Time series analysis via maximum likelihood

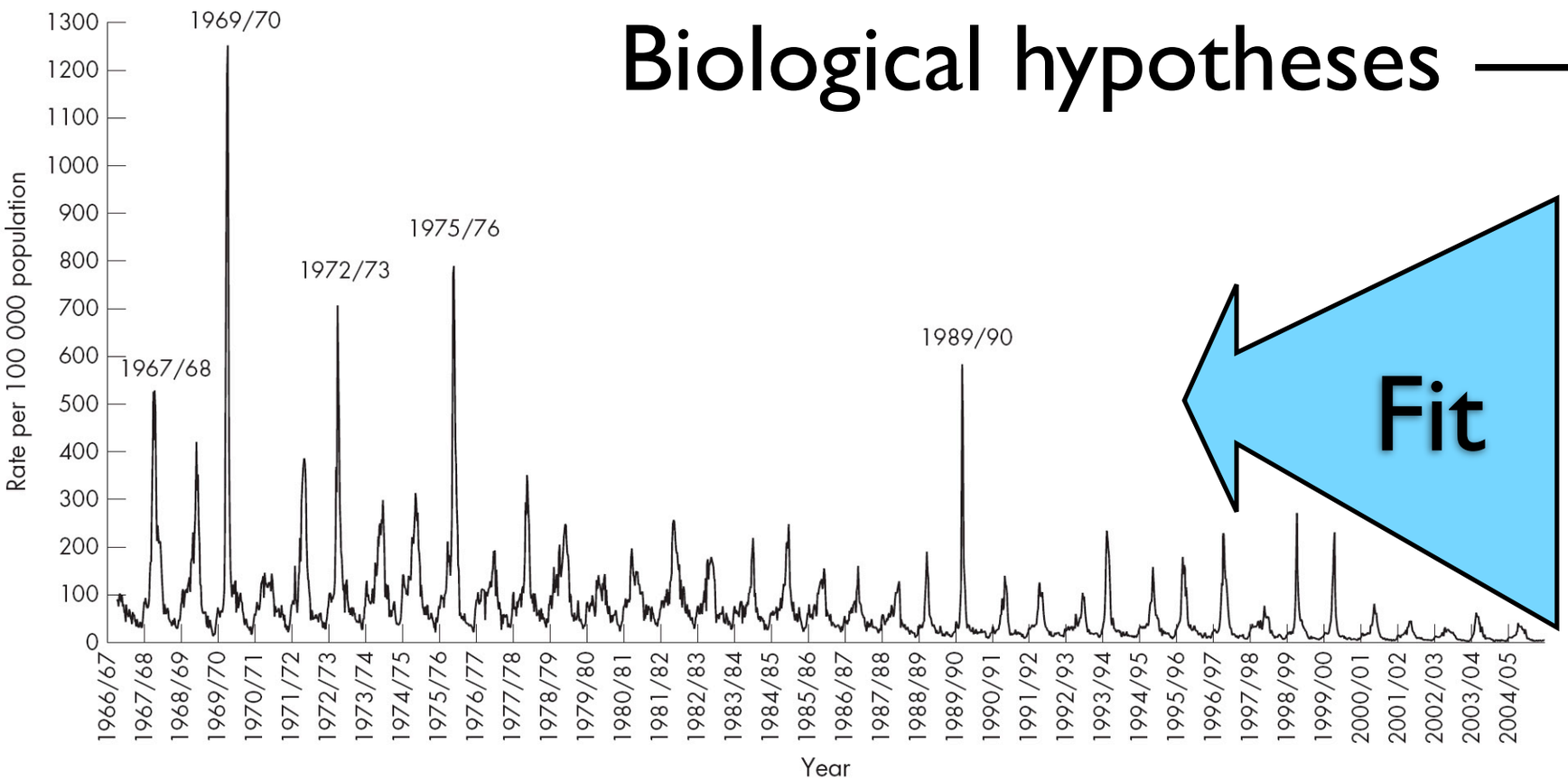
From theory to practice

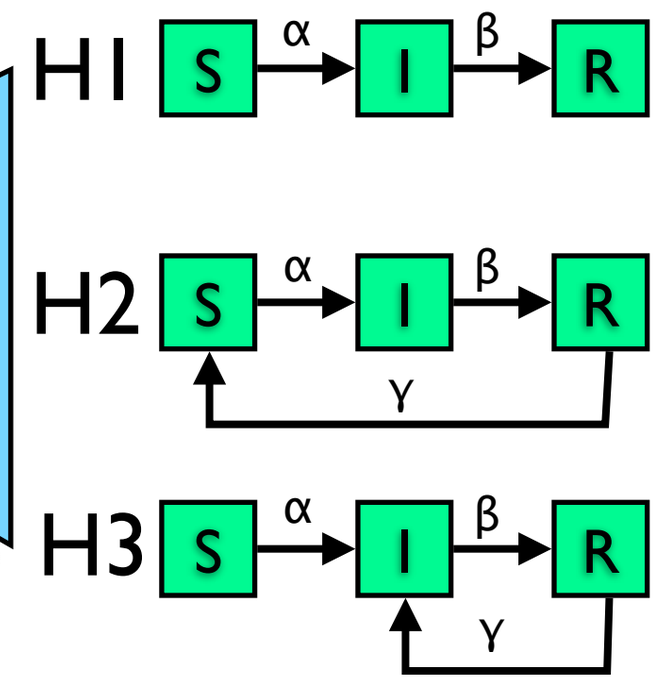
Anton Camacho, Sébastien Ballesteros & Bernard Cazelles

Department of Biology
UMR 7625, UPMC-CNRS-ENS
camacho@biologie.ens.fr

Time series analysis

Biological hypotheses ⟶ Models





Parameter inference:
Identifiability, maximum likelihood estimates, confidence intervals

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

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:

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

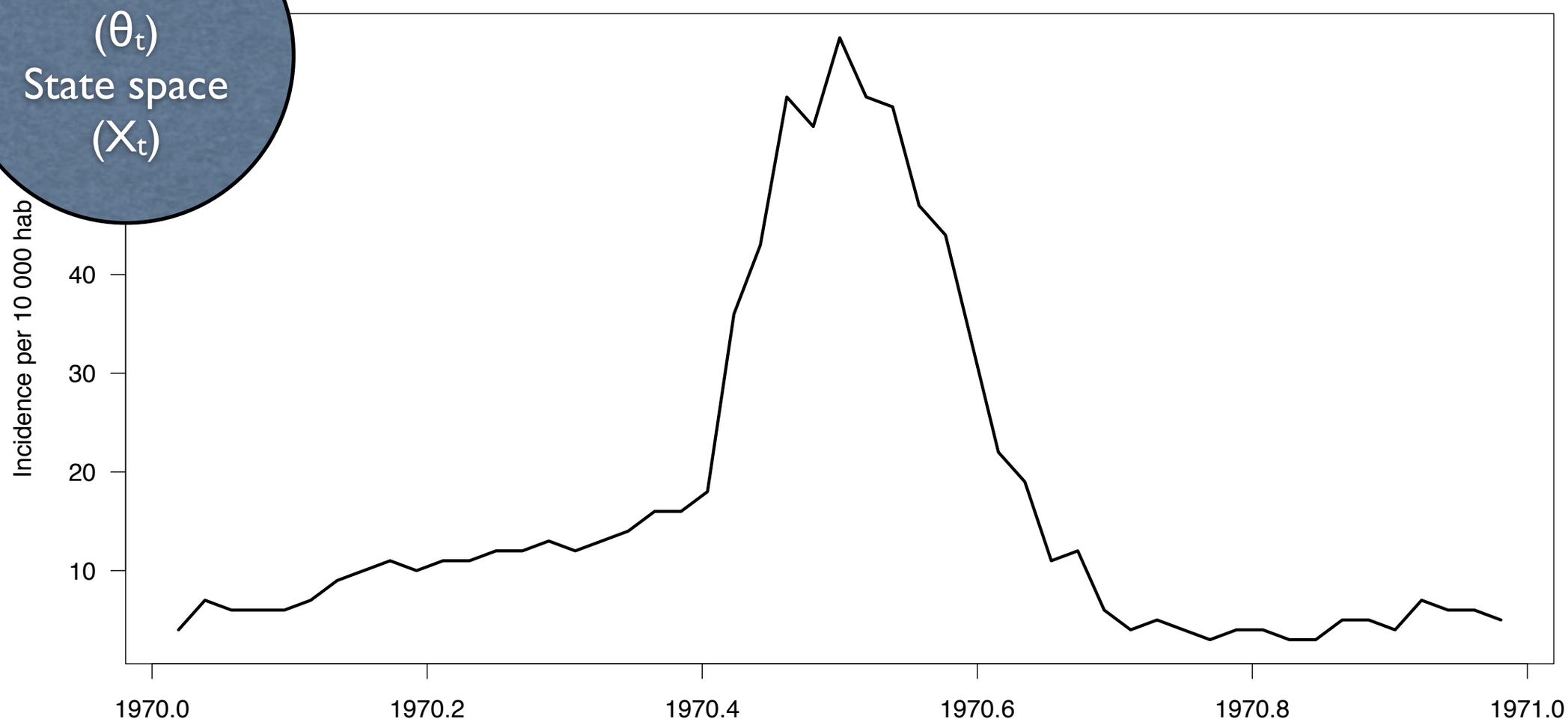
the likelihood is given by the identity:

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

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

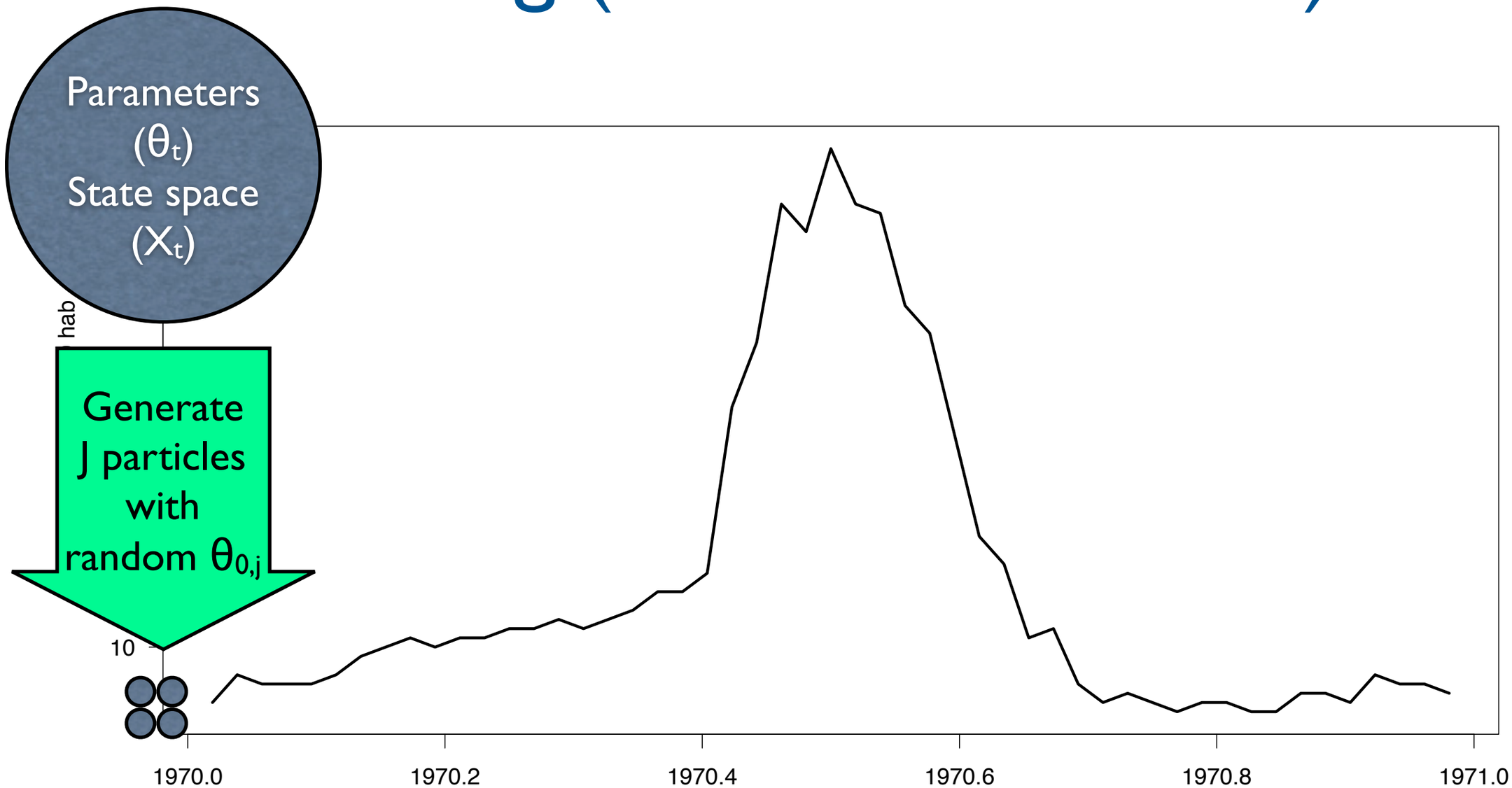
Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)

Parameters
(θ_t)
State space
(X_t)

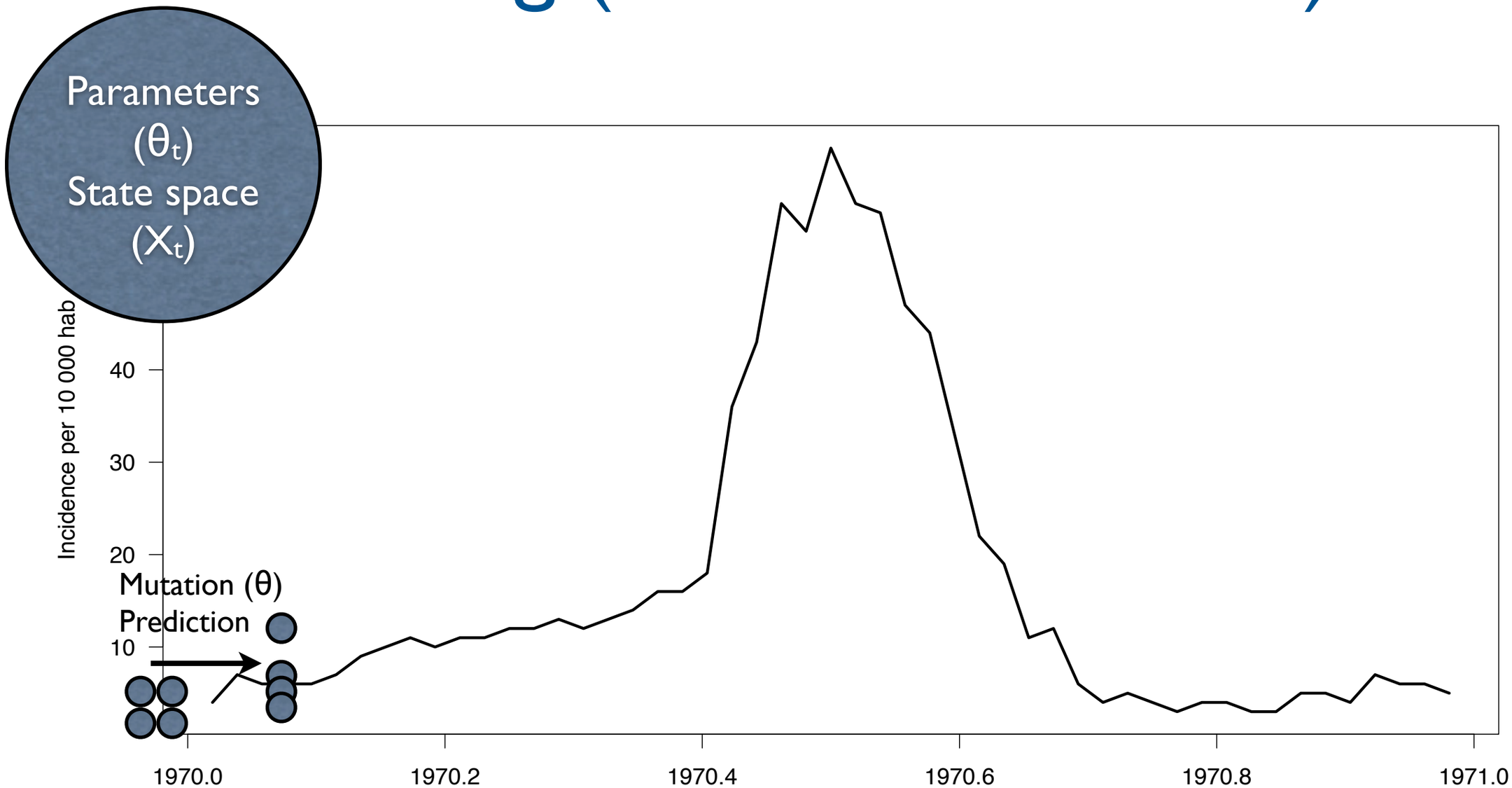


$$M \left\{ \begin{array}{l} f(x_t|x_{t-1},\theta) \\ f(y_t|x_t,\theta) \\ f(x_0|\theta) \end{array} \right. \quad \Rightarrow \quad M' \left\{ \begin{array}{l} f(x_t|x_{t-1},\theta_t) \\ f(y_t|x_t,\theta_t) \\ f(x_0|\theta_t) \end{array} \right.$$

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

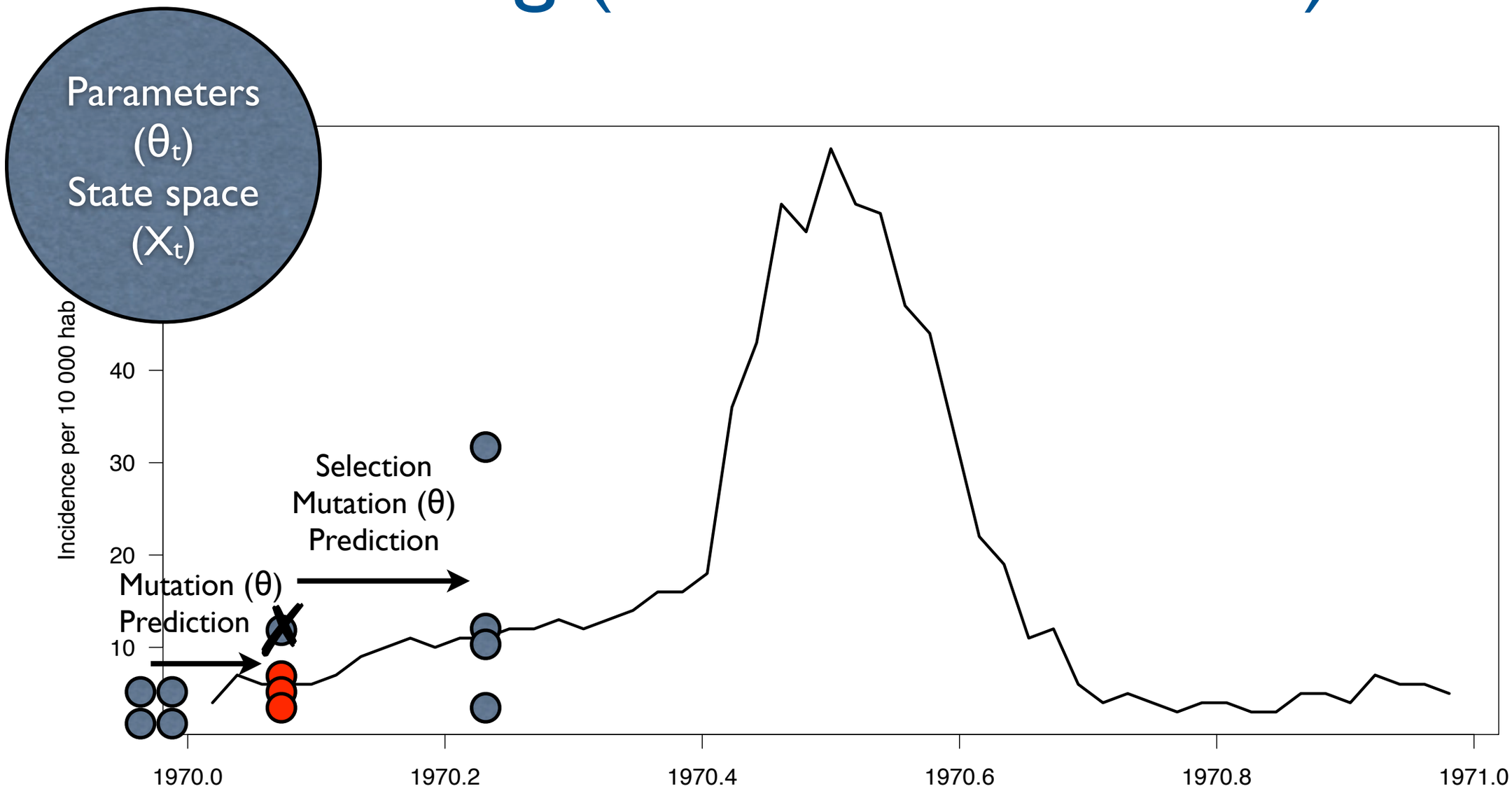


Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



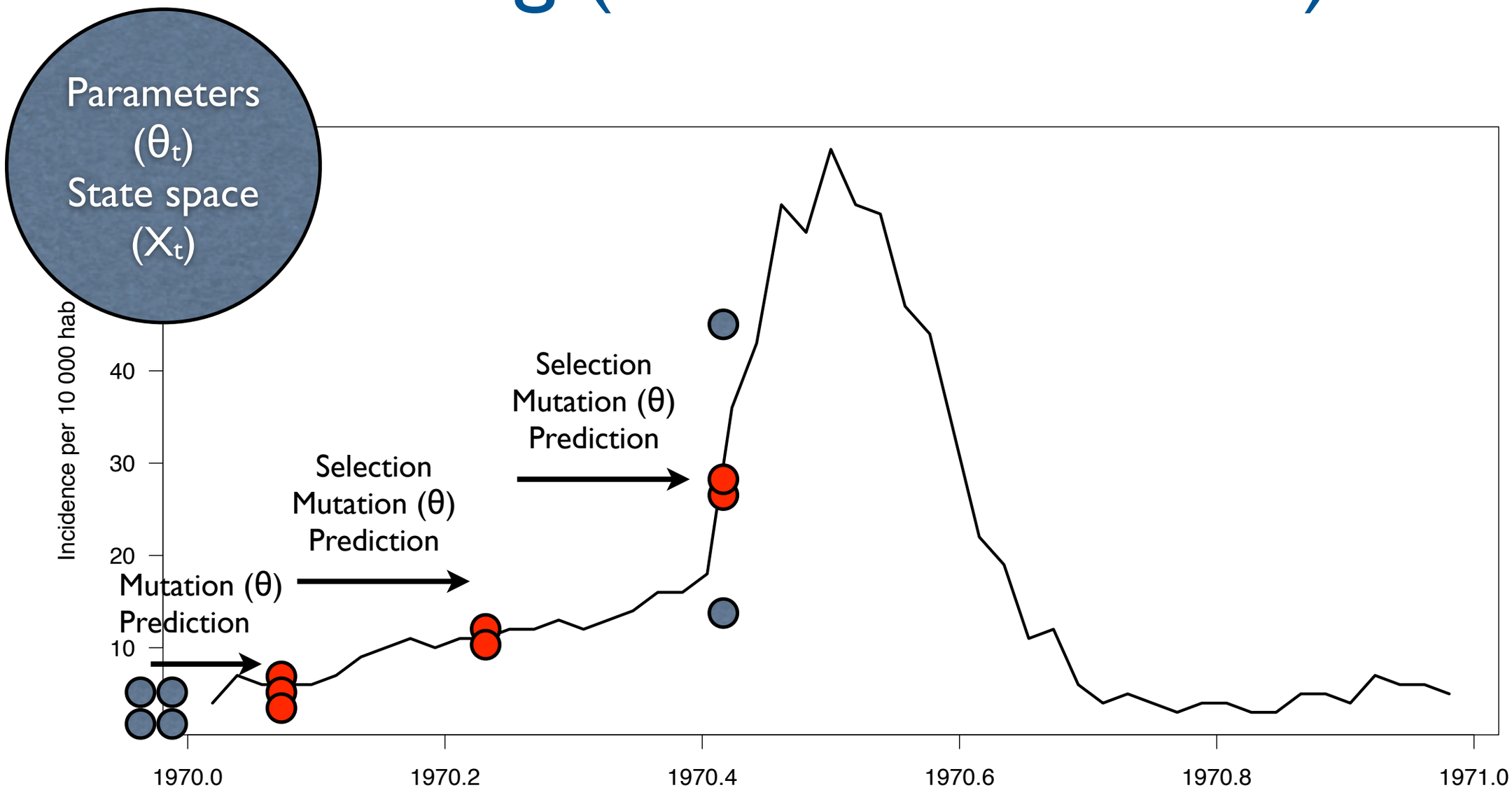
$$\text{For each particle } j: \begin{cases} X_{1,j} & \text{is drawn from} & f(x_1|x_0 = X_{0,j}, \theta_{0,j}) \\ \theta_{1,j} & \text{is drawn from} & \mathcal{N}(\theta_{0,j}, \sigma) \\ w_{1,j} & \text{is equal to} & f(y_1|x_1 = X_{1,j}, \theta_{1,j}) \end{cases}$$

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



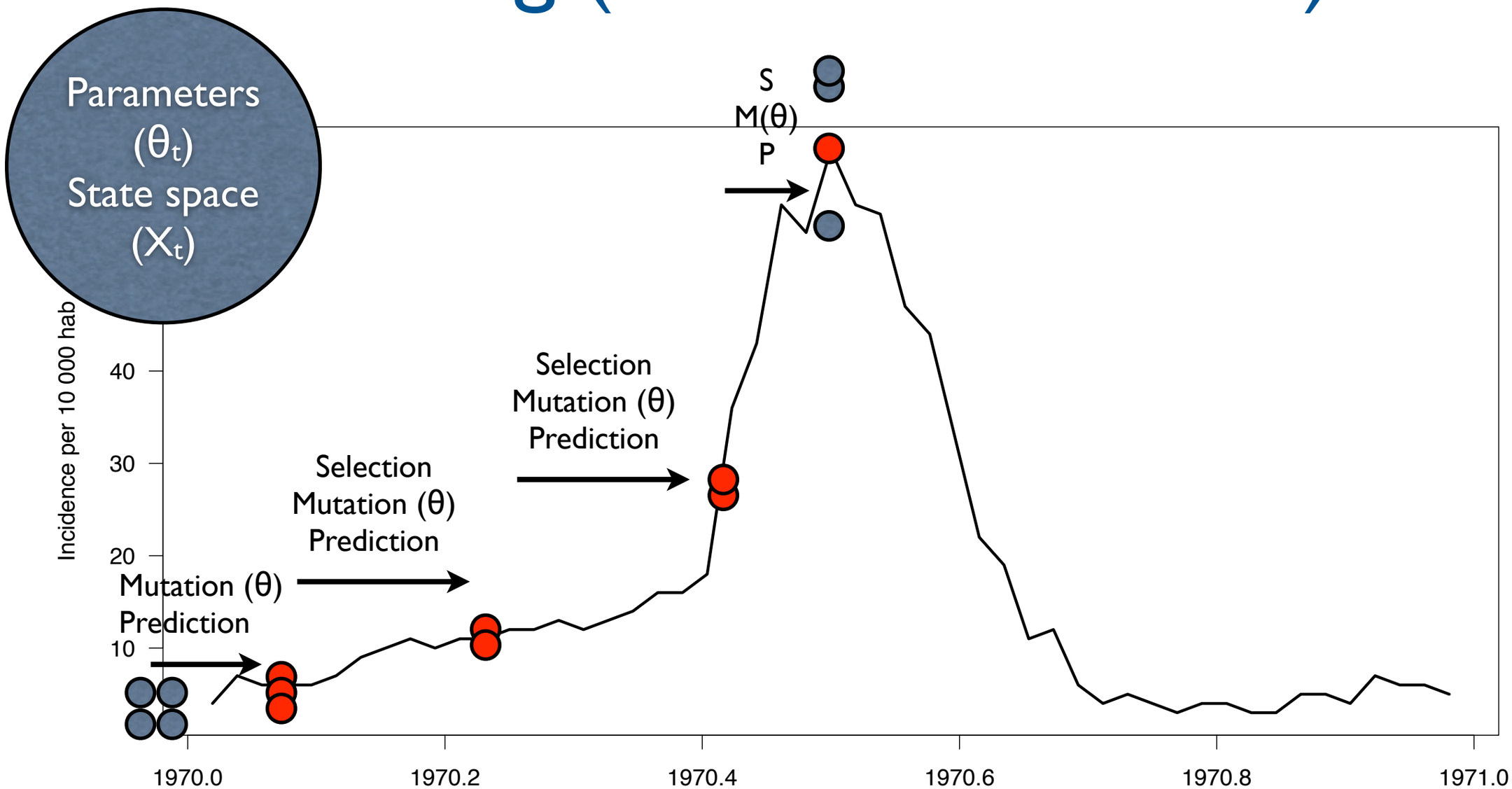
Darwinian selection: particles reproduce proportionally to their weight w_j

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



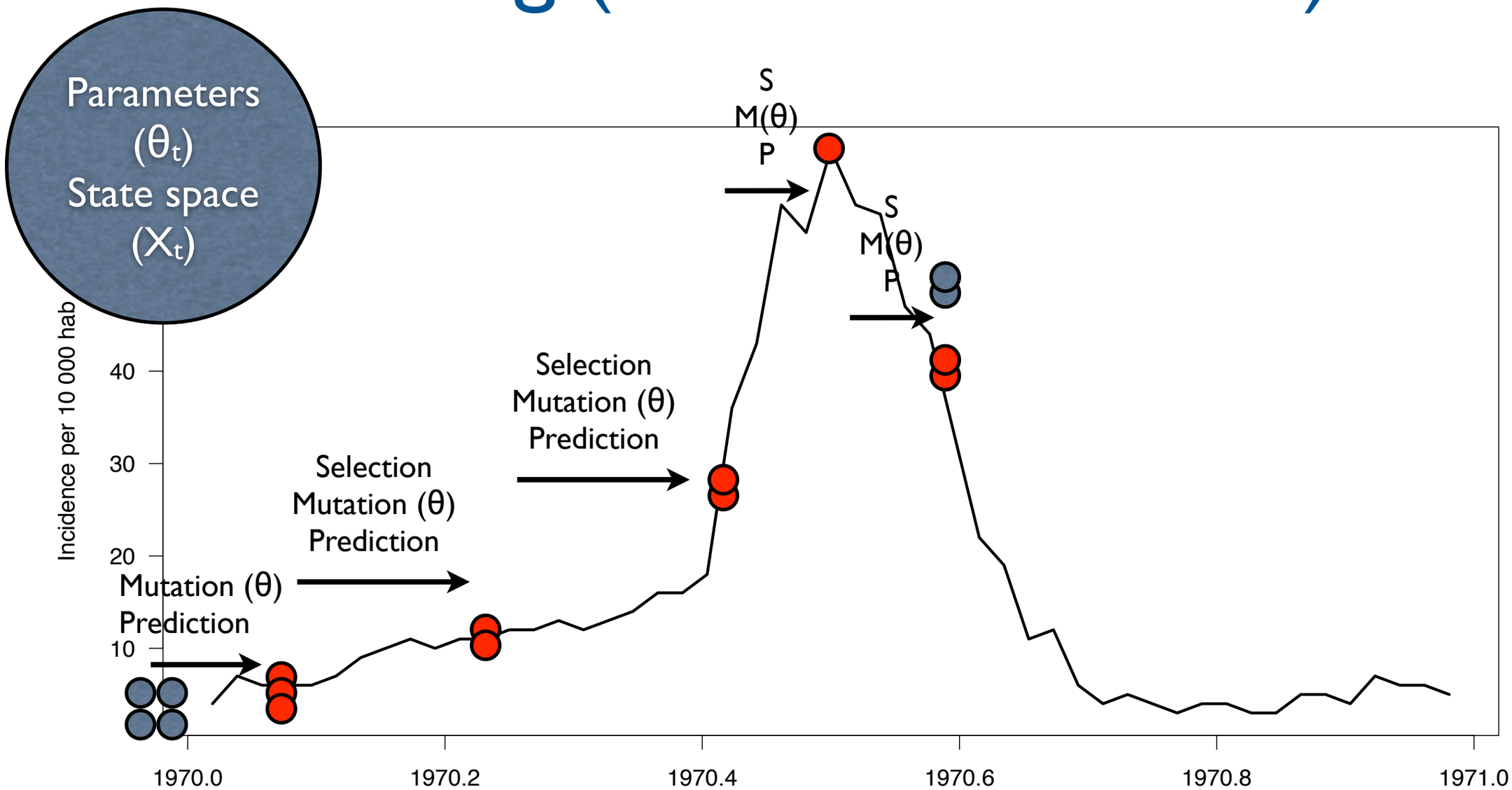
Selection + Mutation + Prediction

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



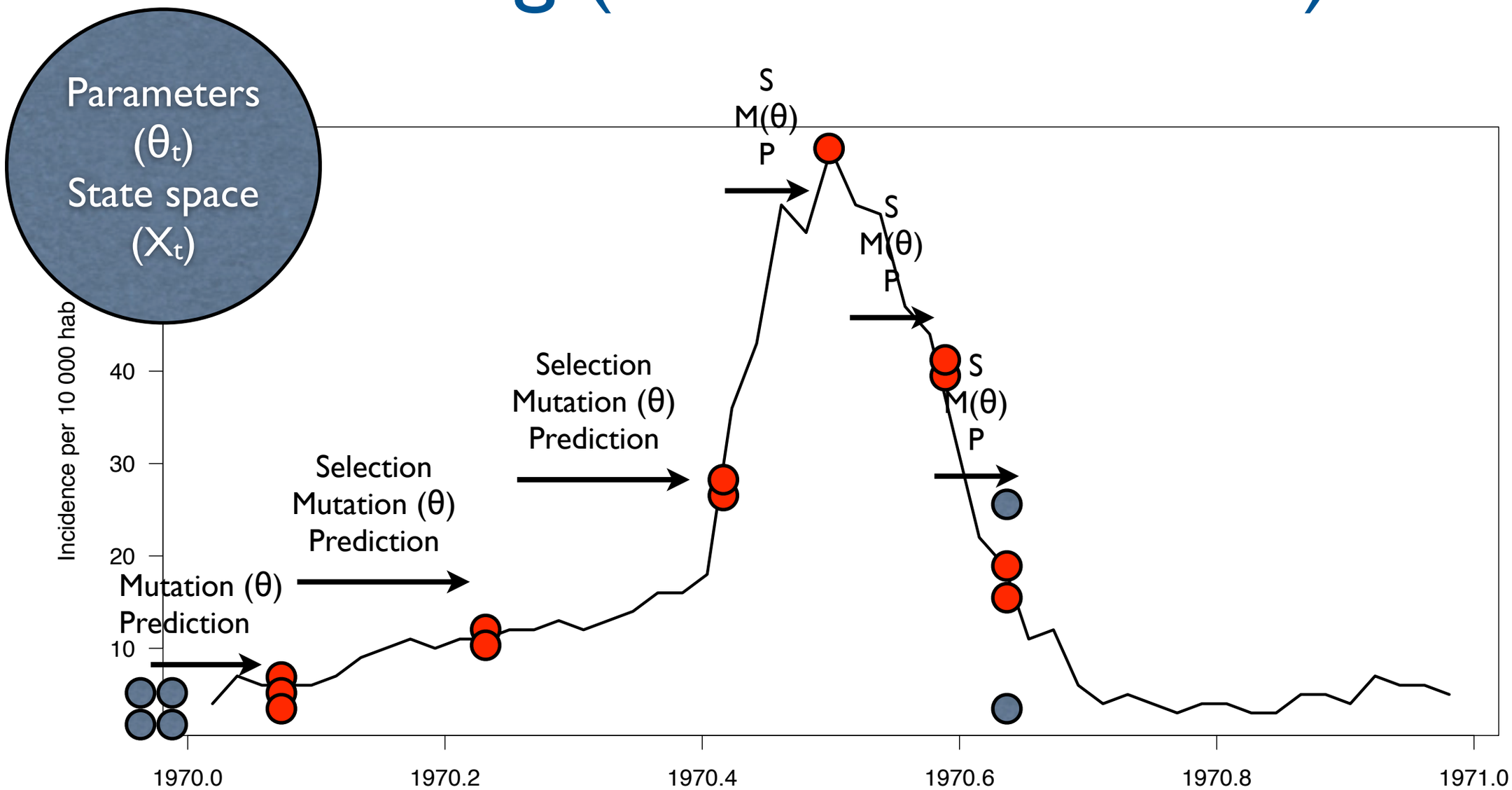
Selection + Mutation + Prediction

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



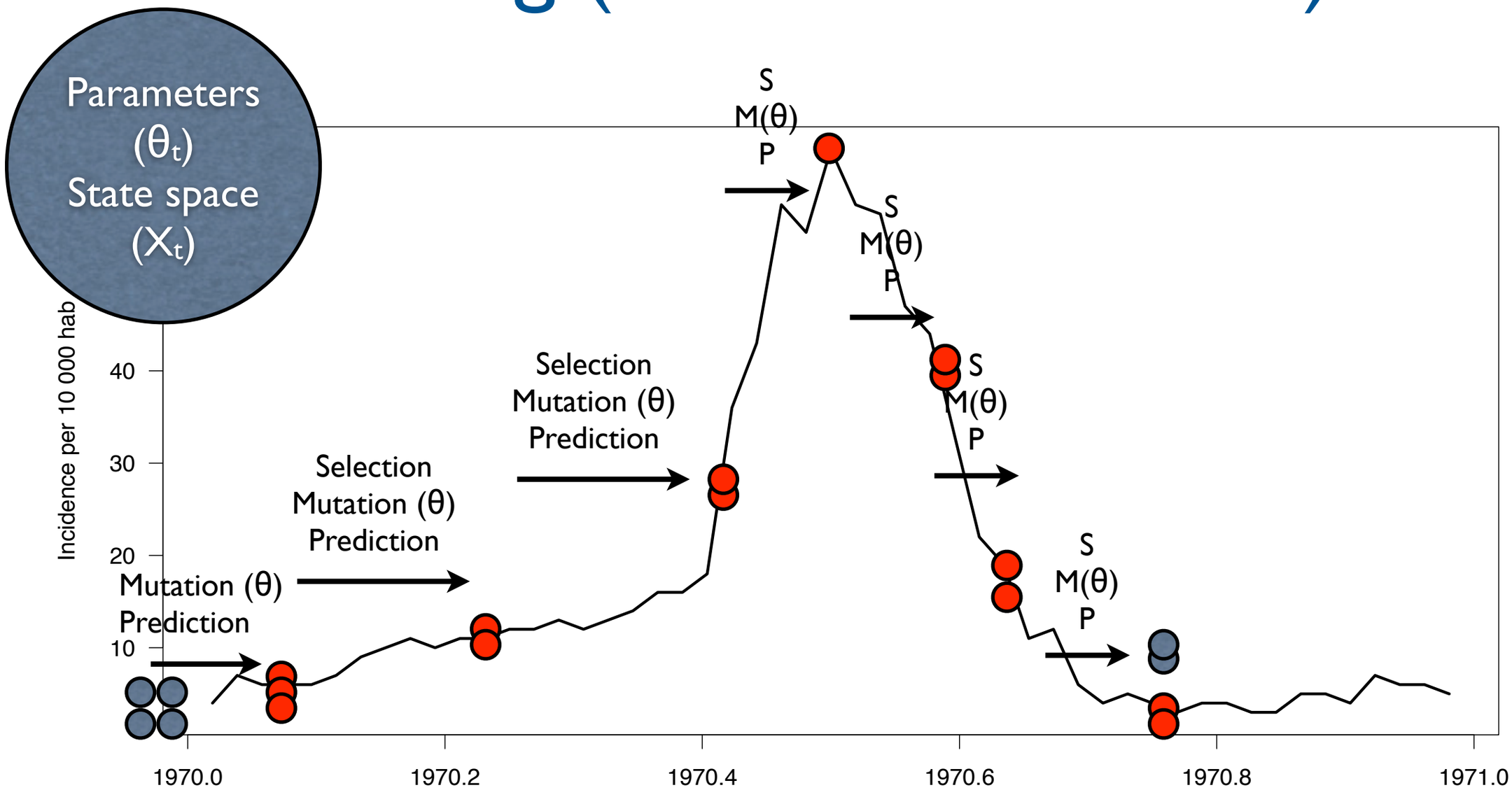
Selection + Mutation + Prediction

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



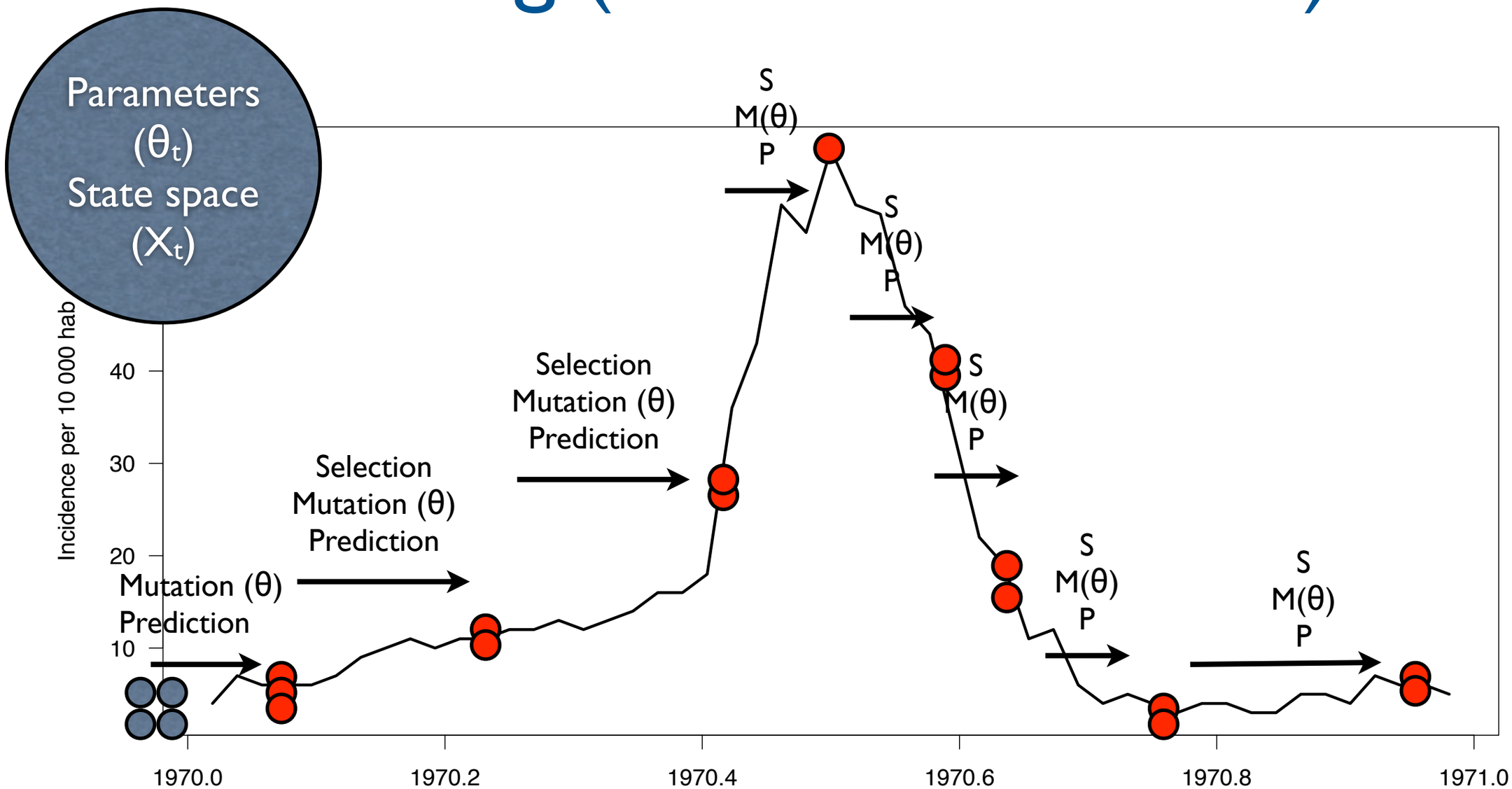
Selection + Mutation + Prediction

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



Selection + Mutation + Prediction

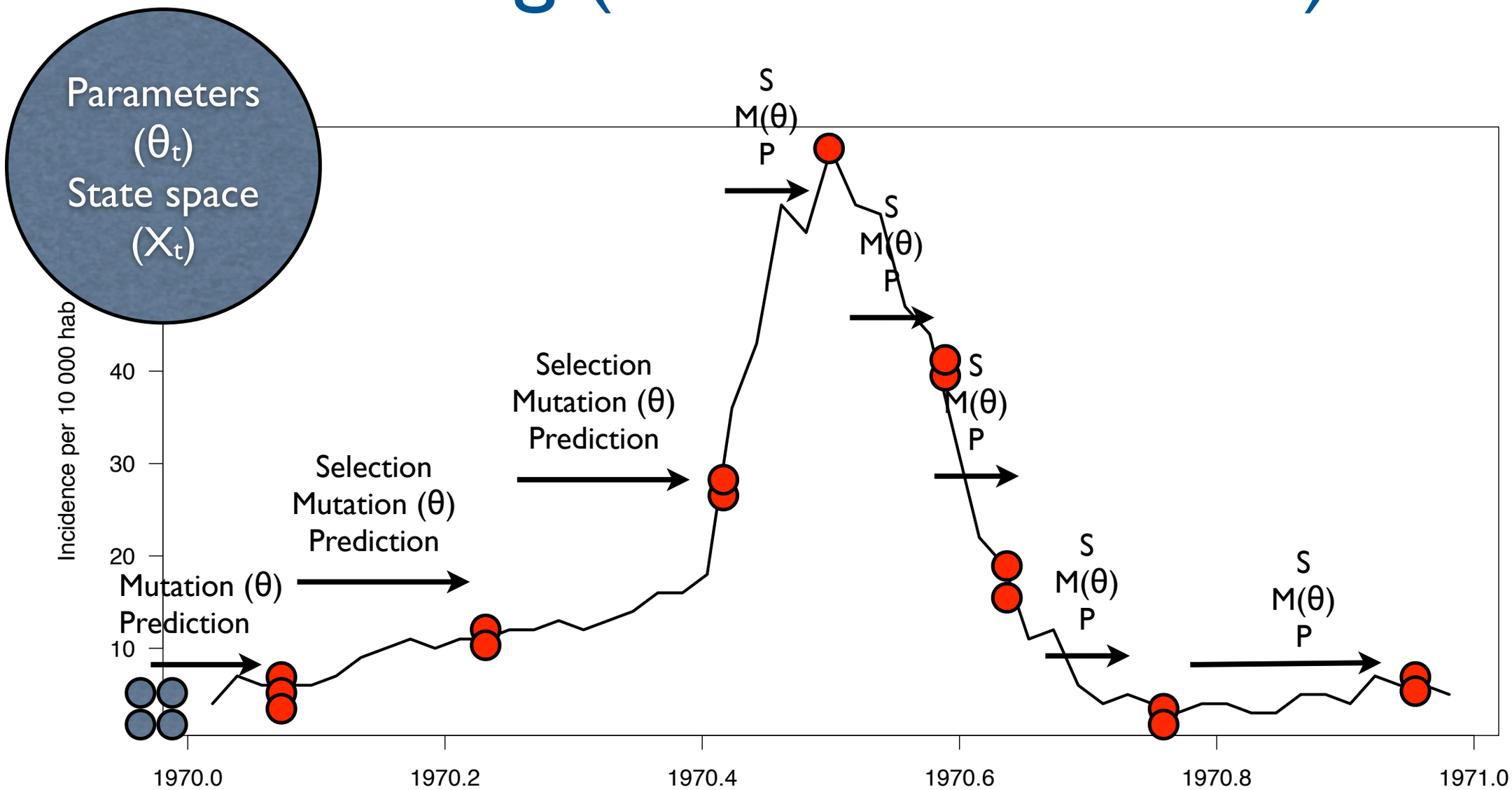
Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



Punctual estimates for each time t from $\{(\theta_{t,j}, w_{t,j})\}_J$

$$\begin{cases} E[\theta_t | y_{1:t}] & \text{by} \quad \hat{\theta}_t \\ Var(\theta_t | y_{1:t-1}) & \text{by} \quad V_t \\ f(y_t | y_{1:t-1}, \theta) & \text{by} \quad l_t(\theta) = \frac{1}{J} \sum_J w_{t,j} \end{cases}$$

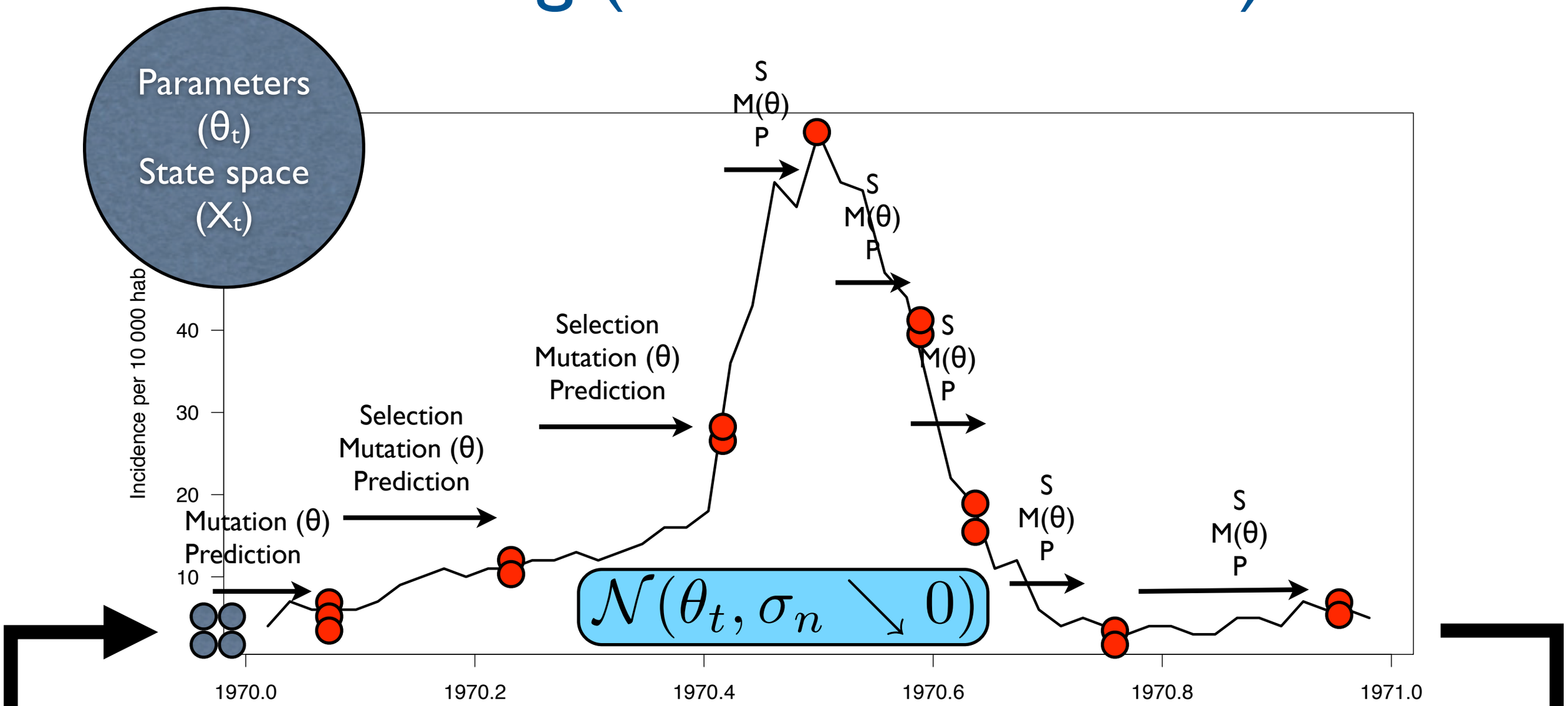
Maximum likelihood via Iterated Filtering (Ionides *et al.* 2006)



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

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

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

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

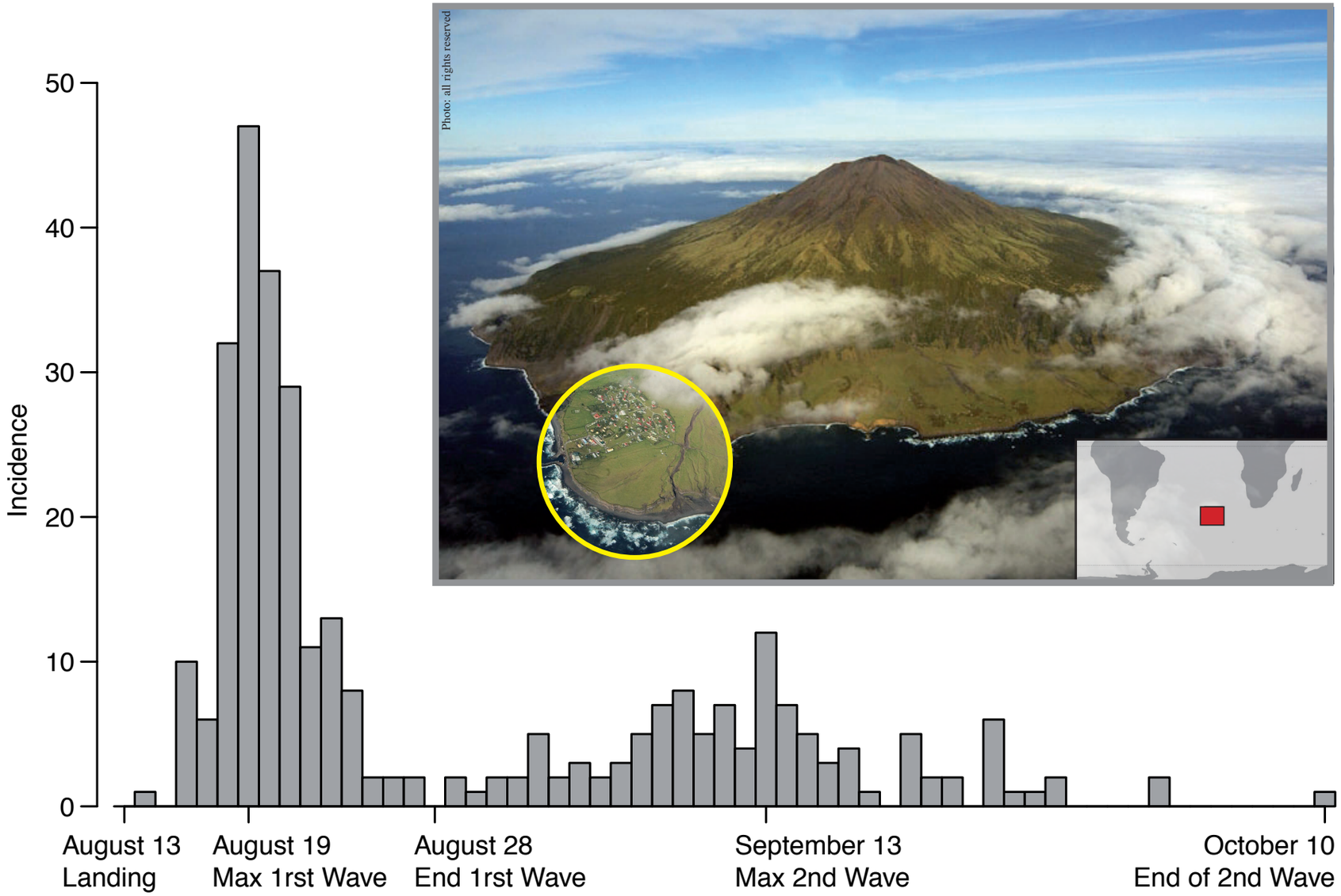
Global estimate: $\hat{\theta}^{(n)} = \hat{\theta}^{(n-1)} + V_1^{(n)} \sum_{t=1}^{T} \frac{\hat{\theta}_t^{(n)} - \hat{\theta}_{t-1}^{(n)}}{V_t^{(n)}}$

As shown by Ionides *et al.* (2006), under rather mild assumptions,

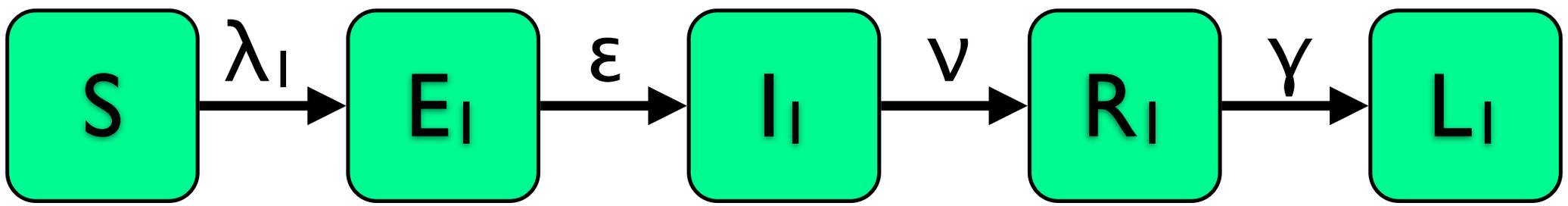
$$\lim_{\sigma \to 0} \sum_{t=1}^{T} \frac{\hat{\theta}_t - \hat{\theta}_{t-1}}{V_t} = \nabla \log f(y_{1:T} | \theta, \sigma = 0)$$

so that, for a sufficiently small σ_n, the algorithm iteratively updates $\hat{\theta}^{(n)}$ in the direction of increasing likelihood, with a fixed point at a local maximum of the likelihood.

A two-wave epidemic on Tristan da Cunha (1971)

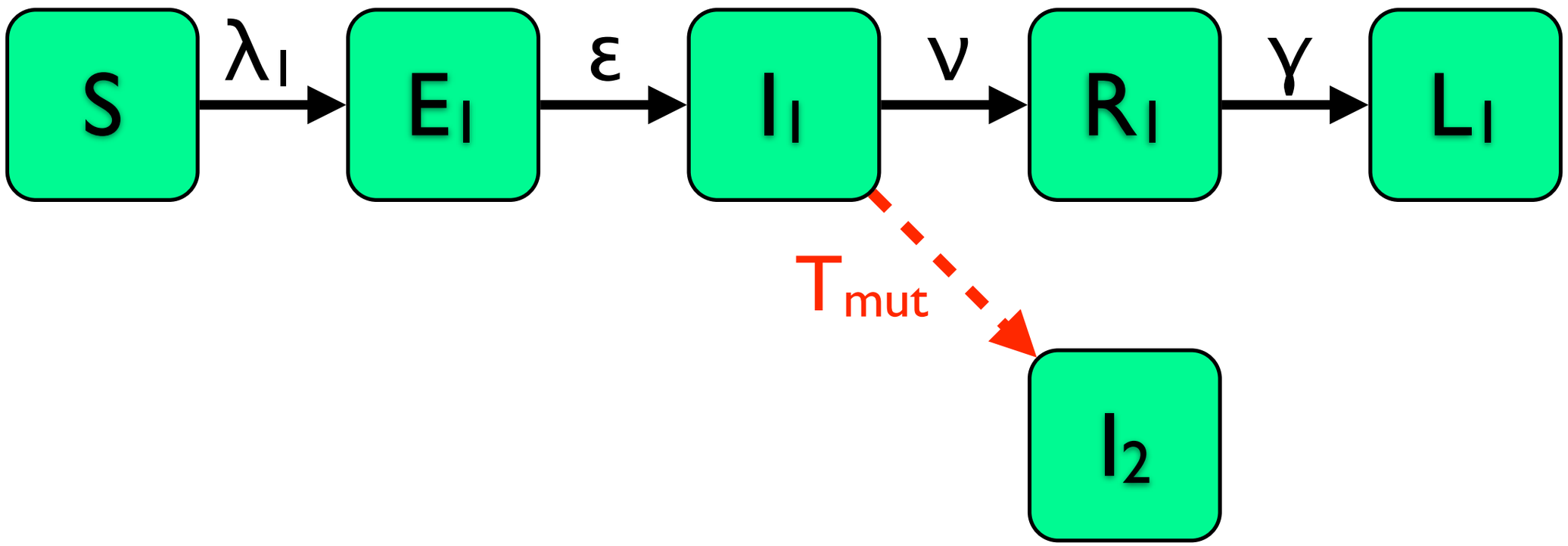


H1: the virus mutated during the first epidemic-wave

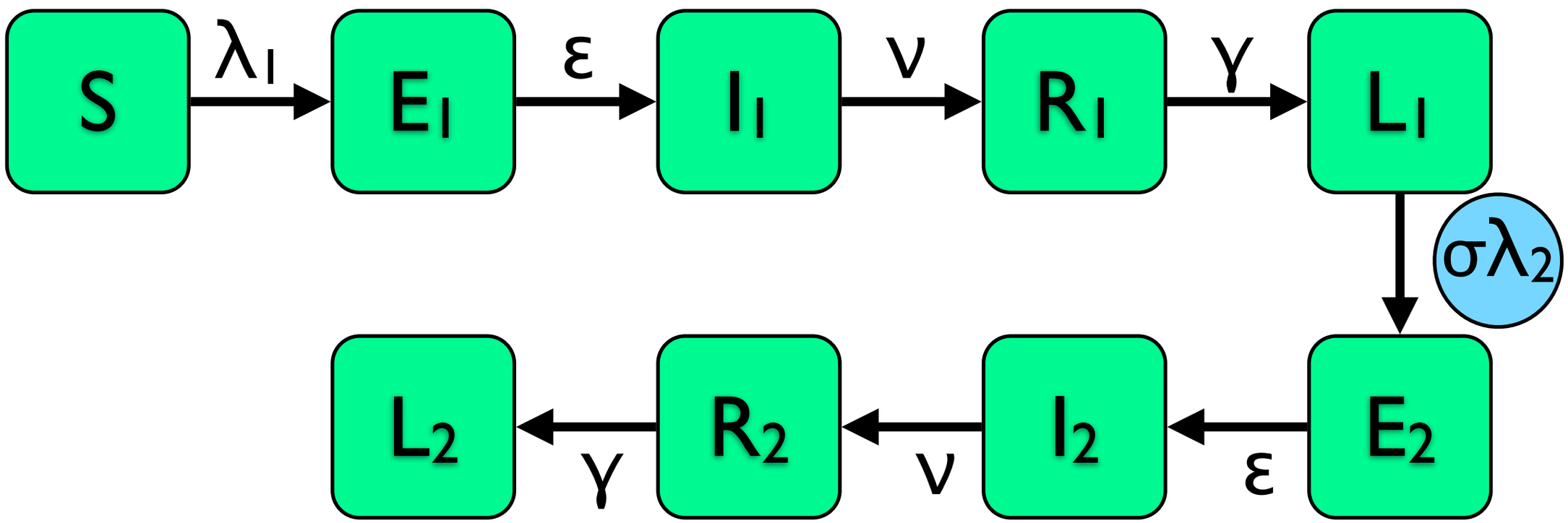


- $\lambda_I = \beta_I I_I/N$ mass-action
- $1/\varepsilon$: mean latent period
- $1/\nu$: mean infectious period
- $1/\gamma$: mean resistance period

H1: the virus mutated during the first epidemic-wave



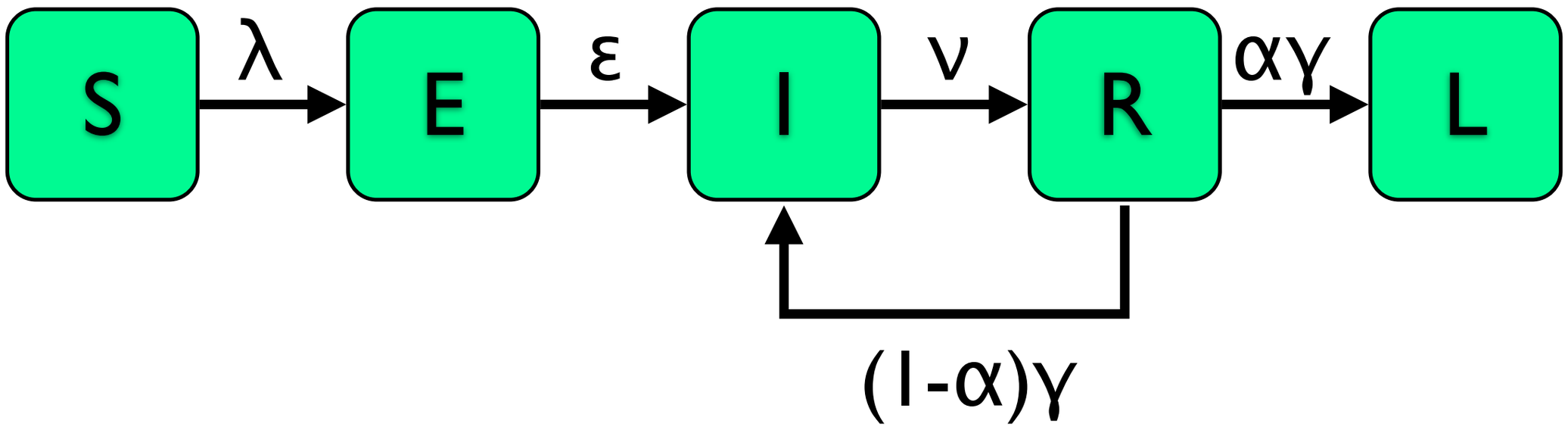
H1: the virus mutated during the first epidemic-wave



- $\lambda_i = \beta_i I_i/N$ mass-action
- $\sigma \in [0,1]$ cross-immunity

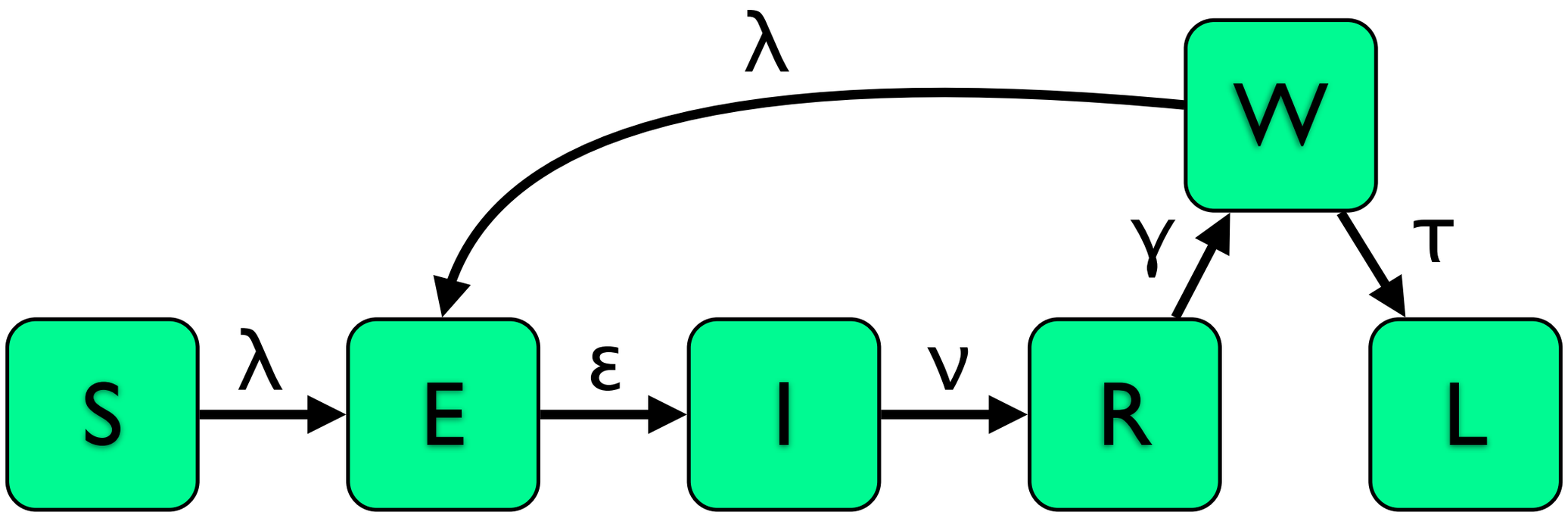
Structural non-identifiability between σ and $\beta_2 \Rightarrow \beta_2 = \beta_1$

H2: intra-host recrudescence of infection



α: the probability to clear the viral load

H3: window of reinfection

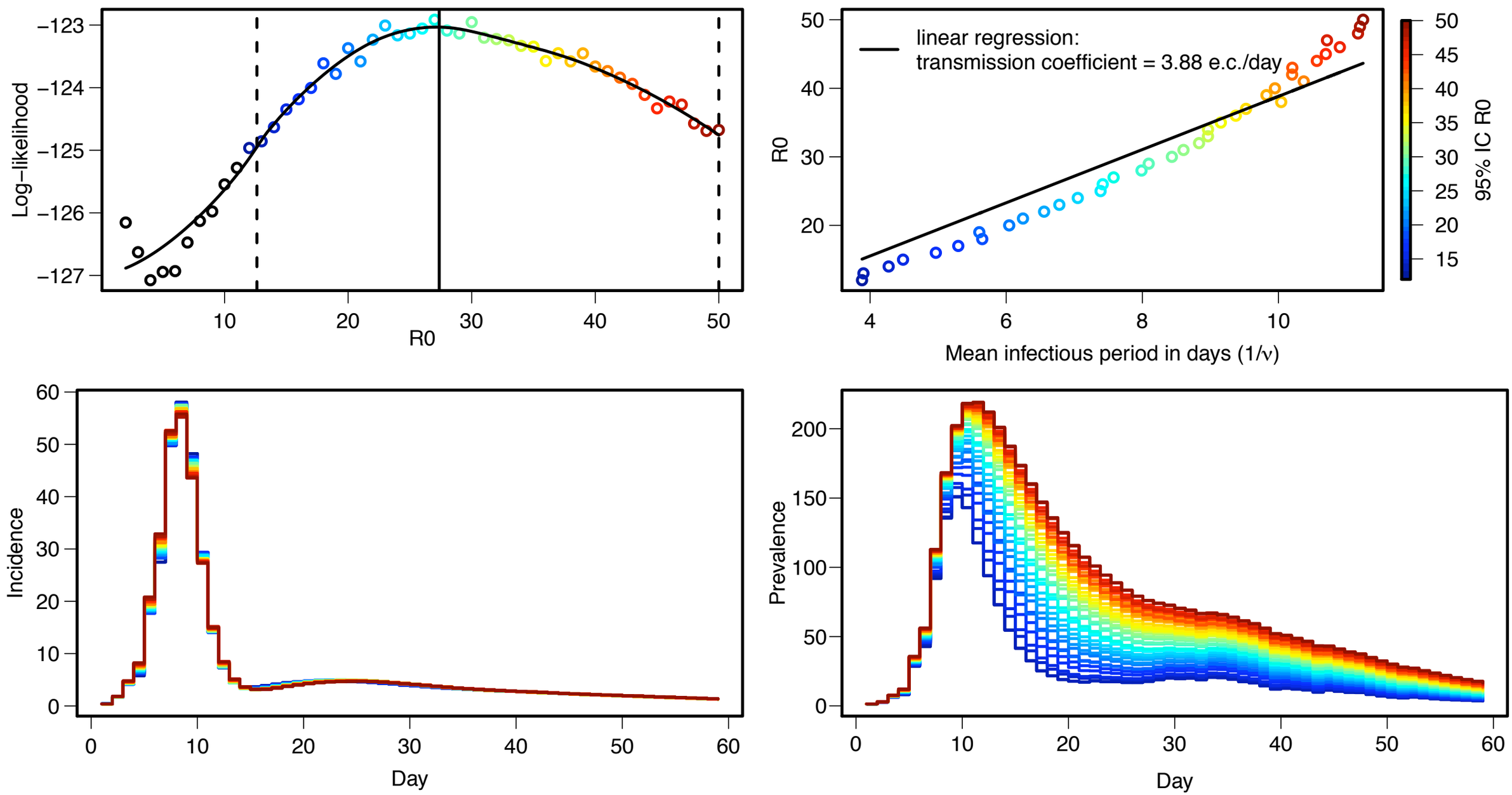


$1/\tau$: the mean duration of the window of susceptibility before developing immunity

Parameter inference

Parameter	Reproduction number ($R_0 = \beta/\nu$)	Infectious period in days ($1/\nu$)
Mut	9.07 [5.05 - 16.39]	0.98 [0.12 - 2.86]
In-Host	25.01 [12.62 - 50*]	5.18 [2.68 - 11*]
Win	10.38 [6.14 - 17.06]	2.01 [0.1* - 4.20]

Practical non-identifiability (In-Host)



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	Mut	In-Host	Win
Log-likelihood	-115.20	-117.50	-112.52
AIC_c	254.98	256.67	246.71
ΔAIC_c	8.27	9.96	0

Conclusion

- Maximum likelihood via Iterated Filtering (MIF, Ionides *et al.* 2006) is a rigorous statistical framework for parameter inference and selection based on AIC for non-linear stochastic models.
- Identifiability analysis and 95% CI via log-likelihood profile
- It could be interesting to compare the results of MIF on Tristan da Cunha with a likelihood-free based statistical framework (ex: ABC)

(1) Ionides EL, Breto CM, King AA. Inference for nonlinear dynamical systems. *PNAS*. 2006

(2) Raue A, Kreutz C, Maiwald T, et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics (Oxford, England)*. 2009

(3) Camacho A, Ballesteros S, Graham AL, Carrat F, Cazelles B. Does reinfection drive multiple-wave influenza outbreaks? Tristan da Cunha epidemic as a case study. *(in preparation)*

Mean behaviour