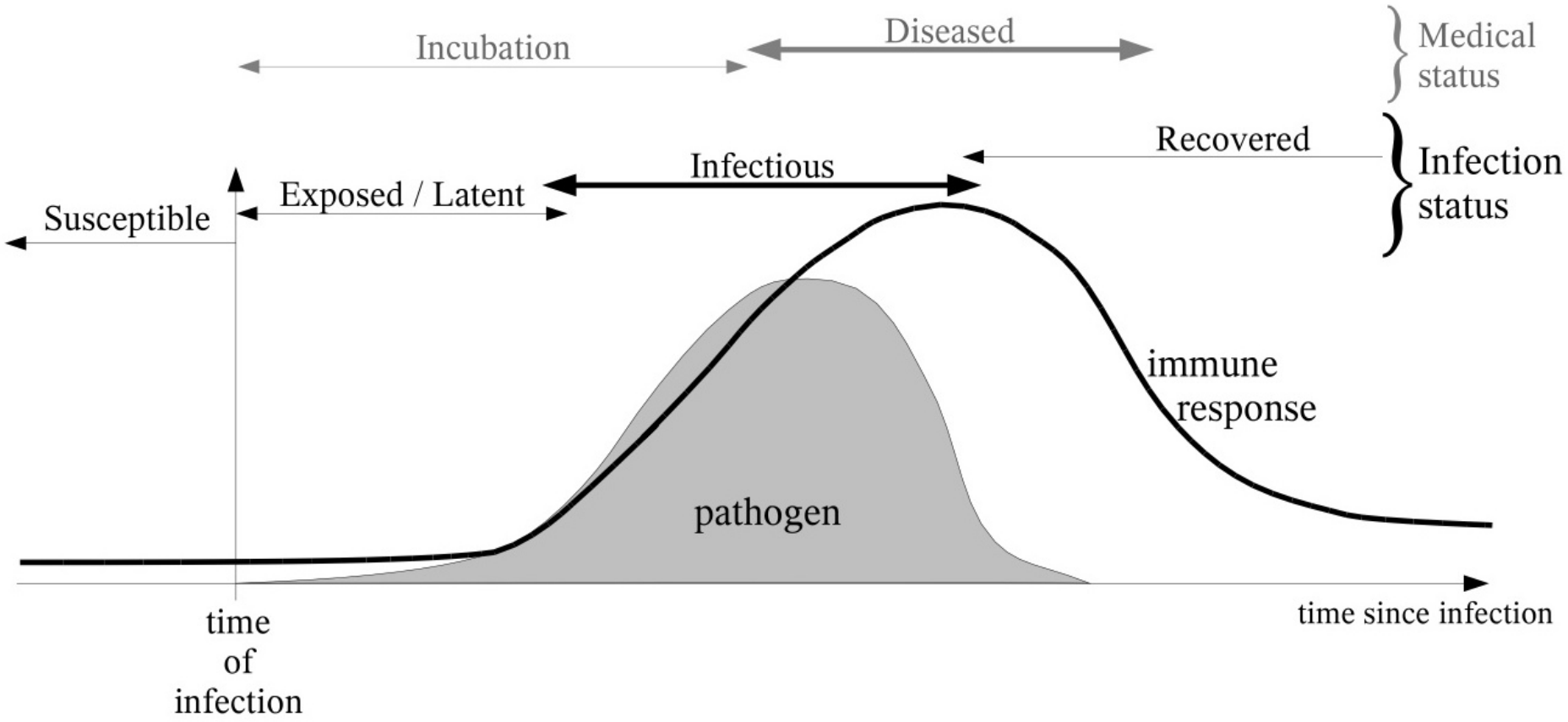
The need of mathematical models, inference and model selection tools in epidemiology

- Epidemiological model since 1927
- Partially observed and noisy data
- Accounting for the characteristics of the dynamics: non linearity and non stationarity
- More and more need of predictive tools in Public Health
- One needs adapted tools for parameter inference, to test hypothesis and to make model selection

Classical Epidemiological Models

Classical Epidemiological Models

- Classical models are known as SIR models and are based on the immune status of the population



Classical Epidemiological Models

- Classical models are known as SIR models and are based on the immune status of the population

Transmission density dependent

$$\frac{dS}{dt} = -\beta . S I$$

$$\frac{dI}{dt} = \beta . S I - \gamma . I$$

$$\frac{dR}{dt} = \gamma . I$$

$$N = S + I + R$$

$$R_0 = \frac{\beta . N}{\gamma}$$

Classical Epidemiological Models

- Classical models are known as SIR models and are based on the immune status of the population

Transmission frequency dependent

$$\frac{dS}{dt} = -\beta . S . \frac{I}{N}$$

$$\frac{dI}{dt} = \beta . S . \frac{I}{N} - \gamma . I$$

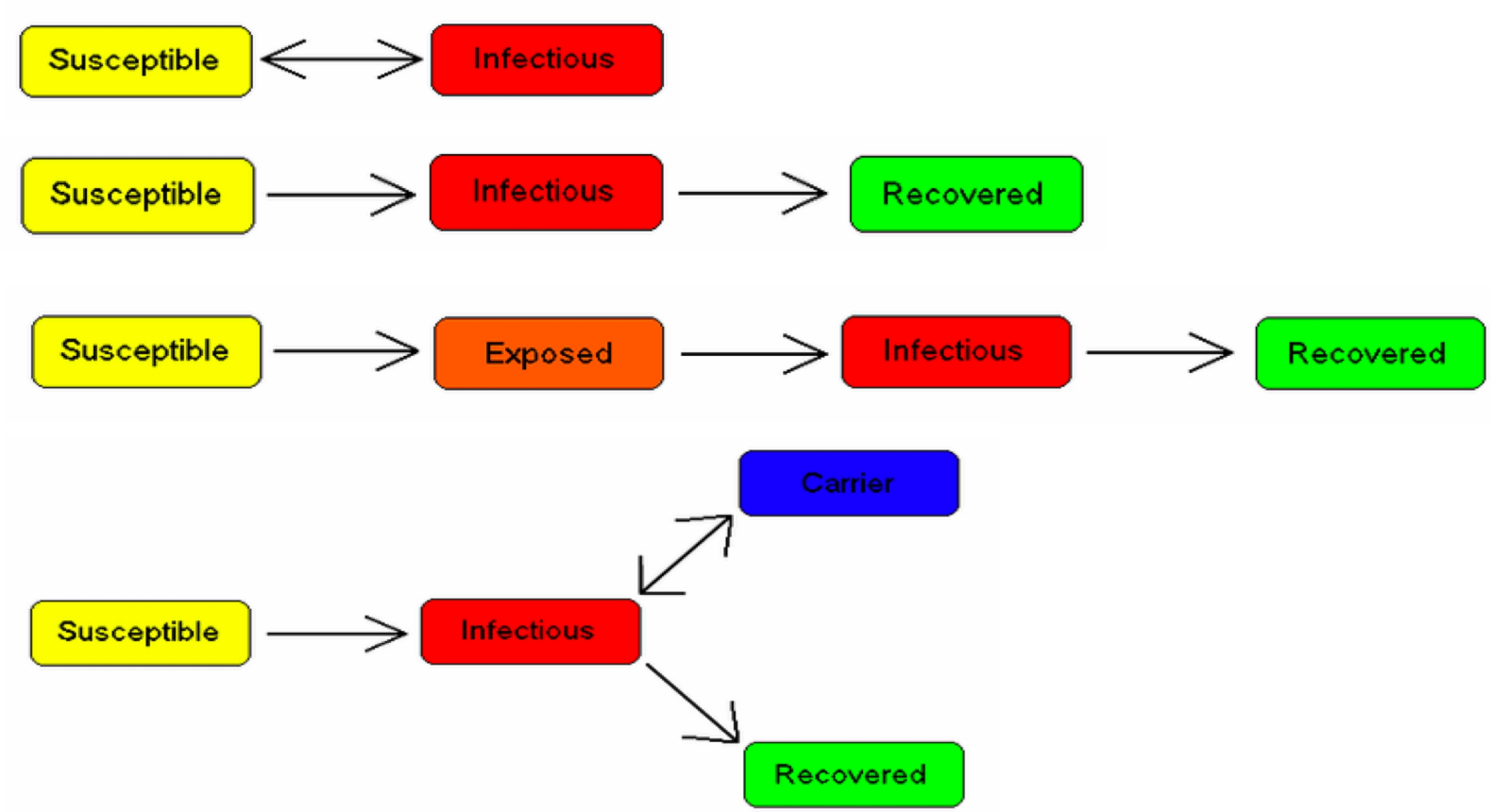
$$\frac{dR}{dt} = \gamma . I$$

$$N = S + I + R$$

$$R_0 = \frac{\beta}{\gamma}$$

Classical Epidemiological Models

- Based on this simple concept different models are possible:



Stochastic Epidemiological Models

We define a possible state of the system as a triplet {*S, I, R*} of random variables. The evolution of the state of the system is led by two transition events:

- Infection of a susceptible by an infective $\{s,i\} \rightarrow \{s-1, i+1\}$
- Removal of an infective $\{s,i\} \rightarrow \{s, i-1\}$

Each event is associated with a transition rate:

- For infection $T(s-1, i+1 \mid s,i) = \beta . \frac{i}{N} . s$
- For recovery and immunity $T(s, i-1 \mid s,i) = \gamma . i$

Stochastic Epidemiological Models

We introduce $P_{s,i}(t)$ as the probability to be in the state $\delta = \{s, i\}$. The evolution of the probability distribution $P_{s,i}(t)$ over the state space $E = \left\{ \{s,i\} \in N^2, s+i \le N \right\}$ is governed by the general **master equation**:

$$\frac{dP_{s,i}(t)}{dt} = \beta \frac{(i-1)}{N}(s+1)P_{s+1,i-1}(t) + \gamma(i+1)P_{s,i+1}(t) - \left(\beta \frac{i}{N} s + \gamma i \right) P_{s,i}(t)$$

$$\frac{dP_t}{dt} = Q.P_t$$

where Q is called the transition matrix (size nxn, n=E is the number of possible states) and P_t is a vector of size N that contains the probabilities for all the states at time t. For our SIR model, n = ½ (N+1)(N+2).

Stochastic Epidemiological Models

To solve the non-linear stochastic equations of the stochastic epidemiological models different methods can be used:

- Monte-Carlo methods: The Gillespie method
- Van Kampen approximation of the master equation (Kurtz approximation for mathematician!)

Stochastic Epidemiological Models

Van Kampen approximation of the master equation

- Make the assumption that *P* depends on *N* and the random variables are rewritten as the sum of a macroscopic deterministic variable and a mesoscopic random variable:

$$S = N\phi(t) + \sqrt{N}\eta_1$$

$$I = N\theta(t) + \sqrt{N}\eta_2$$

- The objective of this approximation is then to extract, from the master equation, the deterministic evolution of $\phi(t)$ and $\theta(t)$ and the probability distribution of η_1 and η_2.
- One obtains a system of ordinary differential equations governing the deterministic variables $\phi(t)$ and $\theta(t)$:

$$\frac{\partial \phi}{\partial t} = -\beta\phi(t)\theta(t)$$

$$\frac{\partial \theta}{\partial t} = \beta\phi(t)\theta(t) - \gamma\theta(t)$$

Stochastic Epidemiological Models

- Van Kampen approximation of the master equation
  - One obtains a system of ordinary differential equations governing the deterministic variables $\phi(t)$ and $\theta(t)$:

$$\frac{\partial \phi}{\partial t} = -\beta \phi(t)\theta(t)$$

$$\frac{\partial \theta}{\partial t} = \beta \phi(t)\theta(t) - \gamma \theta(t)$$

  - And a Fokker-Planck equation for the probability distribution of η_1 and η_2:

$$\frac{\partial P_{\eta_1,\eta_2}}{\partial t} = -\sum_{i,j=1}^{2} A_{i,j} \frac{\partial (\eta_j P_{\eta_1,\eta_2})}{\partial \eta_j} + \frac{1}{2} \sum_{i,j=1}^{2} B_{i,j} \frac{\partial^2 P_{\eta_1,\eta_2}}{\partial \eta_i \partial \eta_j}$$

Stochastic Epidemiological Models

- Van Kampen approximation of the master equation
  - And a Fokker-Planck equation on the probability distribution of η_1 and η_2:

$$\frac{\partial P_{\eta_1,\eta_2}}{\partial t} = -\sum_{i,j=1}^{2} A_{i,j} \frac{\partial(\eta_j P_{\eta_1,\eta_2})}{\partial \eta_j} + \frac{1}{2}\sum_{i,j=1}^{2} B_{i,j} \frac{\partial^2 P_{\eta_1,\eta_2}}{\partial \eta_i \partial \eta_j}$$

  - Then we are able to compute the two first moments of this distribution:

$$\frac{\partial E\left[\eta_1^2\right]}{\partial t} = -2\beta\theta(t)E\left[\eta_1^2\right] - 2\beta\phi(t)E\left[\eta_1\eta_2\right] + \beta\phi(t)\theta(t)$$

$$\frac{\partial E\left[\eta_2^2\right]}{\partial t} = 2(\beta\phi(t) - \gamma)E\left[\eta_2^2\right] + 2\beta\theta(t)E\left[\eta_1\eta_2\right] + \beta\phi(t)\theta(t) + \gamma\theta(t)$$

$$\frac{\partial E\left[\eta_1\eta_2\right]}{\partial t} = \beta\theta(t)E\left[\eta_1^2\right] - \beta\phi(t)E\left[\eta_2^2\right] + (\beta(\phi(t) - \theta(t)) - \gamma)E\left[\eta_1\eta\right] - \beta\phi(t)\theta(t)$$

Stochastic Epidemiological Models

- Van Kampen approximation of the master equation
  - One has to solve a system of ordinary differential equations for the deterministic variables $\phi(t)$ and $\theta(t)$ and for the evolution of the moments of the distribution of their fluctuations:

$$\frac{\partial \phi}{\partial t} = -\beta\phi(t)\theta(t)$$

$$\frac{\partial \theta}{\partial t} = \beta\phi(t)\theta(t) - \gamma\theta(t)$$

$$\frac{\partial E\left[\eta_1^2\right]}{\partial t} = -2\beta\theta(t)E\left[\eta_1^2\right] - 2\beta\phi(t)E\left[\eta_1\eta_2\right] + \beta\phi(t)\theta(t)$$

$$\frac{\partial E\left[\eta_2^2\right]}{\partial t} = 2(\beta\phi(t) - \gamma)E\left[\eta_2^2\right] + 2\beta\theta(t)E\left[\eta_1\eta_2\right] + \beta\phi(t)\theta(t) + \gamma\theta(t)$$

$$\frac{\partial E\left[\eta_1\eta_2\right]}{\partial t} = \beta\theta(t)E\left[\eta_1^2\right] - \beta\phi(t)E\left[\eta_2^2\right] + (\beta(\phi(t) - \theta(t)) - \gamma)E\left[\eta_1\eta\right] - \beta\phi(t)\theta(t)$$

Stochastic Epidemiological Models

- To validate the Van Kampen approximation, we have used a numerical integration of the master equation:

$$\frac{dP_t}{dt} = Q.P_t$$

  - This equation is linear in *P* then:

$$P_t = \exp\left(Qt\right)P_0$$

  - **Expokit*** has been used to numerically solve this equation. The essential advantage of this algorithm lies in the use of Krylov basis that permits the computation without stocking in memory the matrix transition Q.
  - Then, one has a numerical estimation of $P_{s,i}(t)$ for each state $\delta = \{s, i\}$.

*Sidje, R., 1998. Expokit: a software package for computing matrix exponentials. ACM Transactions on Mathematical Software (TOMS).

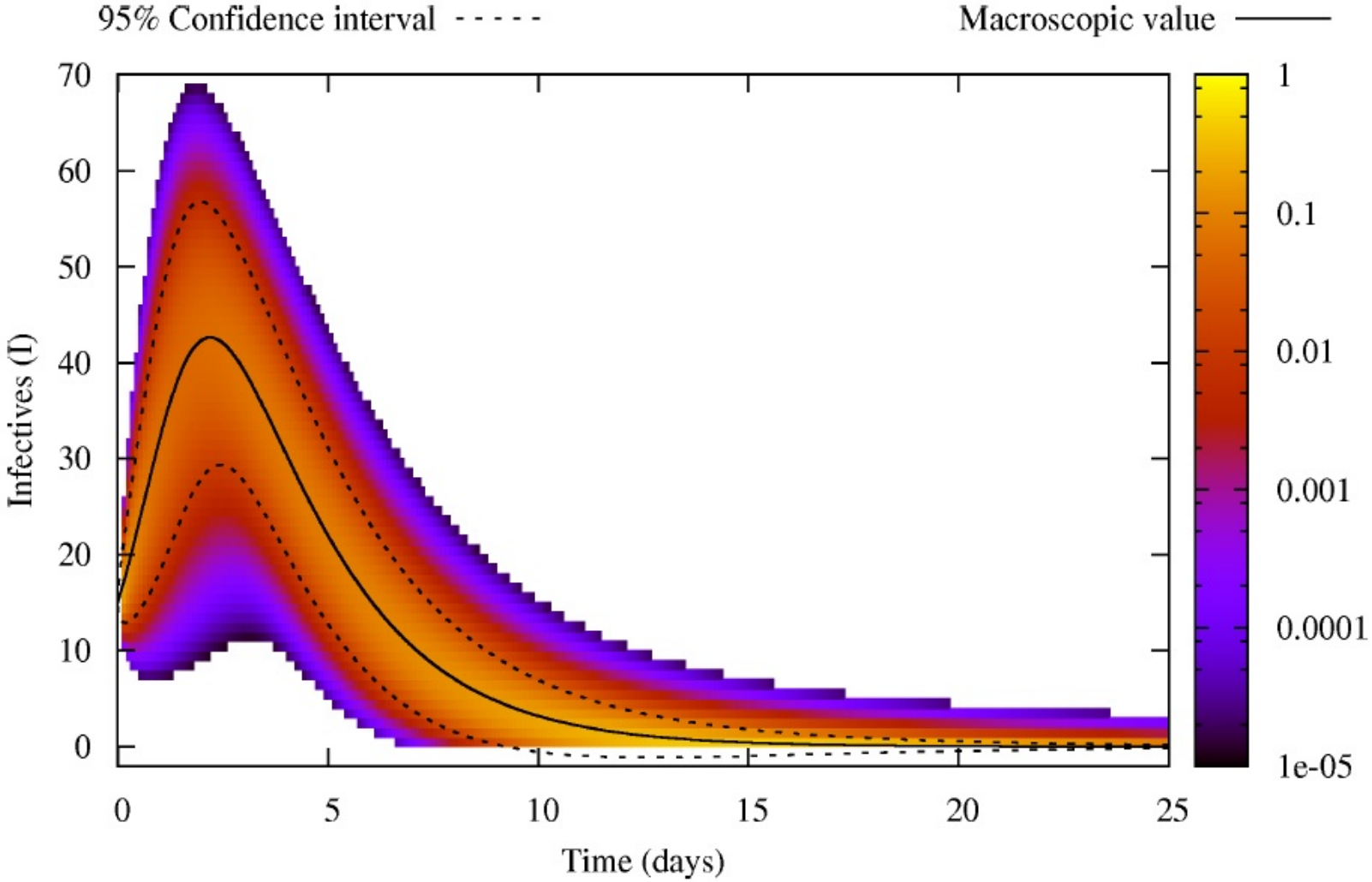
Validation of the *van Kampen* Approximation

$\beta = 1.66$

$\gamma = 0.44$

$N = 100$

$i_0 = 15$



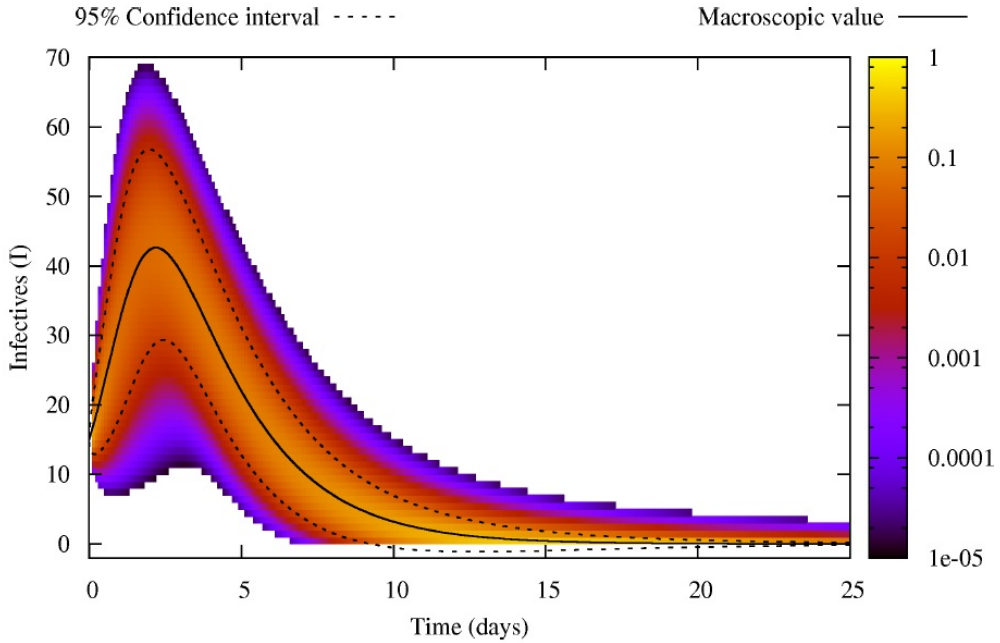
Validation of the *van Kampen* Approximation

$\beta = 1.66$

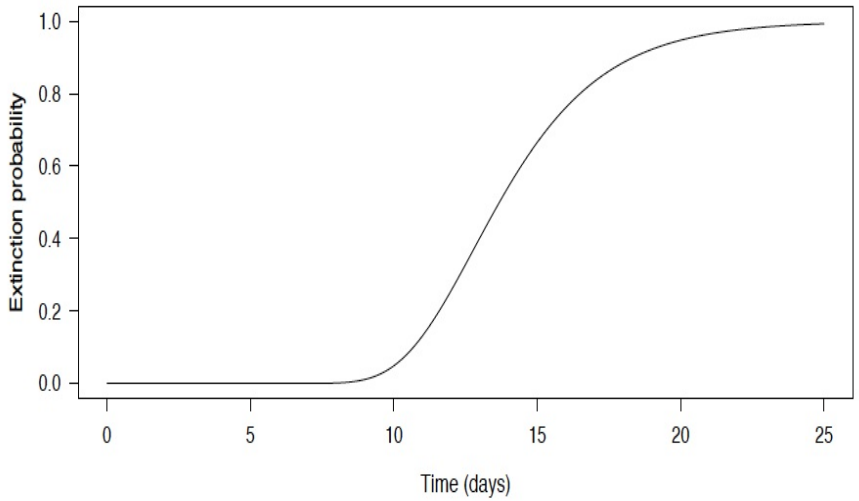
$\gamma = 0.44$

$N = 100$

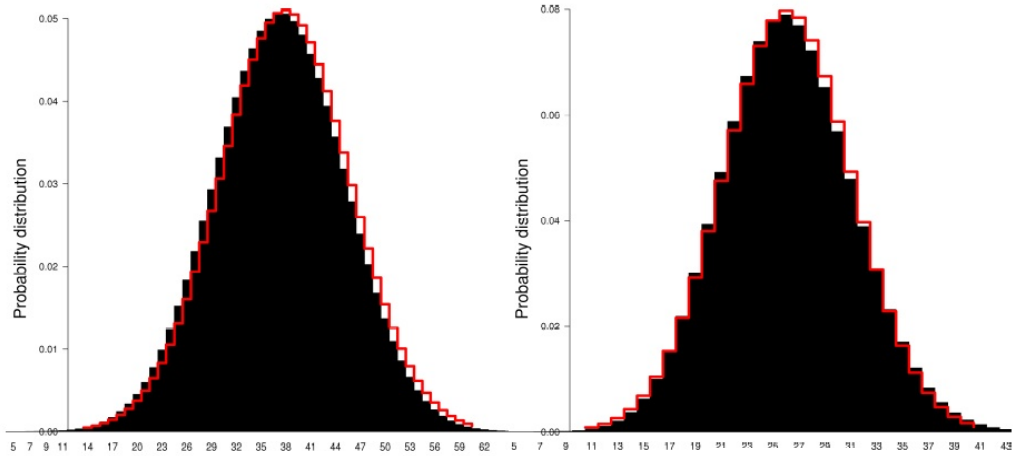
$i_0 = 15$



Probability of extinction



Comparison of the distribution



day 1 day 5

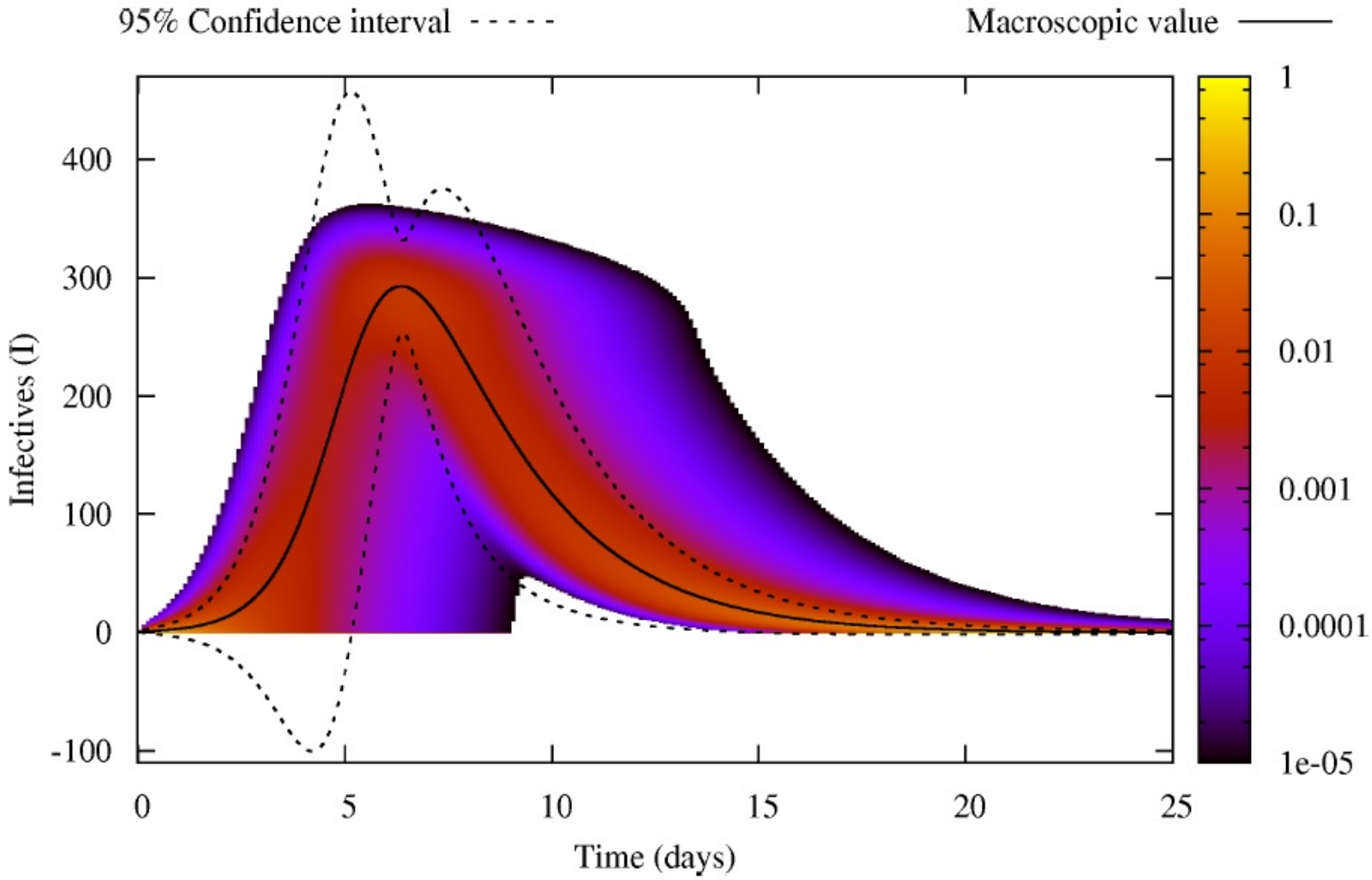
Validation of the *van Kampen* Approximation

$\beta = 1.66$

$\gamma = 0.44$

$N = 763$

$i_0 = 1$



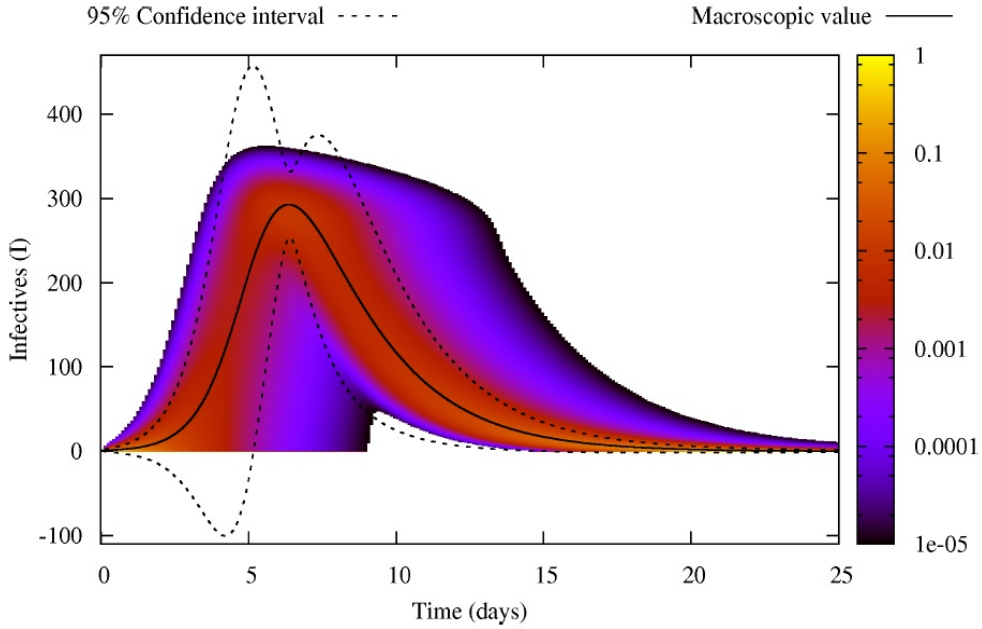
Validation of the *van Kampen* Approximation

$\beta = 1.66$

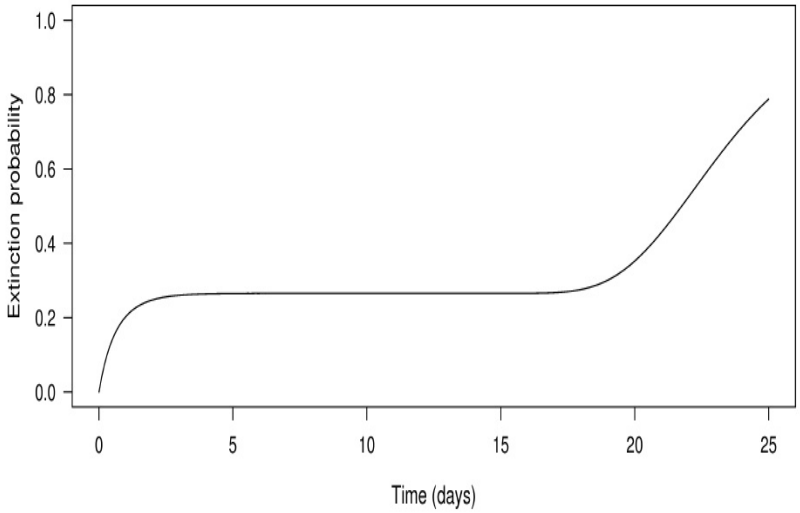
$\gamma = 0.44$

$N = 763$

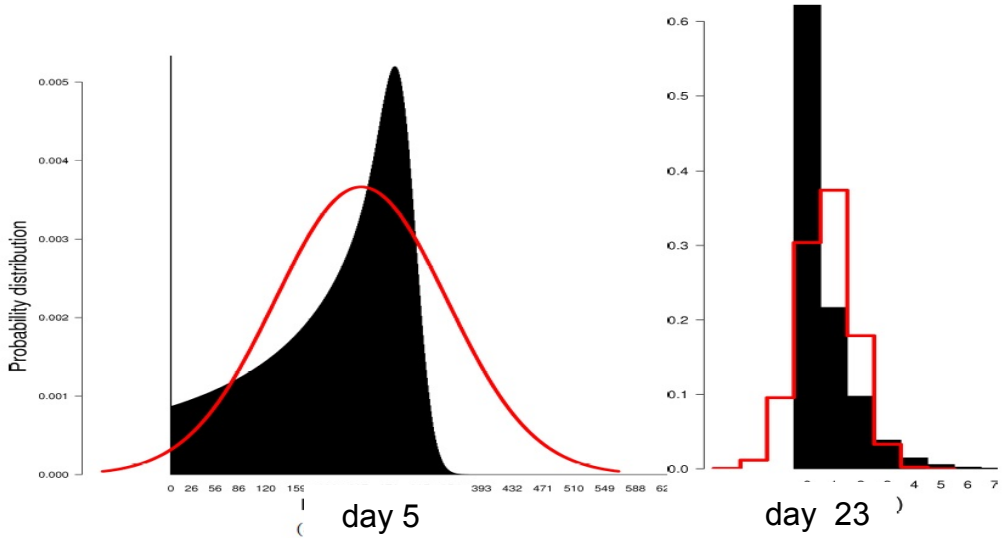
$i_0 = 1$



Probability of extinction



Comparison of the distribution



day 5

day 23

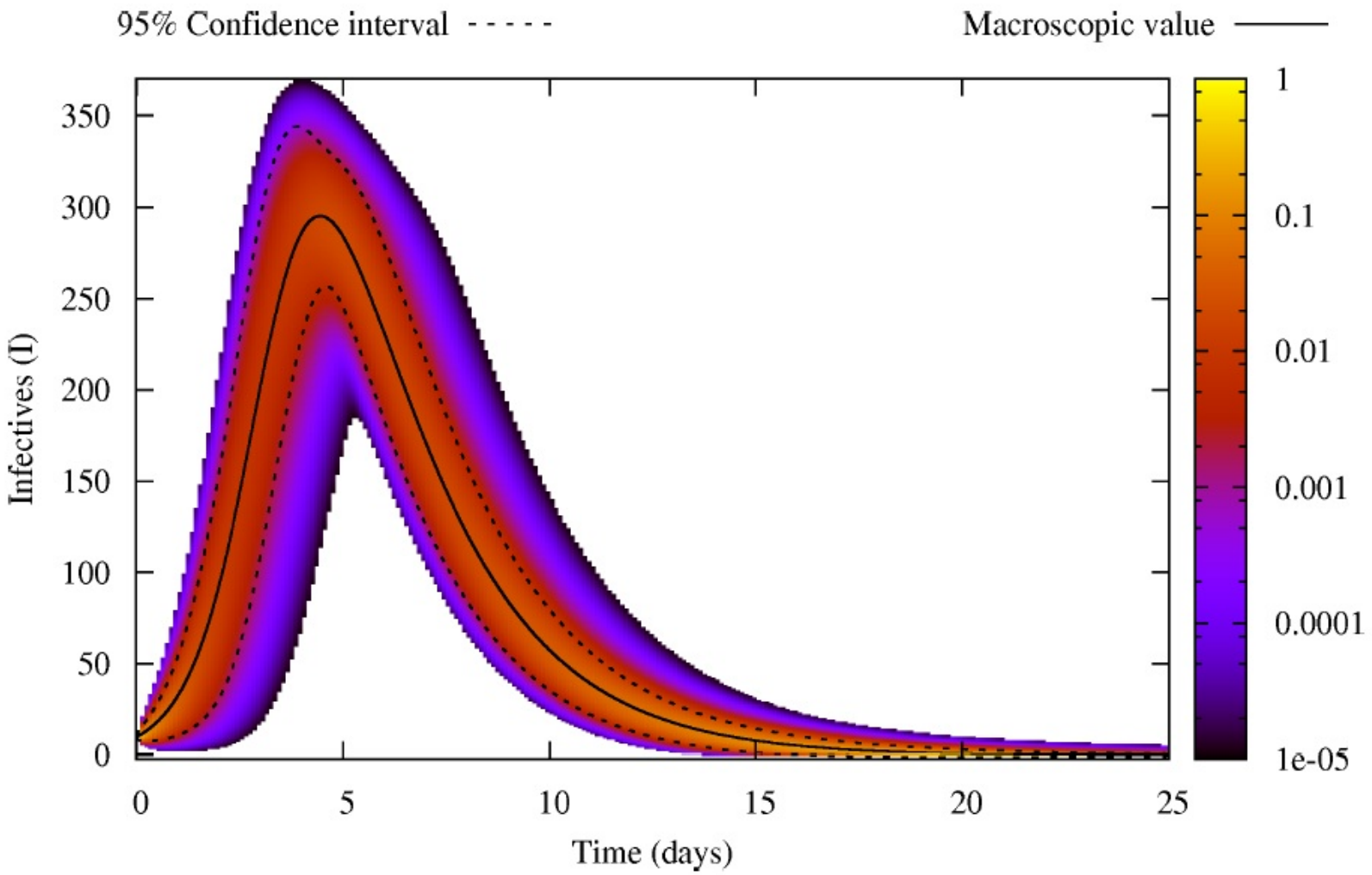
Validation of the *van Kampen* Approximation

$\beta = 1.66$

$\gamma = 0.44$

$N = 763$

$i_0 = 10$



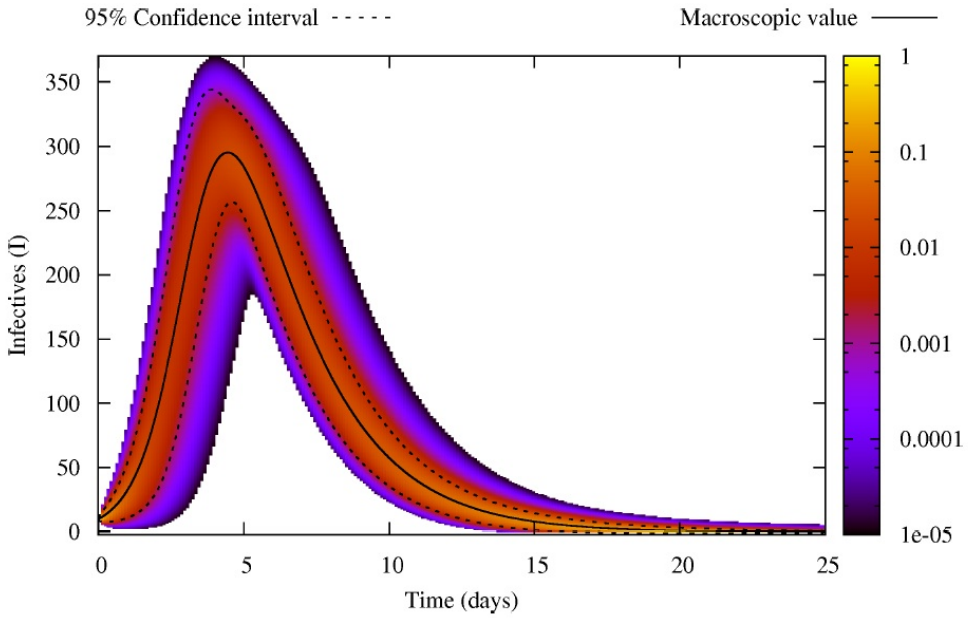
Validation of the *van Kampen* Approximation

$\beta = 1.66$

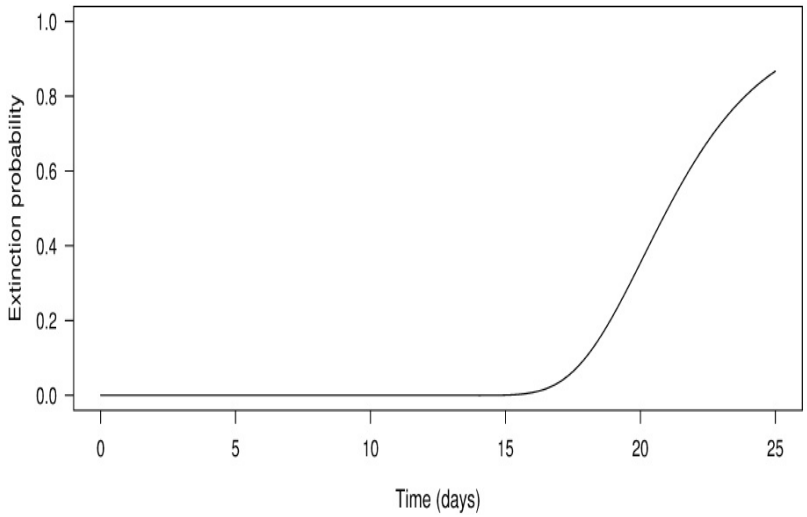
$\gamma = 0.44$

$N = 763$

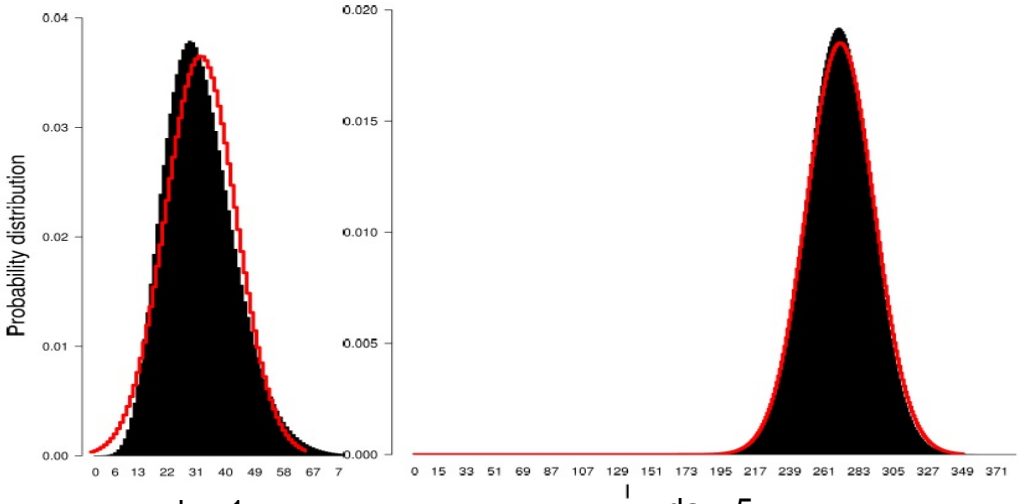
$i_0 = 10$



Probability of extinction



Comparison of the distribution



day 1

day 5

Validation of the *van Kampen* Approximation

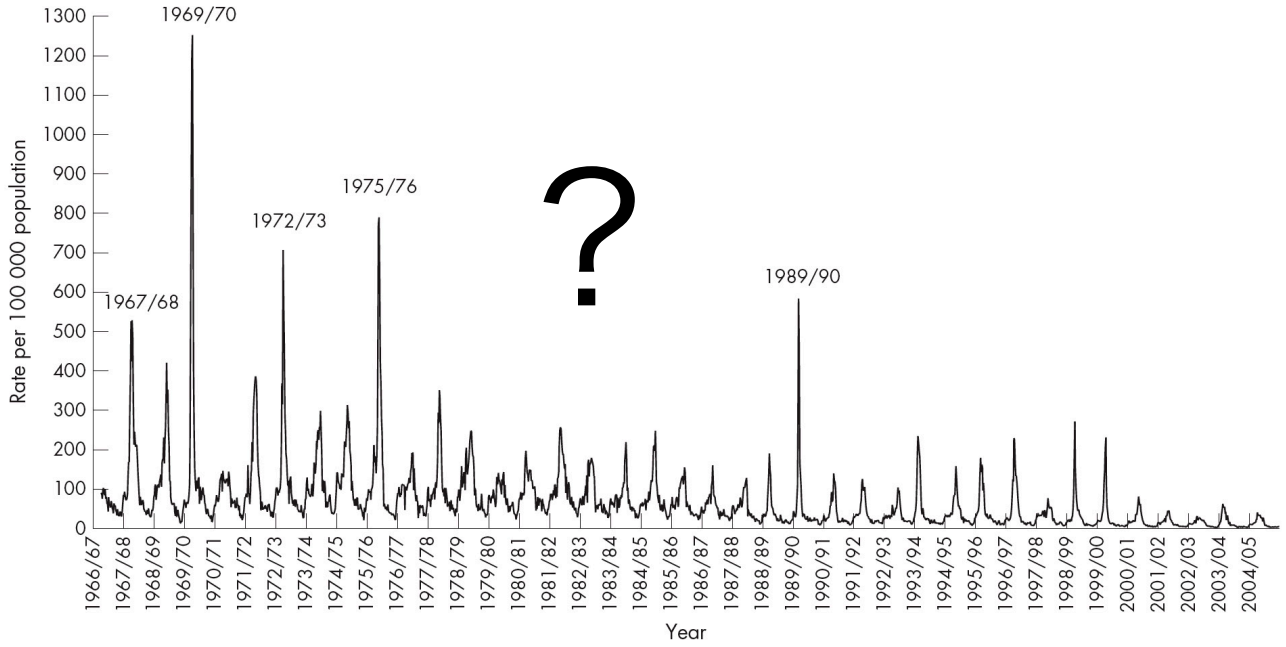
Concluding remarks about the estimation of the variability due to demographic stochasticity:

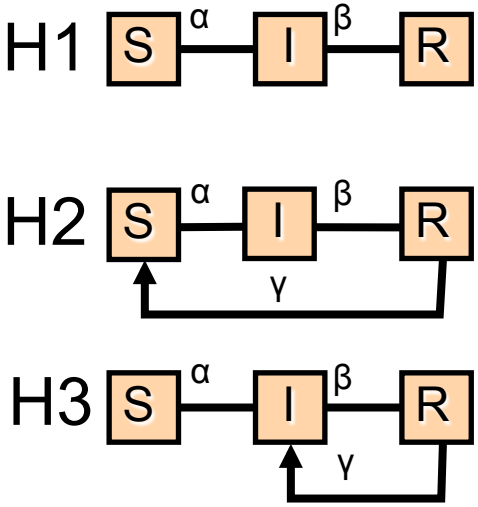
- When the population is small and the model is simple: Expokit
- When the population is intermediate ($< 10^6$): Gillespie's algorithm.
- When the population is large ($> 10^6$): analytical approximations of the master equation as the Van Kampen approximation

Likelihood-Based Inference

Likelihood-Based Inference

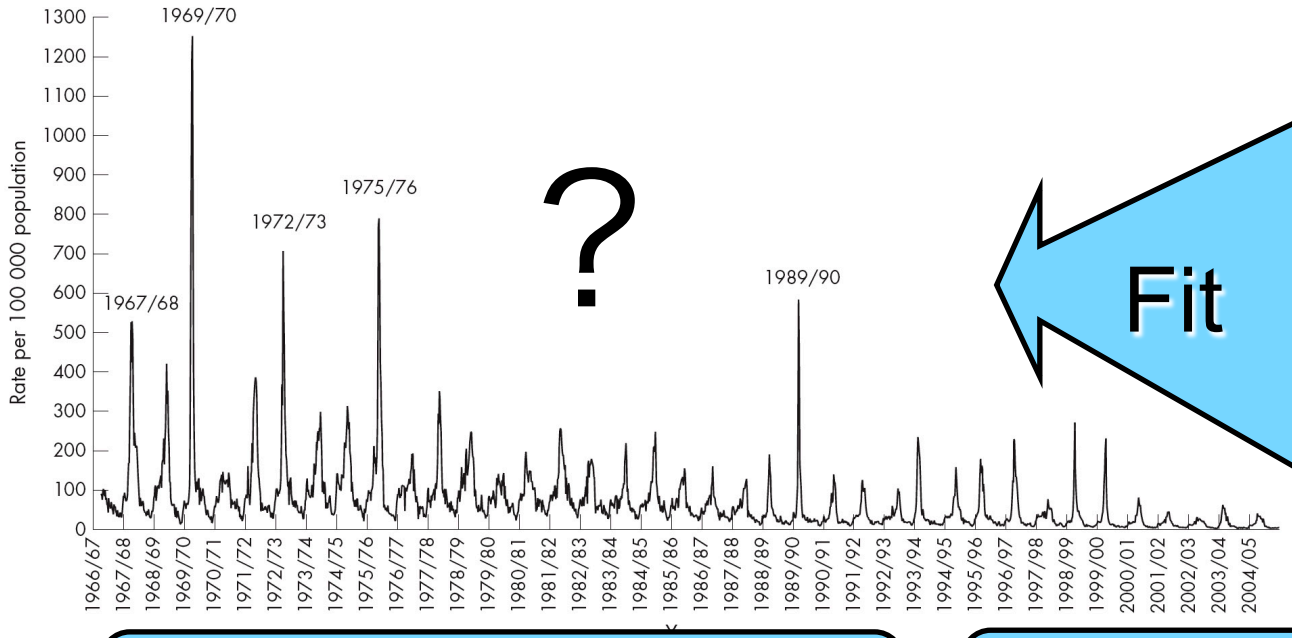
Hypotheses (biological, physical, etc) ⟶ Mechanistic models



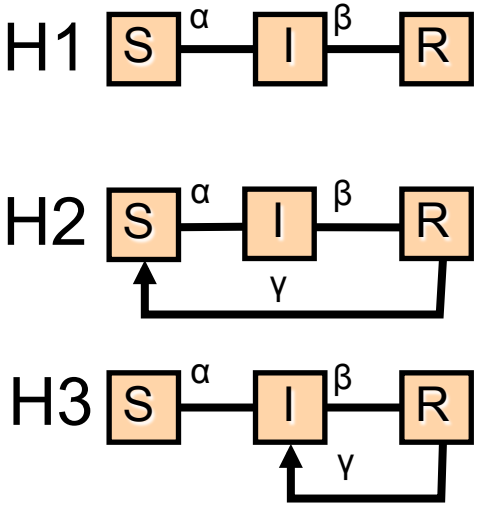


Likelihood-Based Inference

Hypotheses (biological, physical, etc) → Mechanistic models



Fit



Parameter inference:
Identifiability, maximum likelihood estimates, confidence intervals

Model selection:
objective ranking of models, which hypothesis best explains the data?

Particle Filter: MIF

Likelihood-Based Inference

For a given time series: $y_{1:T} = (y_1, y_2, ..., y_T)$
and a state space model completely specified by:

$$M : \begin{cases} f(x_t|x_{t-1}, \theta) & \text{the conditional transition density} \\ f(y_t|x_t, \theta) & \text{the conditional distribution} \\ & \text{of the observation process} \\ f(x_0|\theta) & \text{the initial density} \end{cases}$$

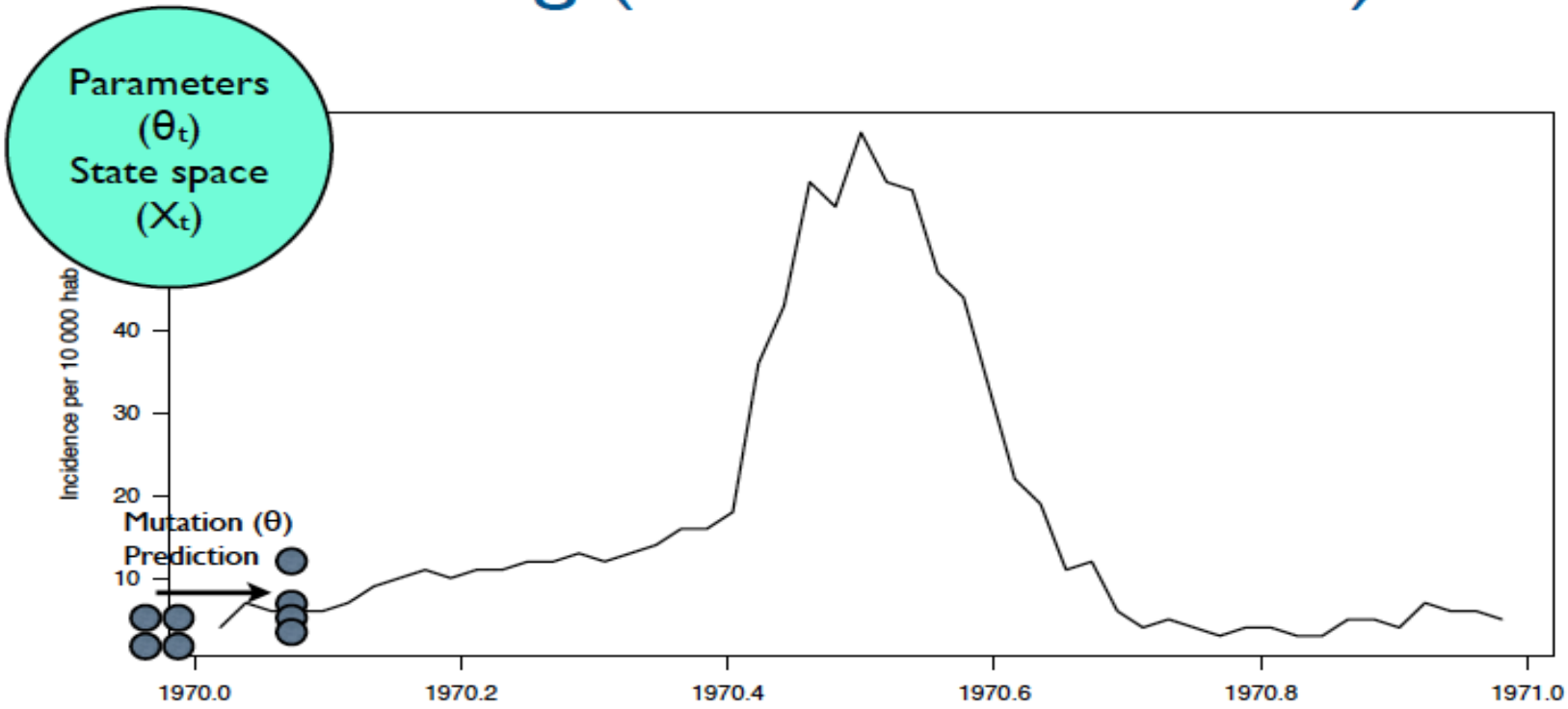
the likelihood is given by the identity:

$$f(y_{1:T}|\theta) = \prod_{t=1}^{T} f(y_t|y_{1:t-1}, \theta)$$

where x_t is the unobserved Markov process, θ is the unknown vector of parameters and $f(.|.)$ is a generic density specified by its arguments

Likelihood-Based Inference

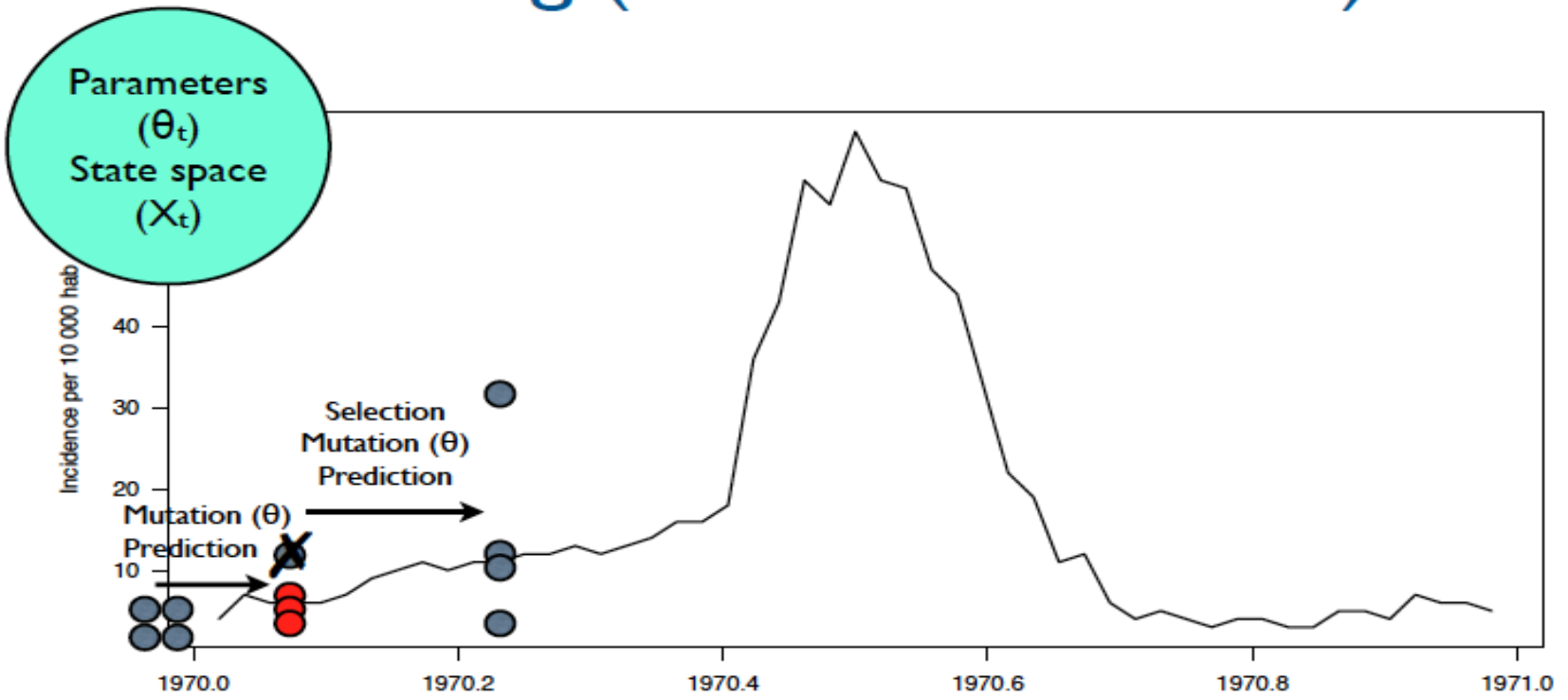
Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



$$\text{For each particle } j: \begin{cases} X_{1,j} & \text{is drawn from} & f(x_1|x_0 = X_{0,j}, \theta_{0,j}) \\ \theta_{1,j} & \text{is drawn from} & \mathcal{N}(\theta_{0,j}, \sigma) \\ w_{1,j} & \text{is equal to} & f(y_1|x_1 = X_{1,j}, \theta_{1,j}) \end{cases}$$

Likelihood-Based Inference

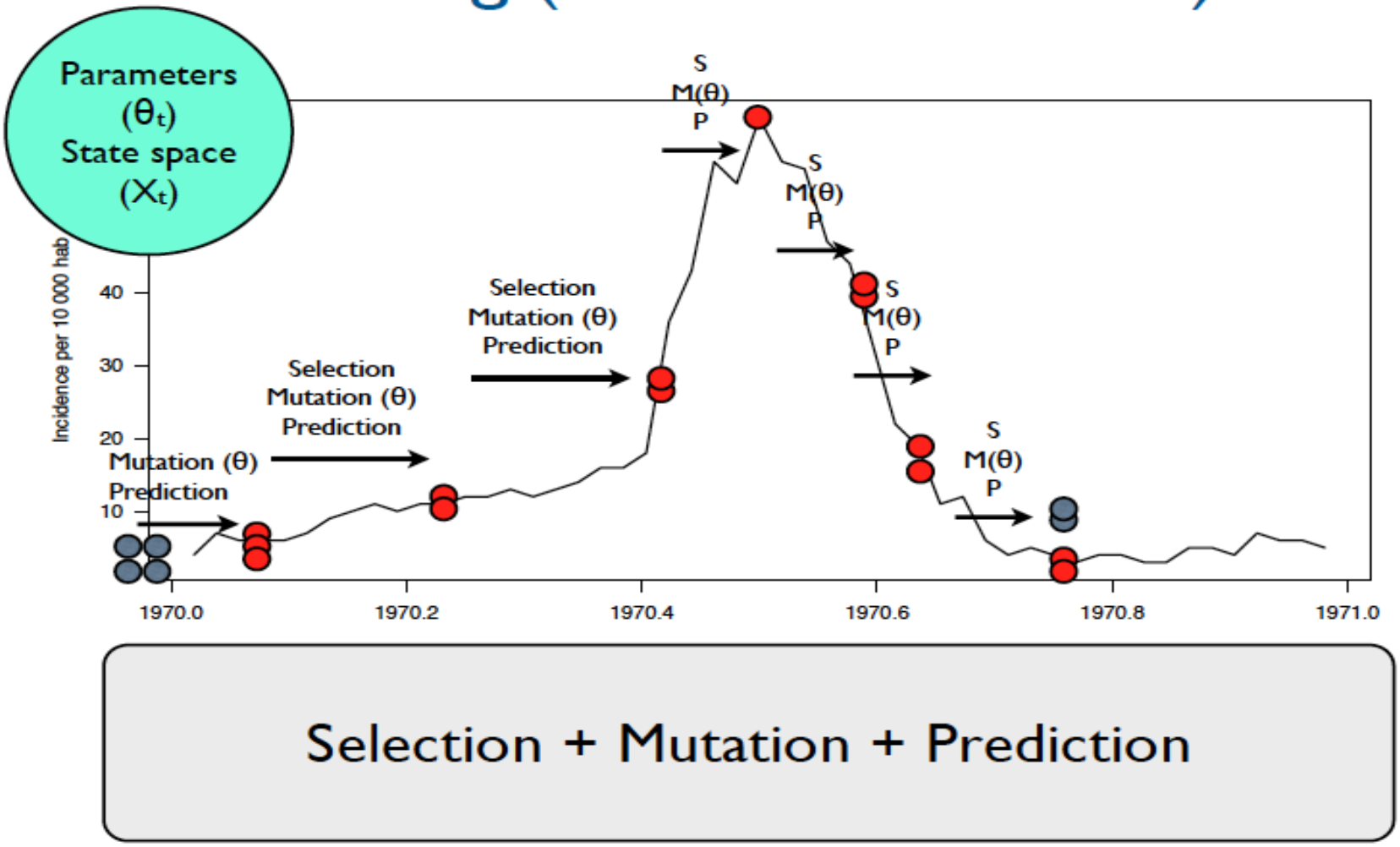
Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



Darwinian selection: particles reproduce proportionally to their weight w_j

Likelihood-Based Inference

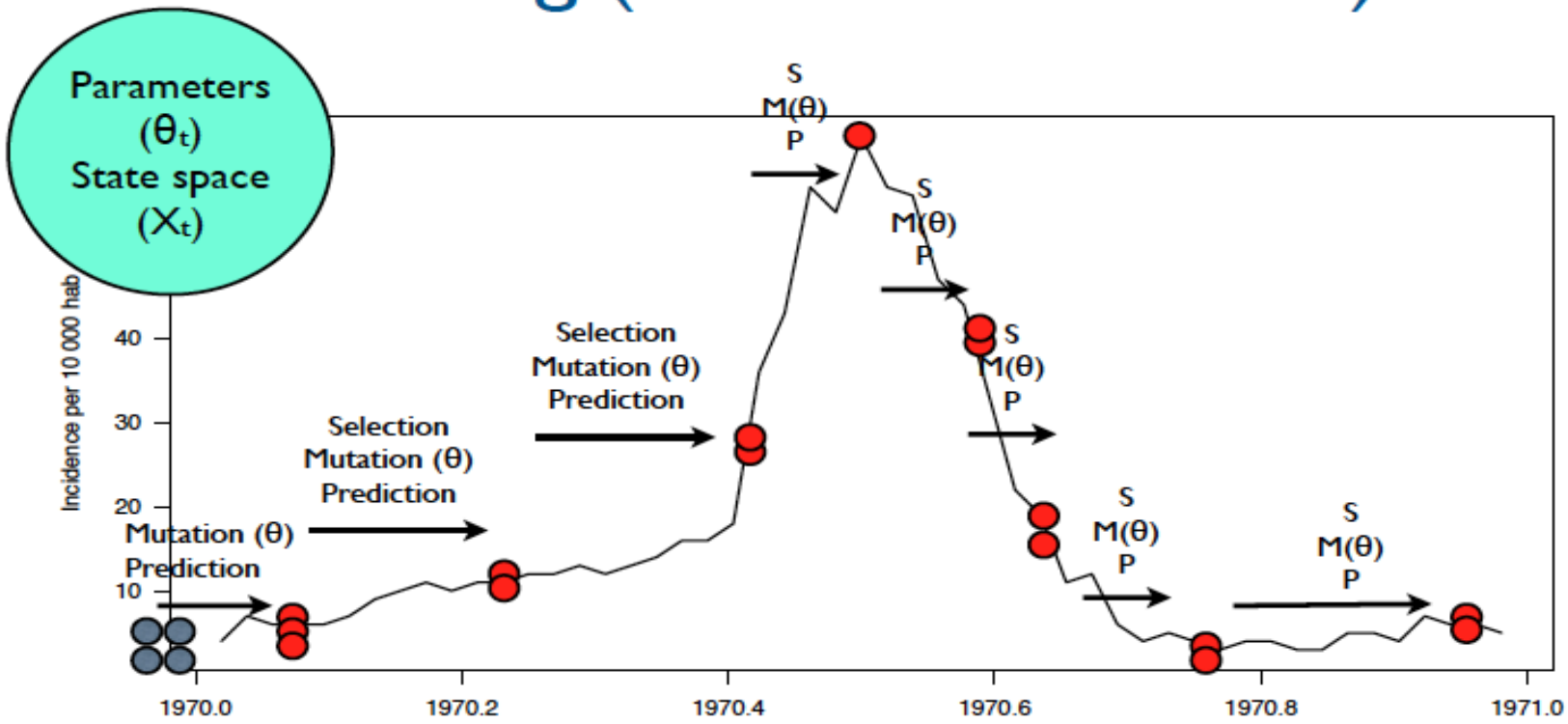
Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



Selection + Mutation + Prediction

Likelihood-Based Inference

Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)

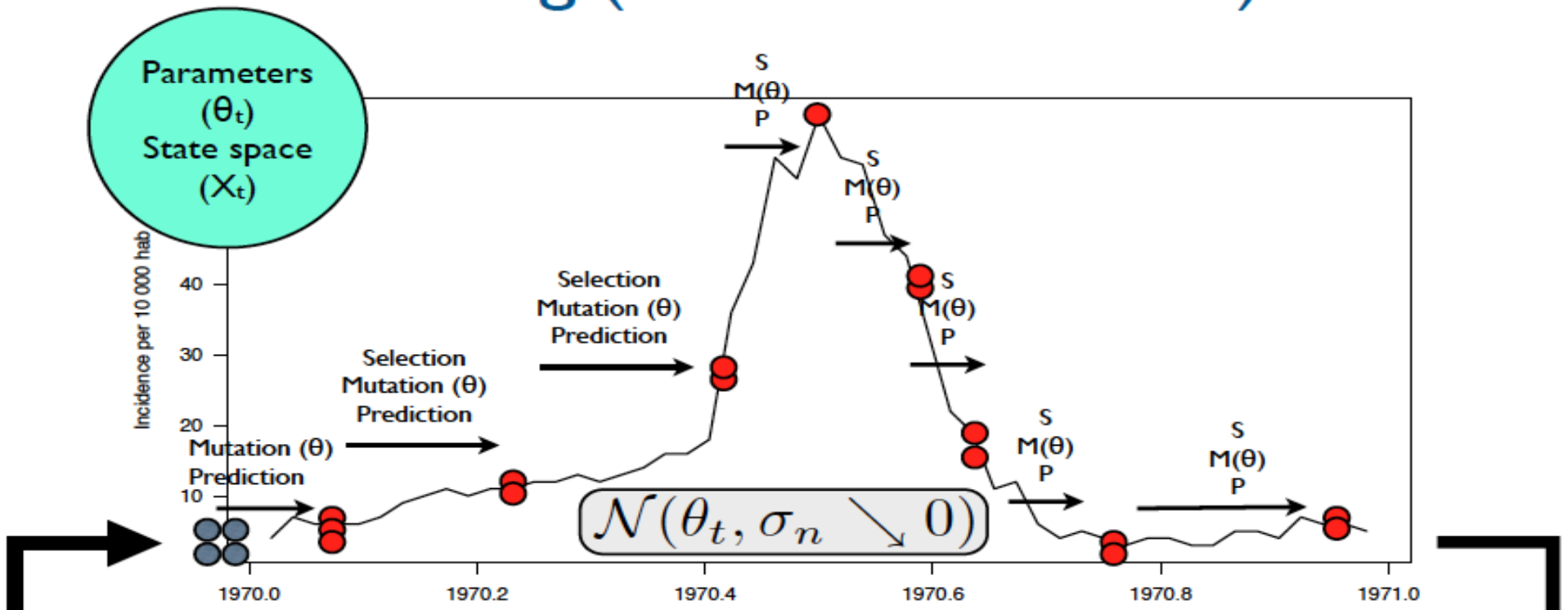


Global estimate: $\hat{\theta} = \hat{\theta}_0 + V_1 \sum_{t=1}^{T} \frac{\hat{\theta}_t - \hat{\theta}_{t-1}}{V_t}$

Log-likelihood: $\mathcal{L}(\hat{\theta}) = \log(\prod_T l_t(\theta))$

Likelihood-Based Inference

Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



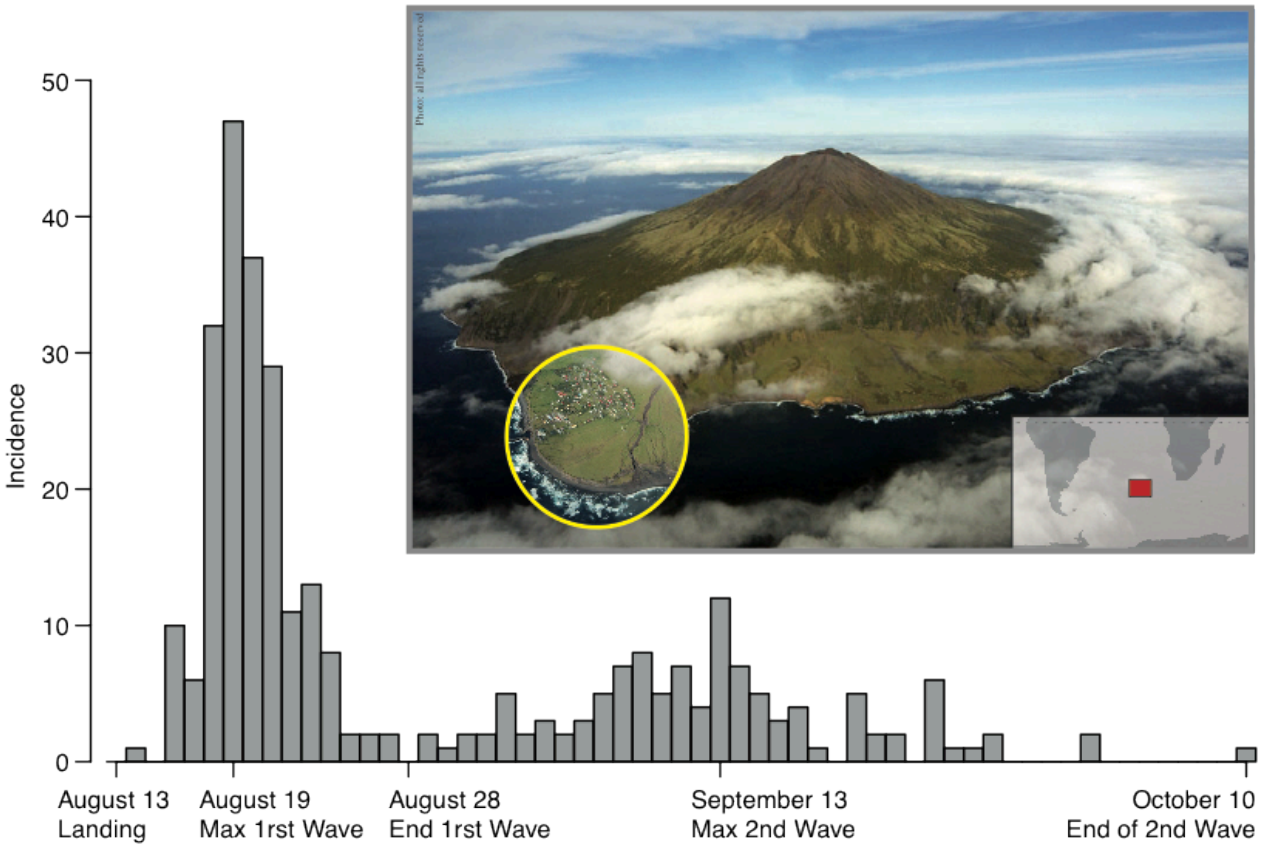
Global estimate: $\hat{\theta}^{(n)} = \hat{\theta}^{(n-1)} + V_1^{(n)} \sum_{t=1}^{T} \frac{\hat{\theta}_t^{(n)} - \hat{\theta}_{t-1}^{(n)}}{V_t^{(n)}}$

Log-likelihood: $\mathcal{L}(\hat{\theta}^{(n)}) = \log(\prod_T l_t^{(n)}(\theta))$

Explaining rapid reinfections in multiple-wave influenza outbreaks

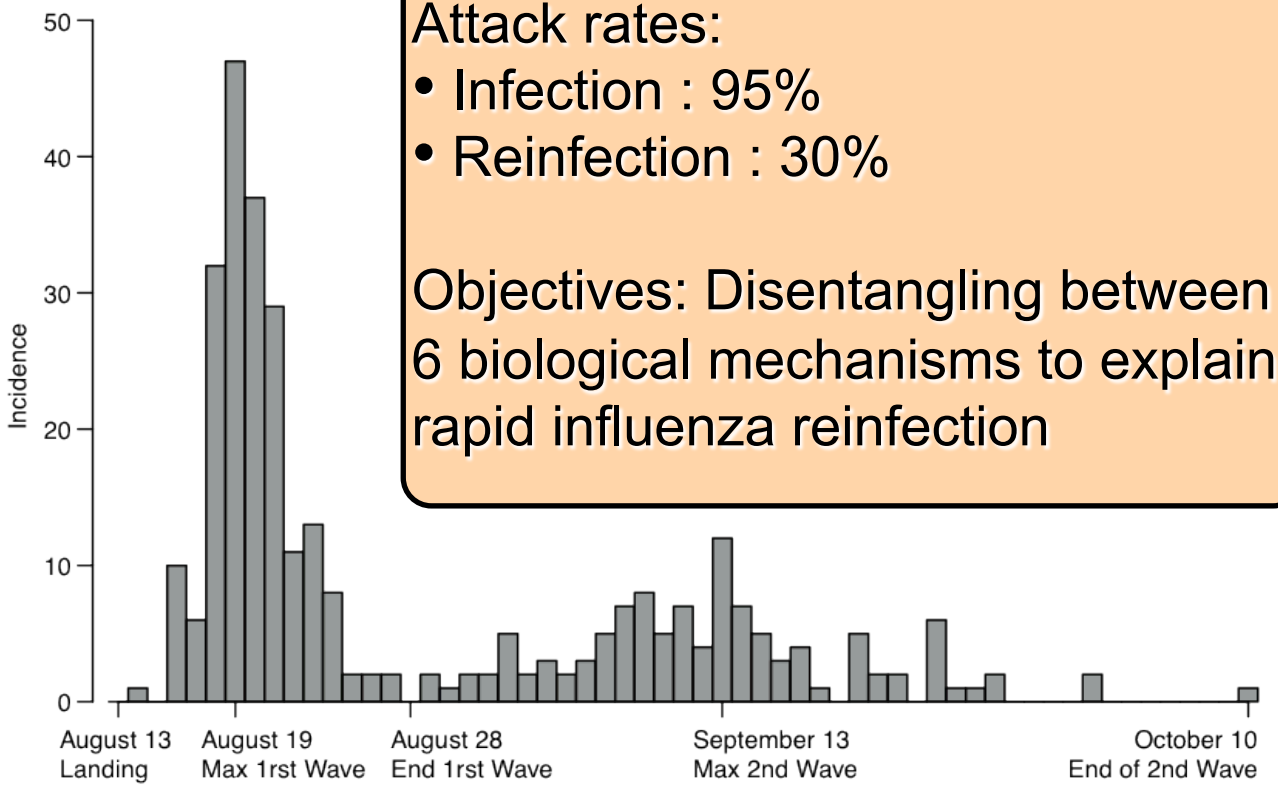
Explaining rapid reinfections in multiple-wave influenza outbreaks

Tristan da Cunha (1971)
a two-wave flu epidemic



Explaining rapid reinfections in multiple-wave influenza outbreaks

Tristan da Cunha (1971)
a two-wave flu epidemic

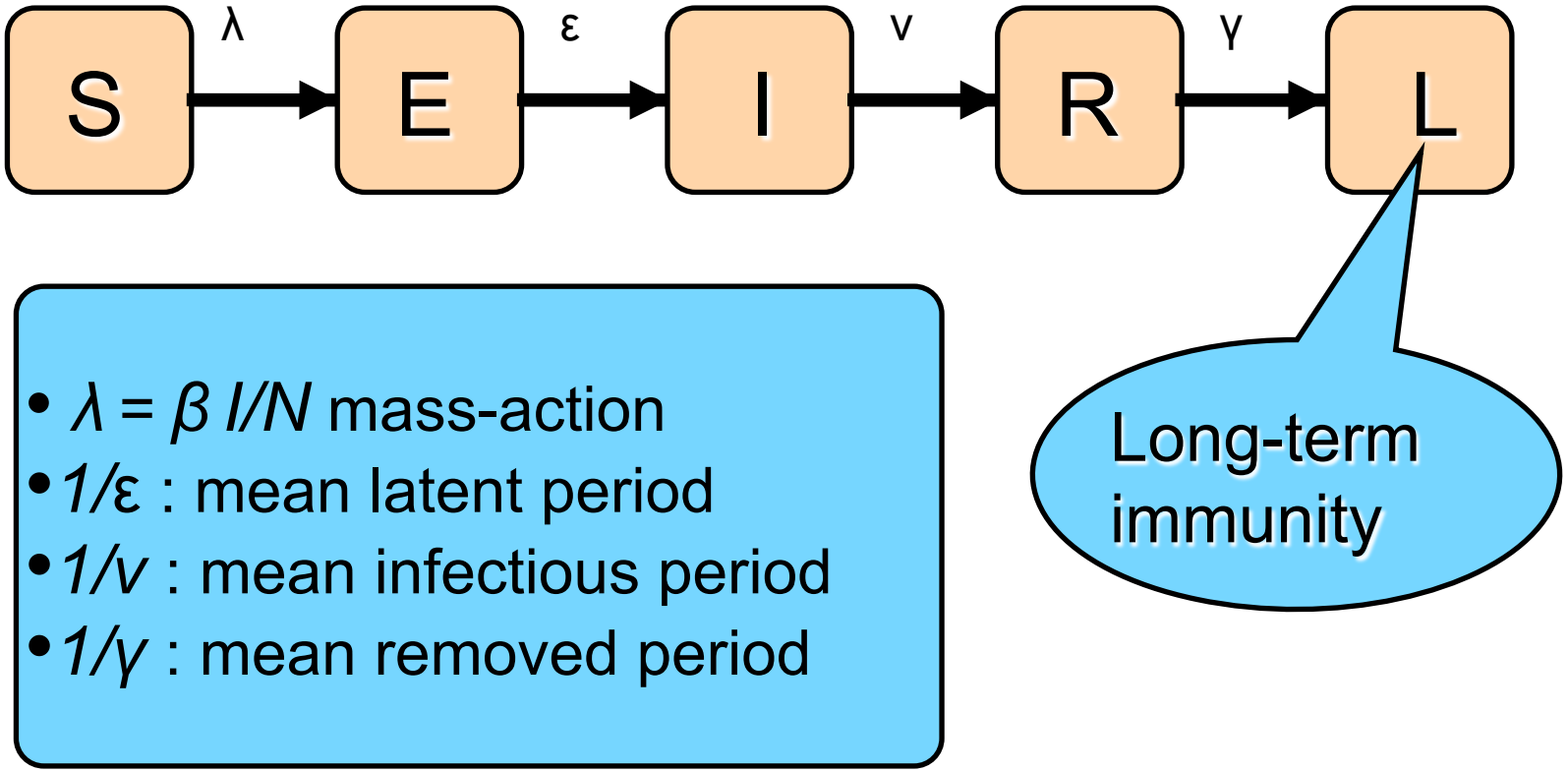


Attack rates:
- Infection : 95%
- Reinfection : 30%

Objectives: Disentangling between 6 biological mechanisms to explain rapid influenza reinfection

Explaining rapid reinfections in multiple-wave influenza outbreaks

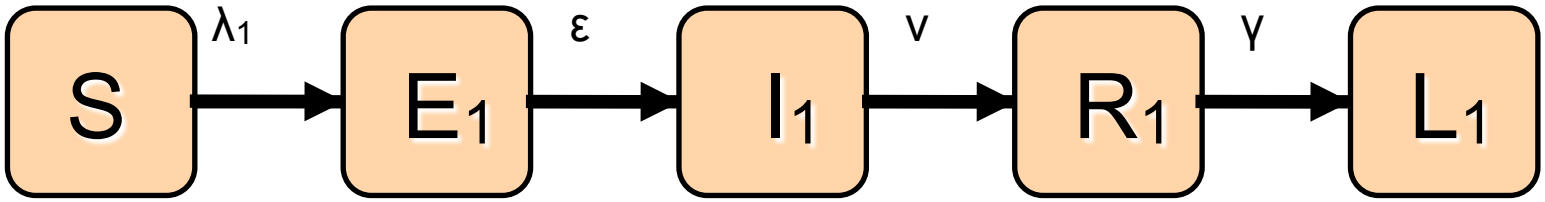
A simple mechanistic approach



- $\lambda = \beta I/N$ mass-action
- $1/\varepsilon$: mean latent period
- $1/\nu$: mean infectious period
- $1/\gamma$: mean removed period

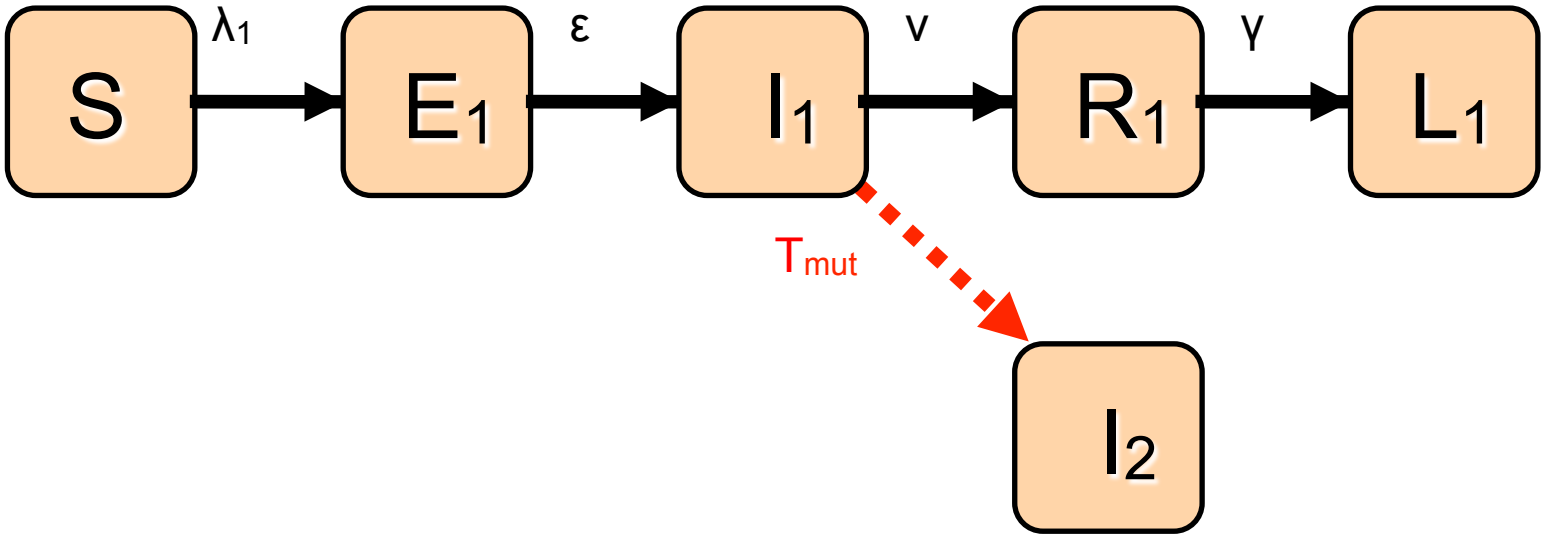
Explaining rapid reinfections in multiple-wave influenza outbreaks

H1: the virus mutated during the first epidemic-wave (Mut)



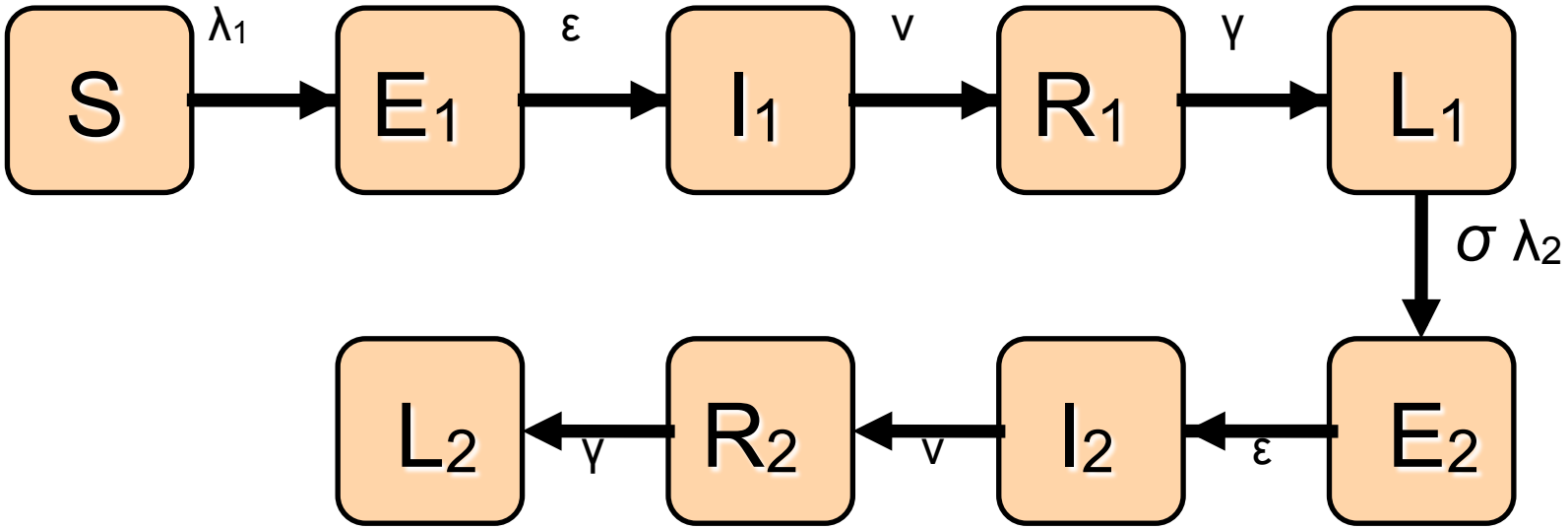
Explaining rapid reinfections in multiple-wave influenza outbreaks

H1: the virus mutated during the first epidemic-wave (Mut)



Explaining rapid reinfections in multiple-wave influenza outbreaks

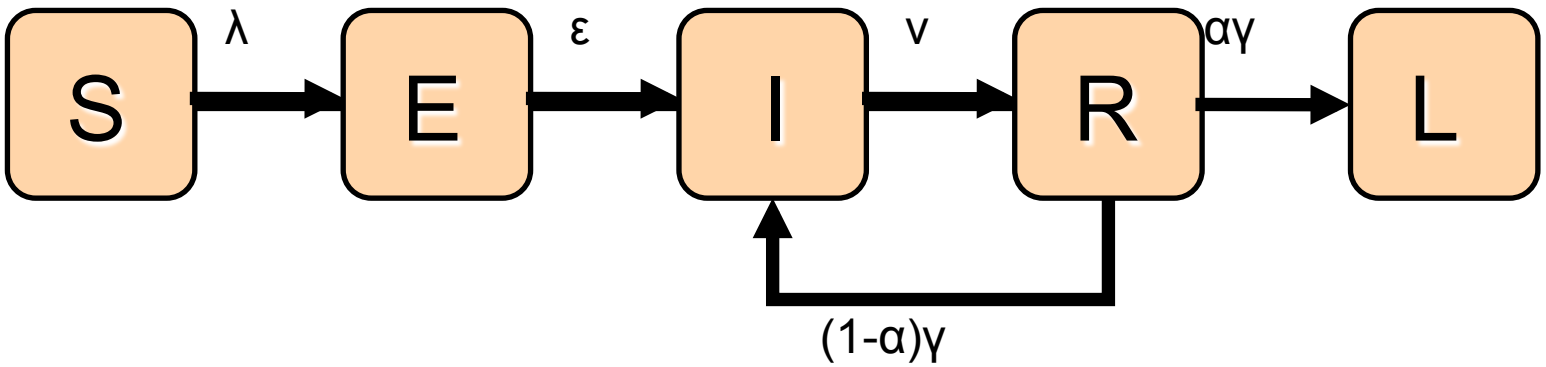
H1: the virus mutated during the first epidemic-wave (Mut)



- $\sigma \in [0,1]$: cross-immunity
- 2-strain history-based model (Rios-Doria & Chowell 2009)

Explaining rapid reinfections in multiple-wave influenza outbreaks

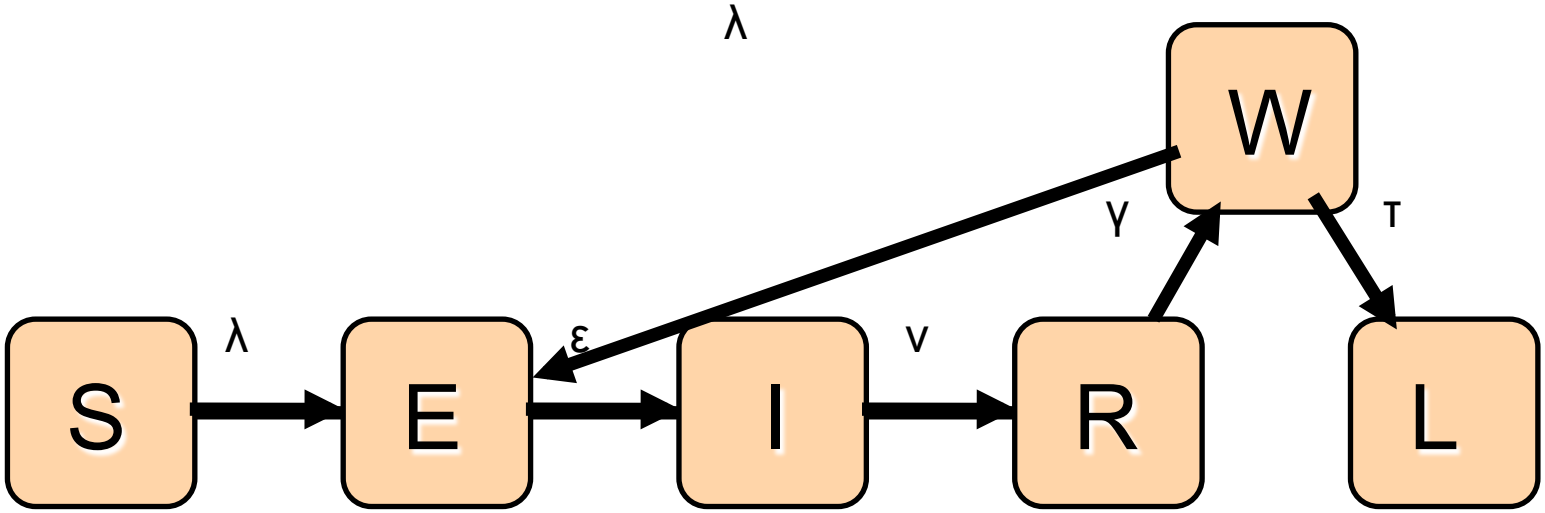
H2: intra-host recrudescence of infection (InH)



α: the probability to clear the viral load

Explaining rapid reinfections in multiple-wave influenza outbreaks

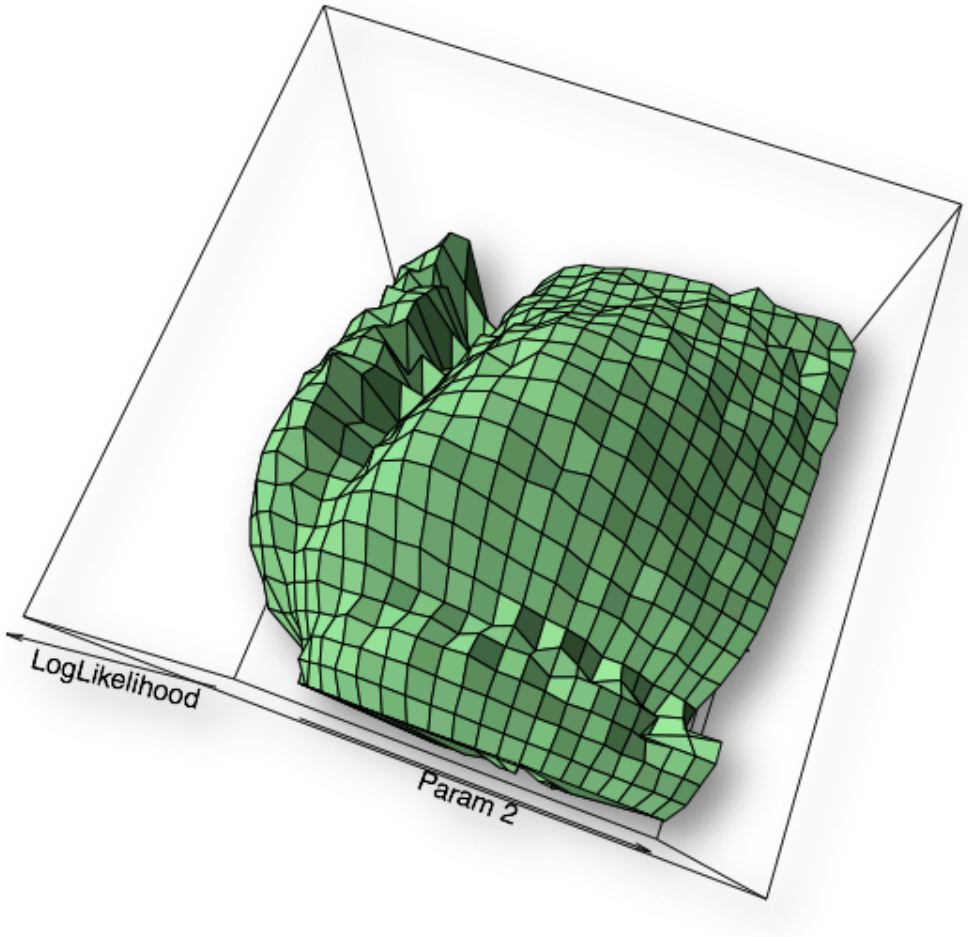
H3: window-of-reinfection (Win)



1/τ: the mean duration of the window of susceptibility before developing immunity

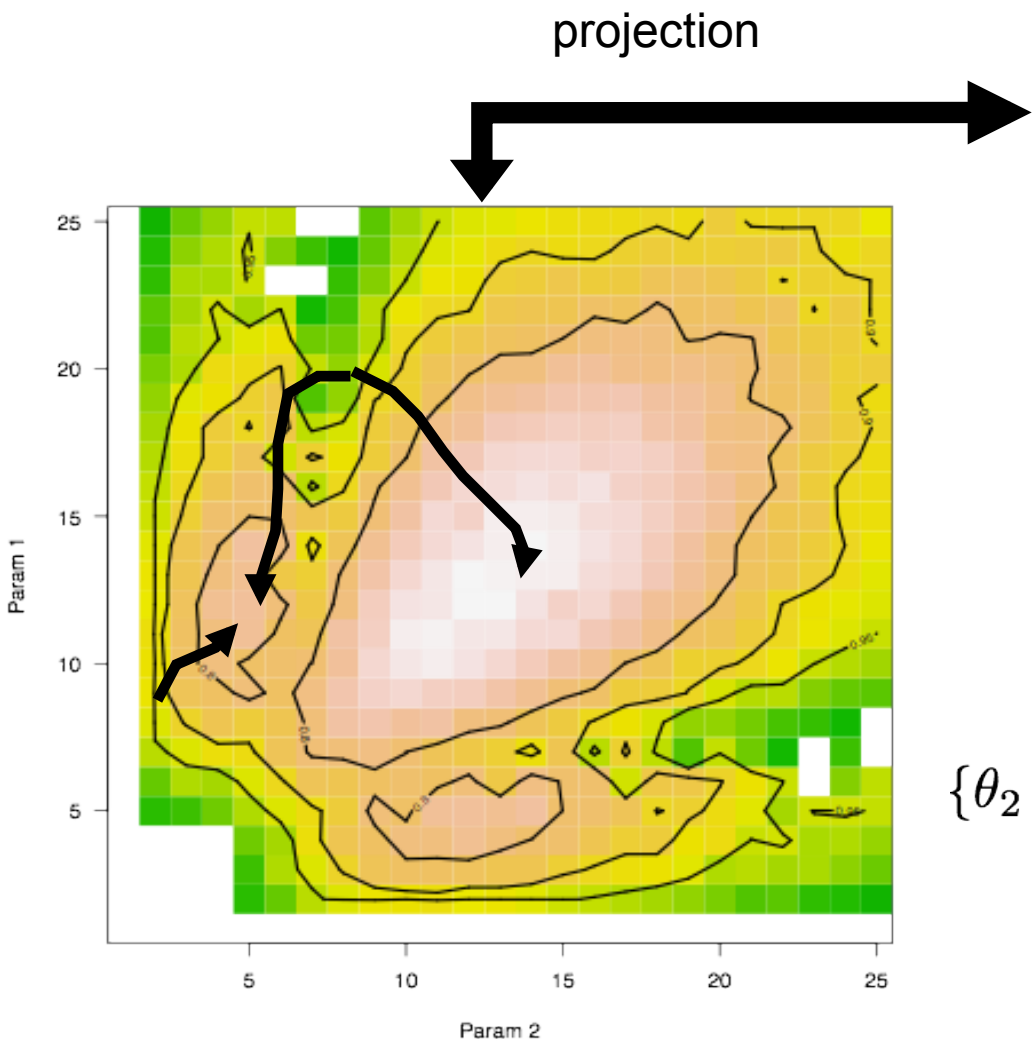
Explaining rapid reinfections in multiple-wave influenza outbreaks

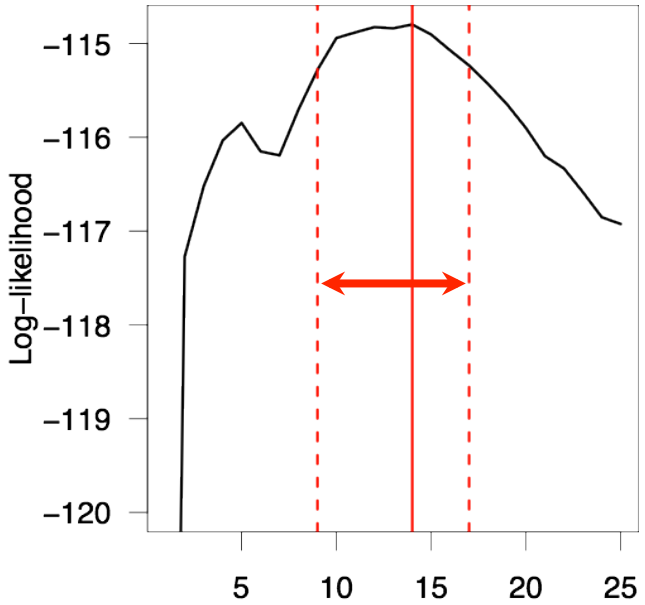
Exploring the likelihood surface



Explaining rapid reinfections in multiple-wave influenza outbreaks

Log-likelihood profile





95% confidence interval:

$$\{\theta_2 : 2[l_{(p)}(\hat{\theta}_2) - l_{(p)}(\theta_2)] < \chi^2_{0.95}(1)\}$$

$$\text{with } \hat{\theta}_2 = \operatorname{argmax} l_{(p)}(\theta_2)$$

Explaining rapid reinfections in multiple-wave influenza outbreaks

Model selection: Akaike information criterion

$$AIC_c = -2\mathcal{L}(\theta_{MLE}) + 2k + \frac{2k(k+1)}{T-k-1} \text{ with } k = ||\theta||$$

Model	Win	Mut	In-Host
k	9	10	9
Log-Like	-112.52	-115.20	-117.50
ΔAIC_c	0	8.27	9.96

Camacho et al, 2011: PRSB

Explaining rapid reinfections in multiple-wave influenza outbreaks

Model selection: Akaike information criterion

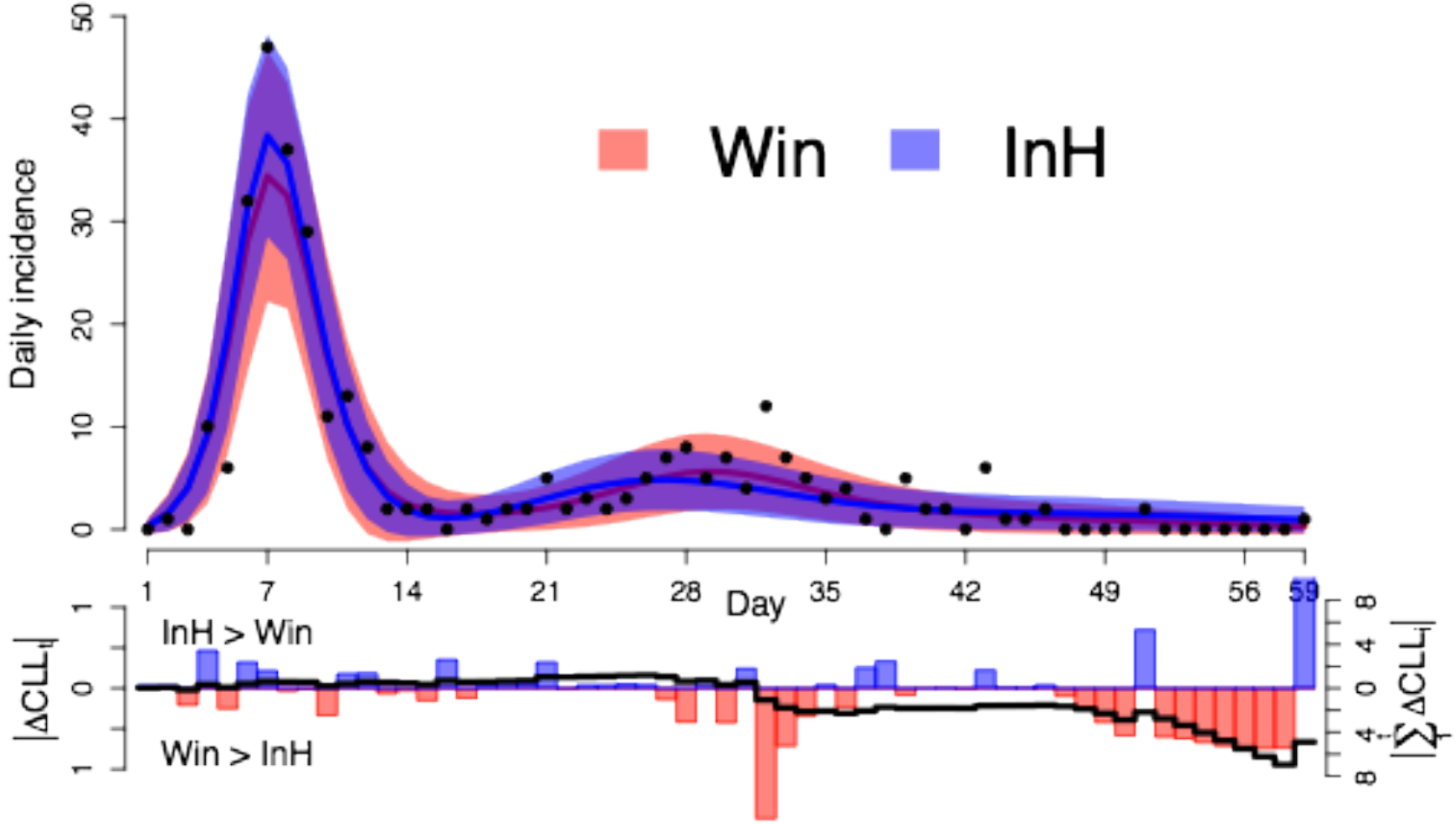
$$AIC_c = -2\mathcal{L}(\theta_{MLE}) + 2k + \frac{2k(k+1)}{T-k-1} \text{ with } k = ||\theta||$$

Model	Win	Mut	In-Host
k	9	10	9
Log-Like	-112.52	-115.20	-117.50
ΔAIC_c	0	8.27	9.96

Camacho et al, 2011: PRSB

Explaining rapid reinfections in multiple-wave influenza outbreaks

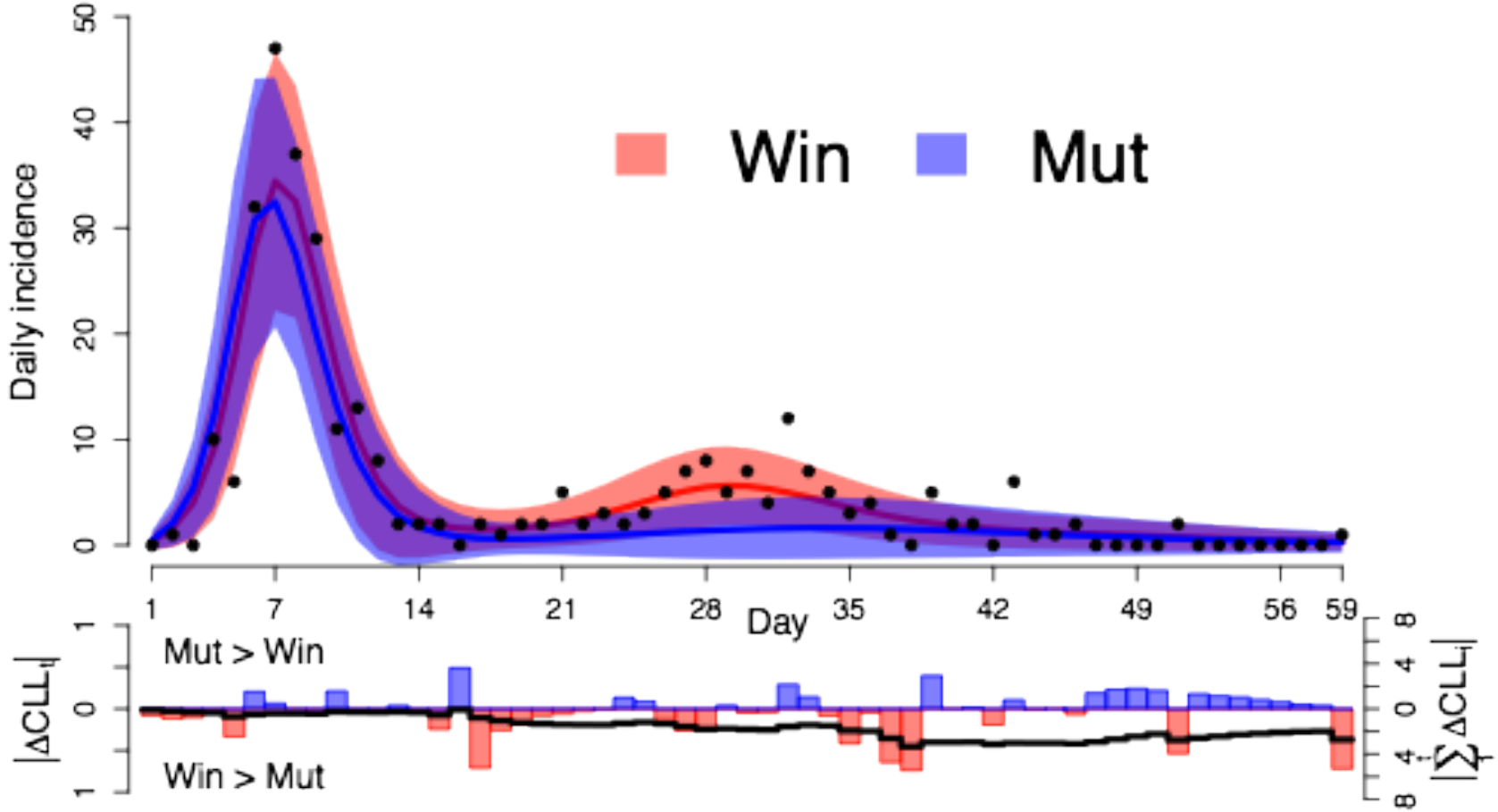
Dynamics comparison



Camacho et al, 2011: PRSB

Explaining rapid reinfections in multiple-wave influenza outbreaks

Dynamics comparison



Camacho et al, 2011: PRSB

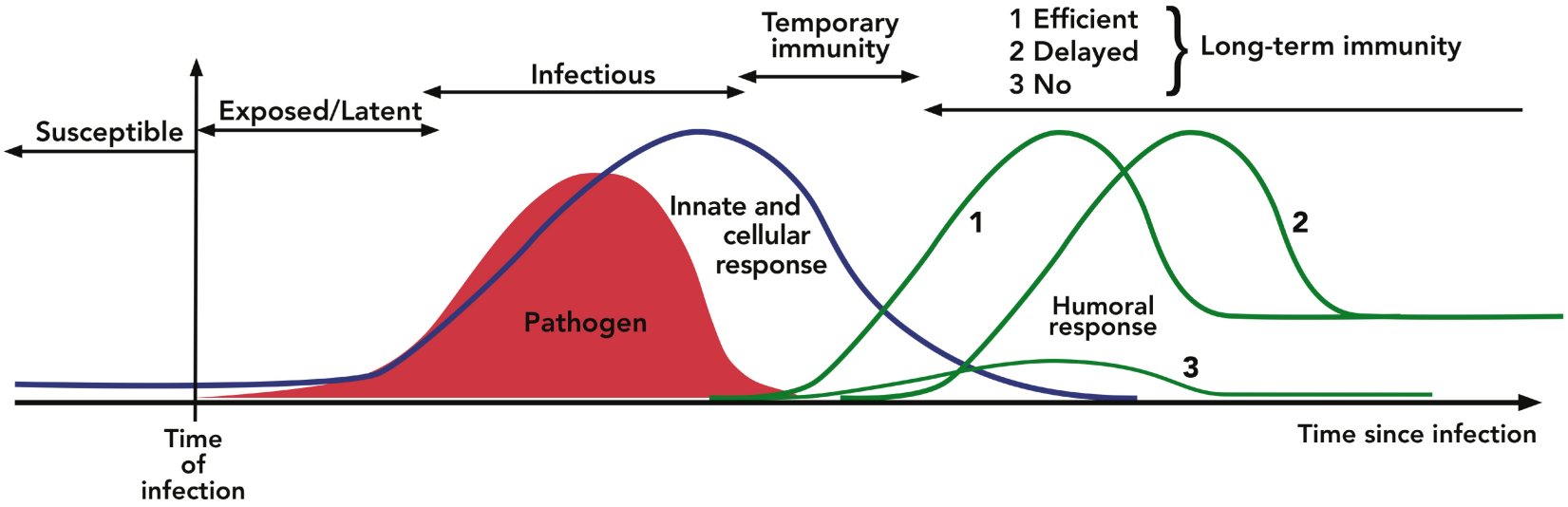
Explaining rapid reinfections in multiple-wave influenza outbreaks

- A stochastic formulation is essential to capture demographic stochasticity induced by small populations.
- Heterogeneity among hosts is a significantly more likely explanation for 1971's two-wave than viral heterogeneity.
- Studies assuming that the immune response always provides a long-term humoral protection should overestimate the amount of immune escape required to sequential influenza variants to cause rapid reinfection.

Accounting for immunodynamics in epidemiological models

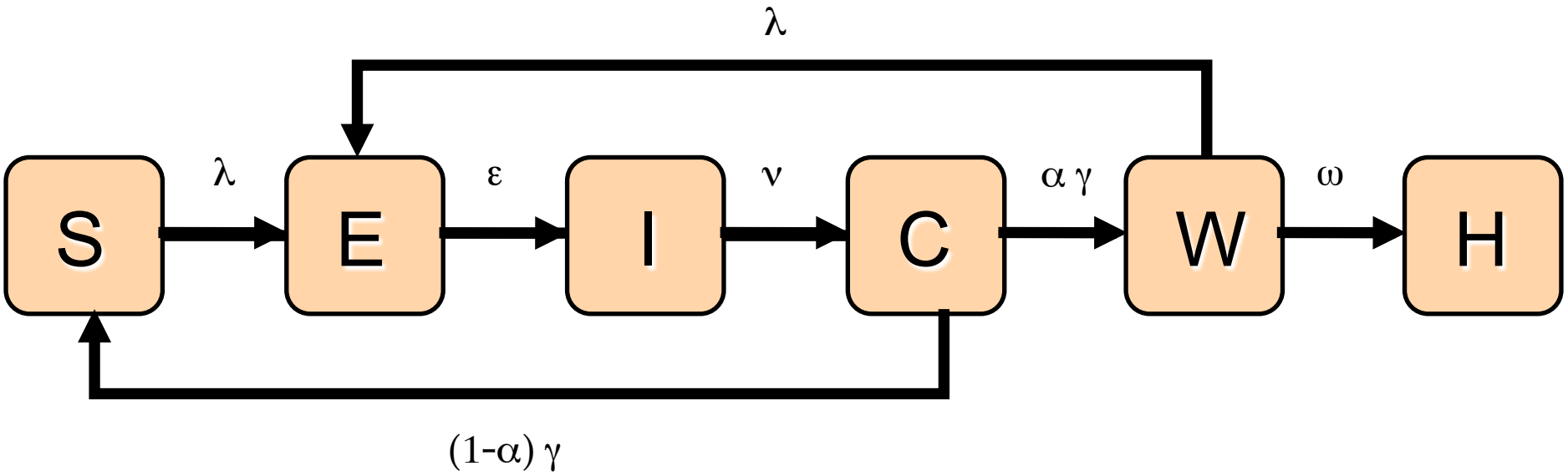
Accounting for immunodynamics in epidemiological models

There is two main immunological responses:
(i) The cellular immune response with the Cytotoxic T Lymphocytes that can eliminate infected cells and then prevent viral release
(ii) The humoral response with T cells that can neutralize the virus.



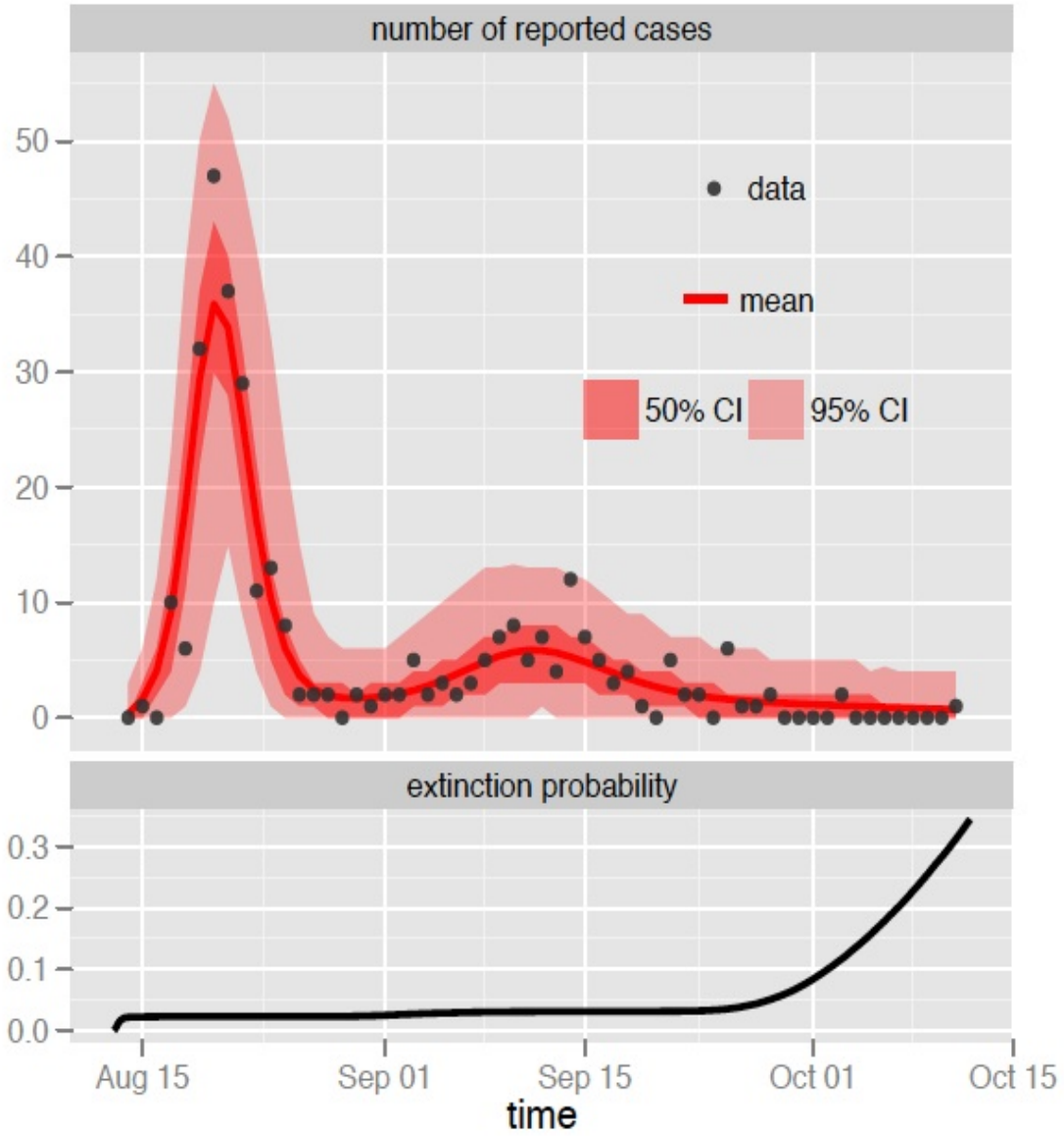
Mechanistic modelling of the primary immune response to influenza. A: schematized dynamics of the viral load as well as the innate and adaptive immune responses

Accounting for immunodynamics in epidemiological models



α : probability of developing an humoral response
1/γ : the mean duration in the cellular protected stage
*1/*ω : the mean duration of the window of susceptibility before developing humoral immunity

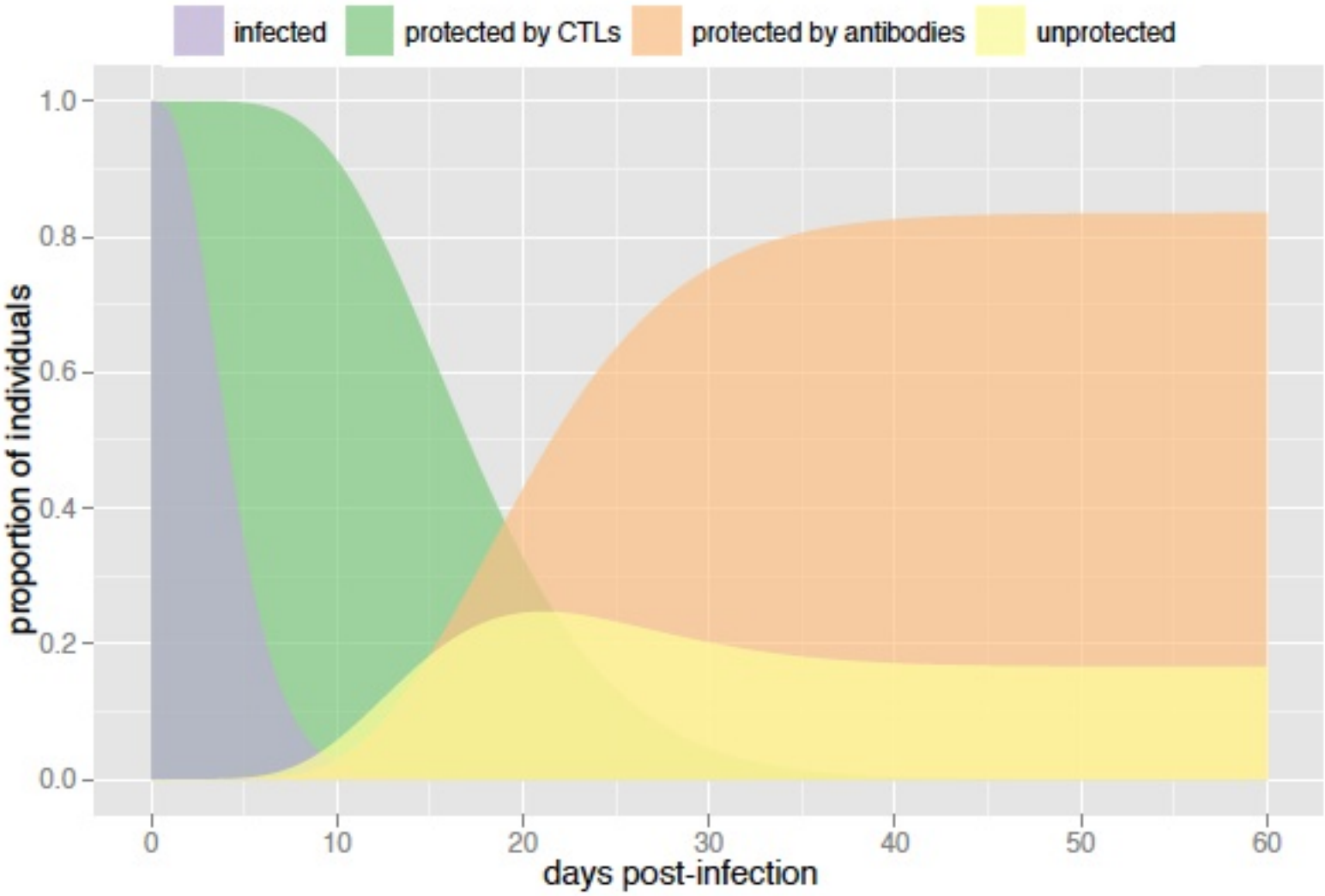
Accounting for immunodynamics in epidemiological models



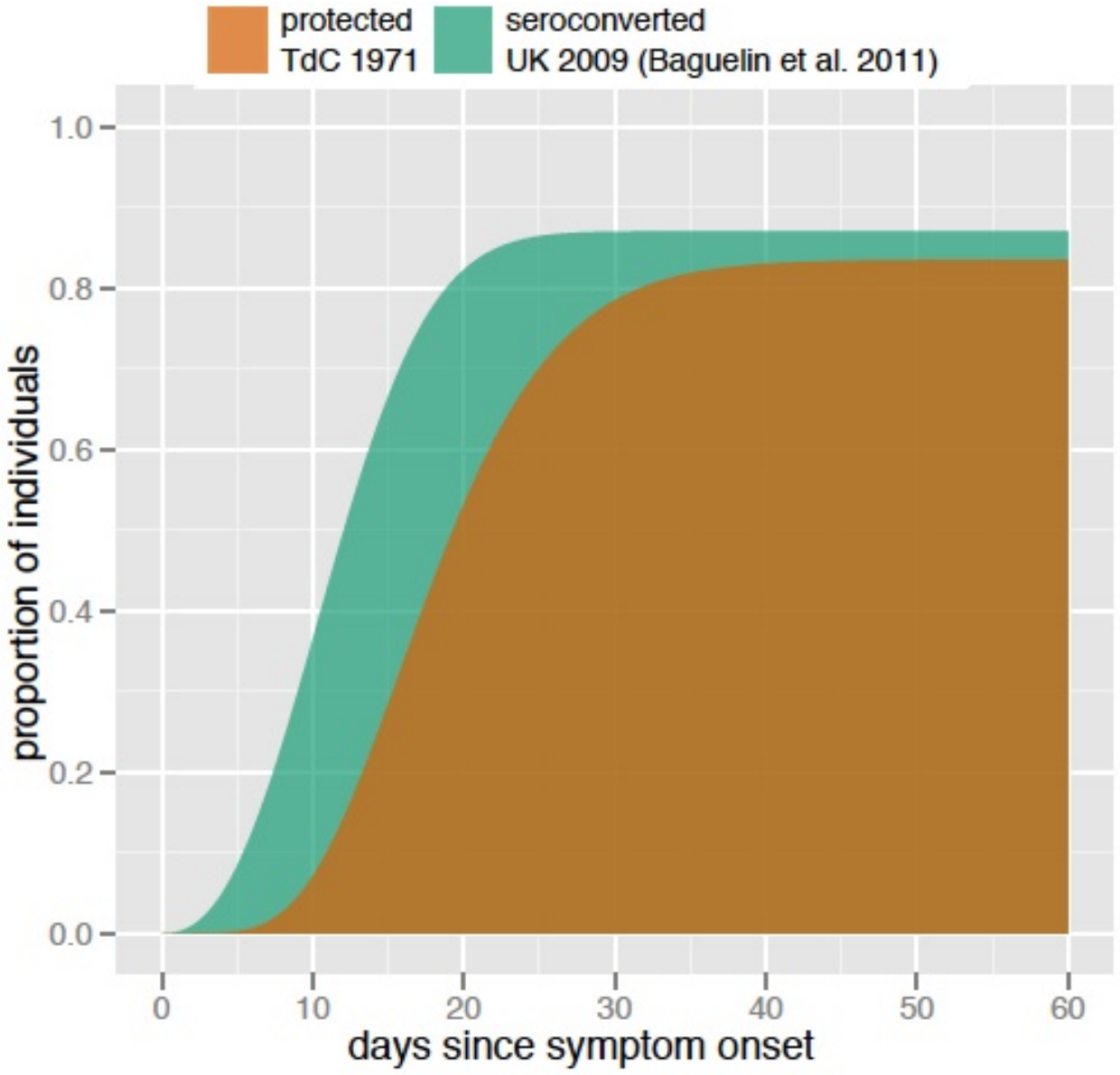
Accounting for immunodynamics in epidemiological models

Symbol	Description	Estimate	95% CI
$R_0 = \beta/\nu$	basic reproduction number	11.78	7.70 – 25.50
$1/\epsilon$	mean latent period (days)	2.18	1.53 – 2.96
$1/\nu$	mean infective period (days)	2.32	0.70 – 5.03
$1/\gamma$	mean temporary removed period (days)	13.37	10.37 – 16.31
$1/\omega$	mean duration of the reinfection window (days)	2.75	0 – 6.03
α	probability to develop long-term immunity	0.83	0.49 – 1
ρ	reporting rate for observation	0.71	0.62 – 0.82
I_0	number of initially infective individuals	1	1 – 3
S_0	number of initially susceptible individuals	277	275 – 280
$l(\theta_{\mathrm{ML}})$	maximized log-likelihood	-112.19	–

Accounting for immunodynamics in epidemiological models

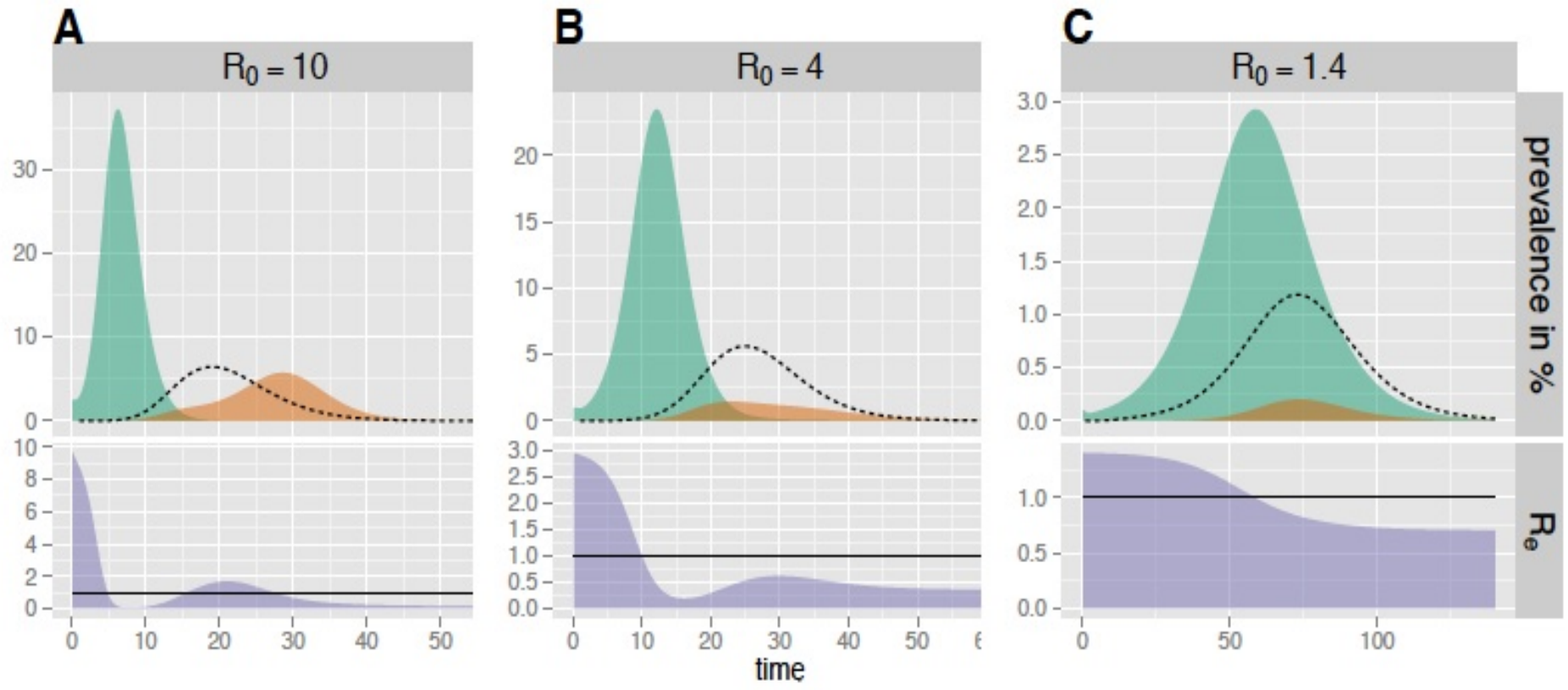


Accounting for immunodynamics in epidemiological models



Accounting for immunodynamics in epidemiological models

Interplay between the immunological and epidemiological dynamics



Accounting for immunodynamics in epidemiological models

- Host heterogeneity in the timely development of a protective immunity can explain reinfection.
- In TdC the reinfection wave was a natural consequence of the exceptional contact configuration and high susceptibility of this small and isolated community.
- In larger, less mixed and partially protected populations, reinfection alone can not generate multiple-wave outbreaks. But, this type of model can quantify the proportion of unprotected at the end of epidemics.