Probabilistic projections of HIV prevalence using Bayesian melding

Leontine Alkema, Adrian Raftery and Sam Clark

Department of Statistics, University of Washington, Seattle, USA

Outline

- HIV projections

Outline

- HIV projections
- Bayesian melding

Outline

- HIV projections
- Bayesian melding
- Results for Uganda

Outline

- HIV projections
- Bayesian melding
- Results for Uganda
- Summary

Outline

- HIV projections
- Bayesian melding
- Results for Uganda
- Summary

HIV projections

- HIV/AIDS statistics in 2005:
  40 million infected, 3 million died, 5 million newly infected

HIV projections

- HIV/AIDS statistics in 2005:
  40 million infected, 3 million died, 5 million newly infected
- HIV/AIDS is widely spread in sub-Saharan Africa

HIV projections

- HIV/AIDS statistics in 2005:
  40 million infected, 3 million died, 5 million newly infected
- HIV/AIDS is widely spread in sub-Saharan Africa
- The epidemic has a huge effect on populations

HIV projections

- HIV/AIDS statistics in 2005:
  40 million infected, 3 million died, 5 million newly infected
- HIV/AIDS is widely spread in sub-Saharan Africa
- The epidemic has a huge effect on populations
- Need for projections of HIV prevalence

HIV projections

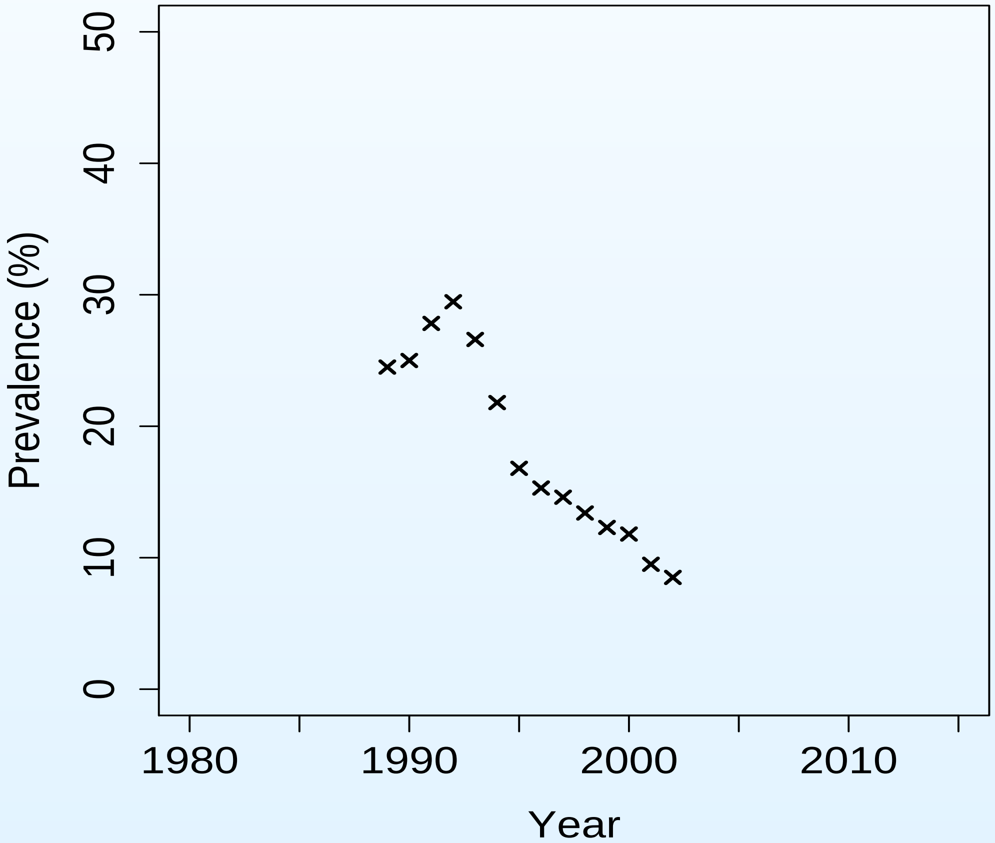
- HIV/AIDS statistics in 2005:
  40 million infected, 3 million died, 5 million newly infected
- HIV/AIDS is widely spread in sub-Saharan Africa
- The epidemic has a huge effect on populations
- Need for projections of HIV prevalence
- United Nations program on HIV/AIDS (UNAIDS)
  produces estimates and projections,
  using the Estimation and Projection Package (EPP)

HIV projections

- HIV/AIDS statistics in 2005:
  40 million infected, 3 million died, 5 million newly infected
- HIV/AIDS is widely spread in sub-Saharan Africa
- The epidemic has a huge effect on populations
- Need for projections of HIV prevalence
- United Nations program on HIV/AIDS (UNAIDS) produces estimates and projections, using the Estimation and Projection Package (EPP)
- EPP is designed to fit an epidemic curve to various epidemics based on little data

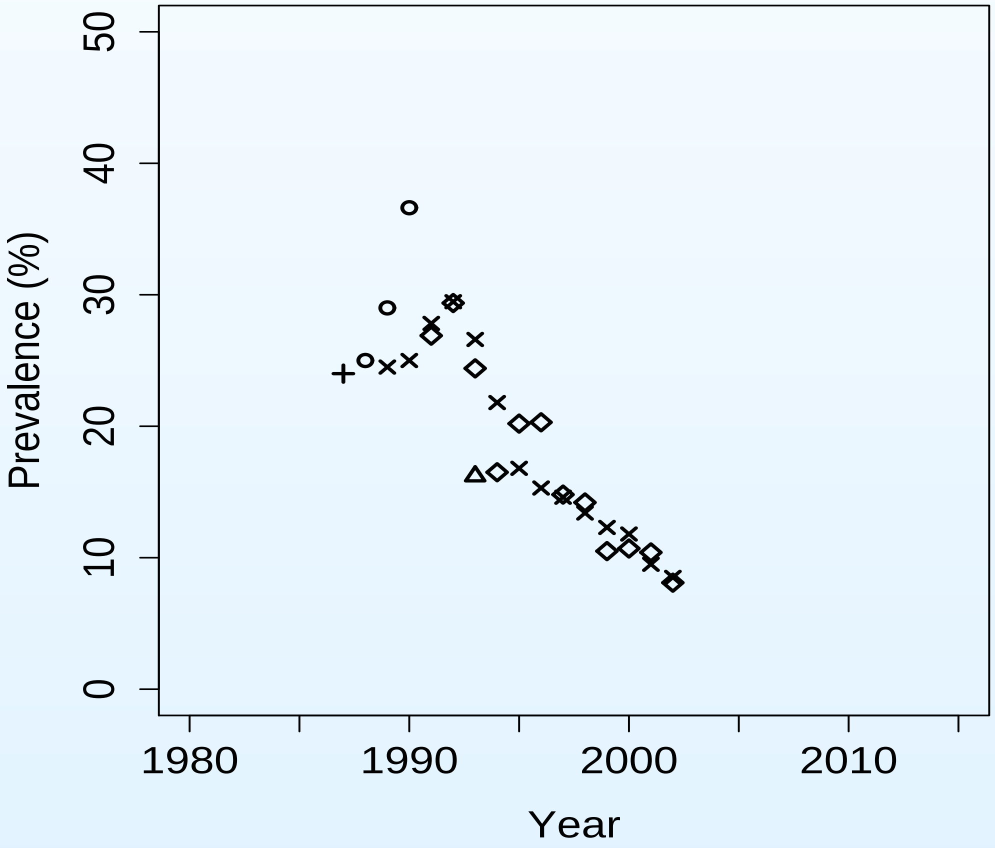
HIV prevalence prediction for Uganda urban

- Observed HIV prevalence in five urban antenatal clinics



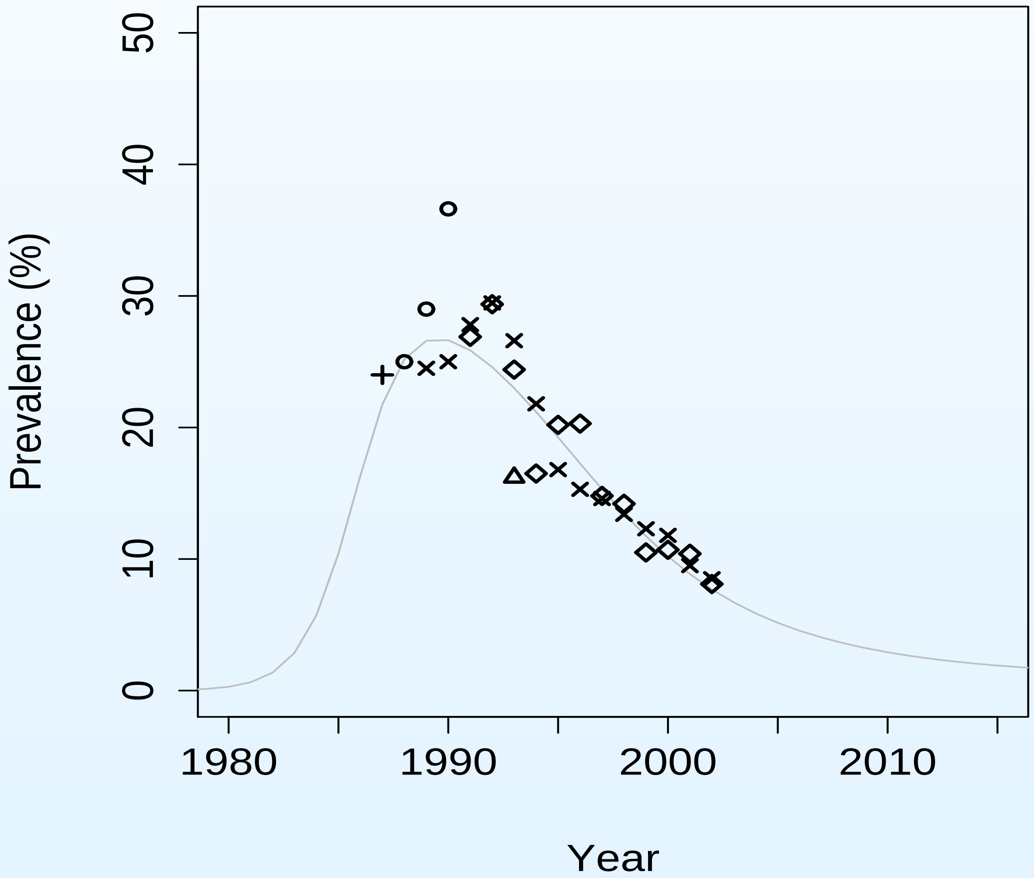
HIV prevalence prediction for Uganda urban

- Observed HIV prevalence in five urban antenatal clinics



HIV prevalence prediction for Uganda urban

- Observed HIV prevalence in five urban antenatal clinics
- The EPP model fits an epidemic curve



UNAIDS EPP model

- The population 15-49 is divided into three groups: Not at risk of getting infected, at risk, infected

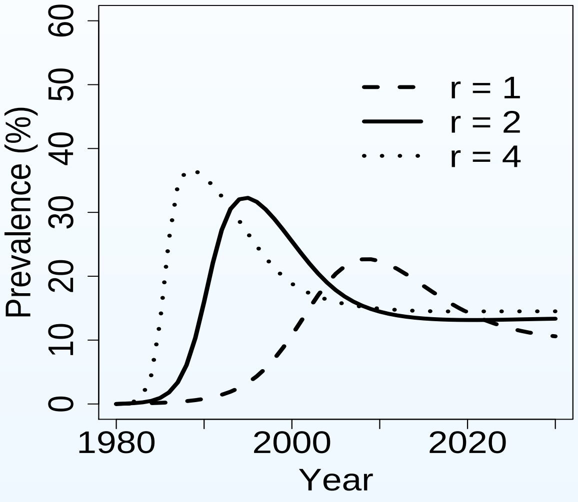
UNAIDS EPP model

- The population 15-49 is divided into three groups:
  Not at risk of getting infected, at risk, infected
- Three differential equations describe the changes in those groups over time, thus in prevalence over time

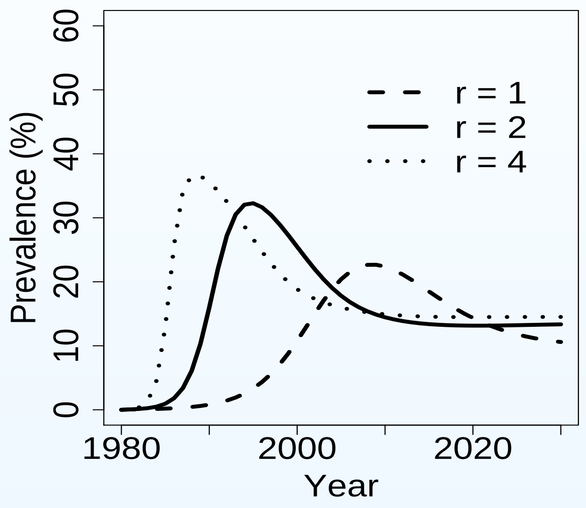
UNAIDS EPP model

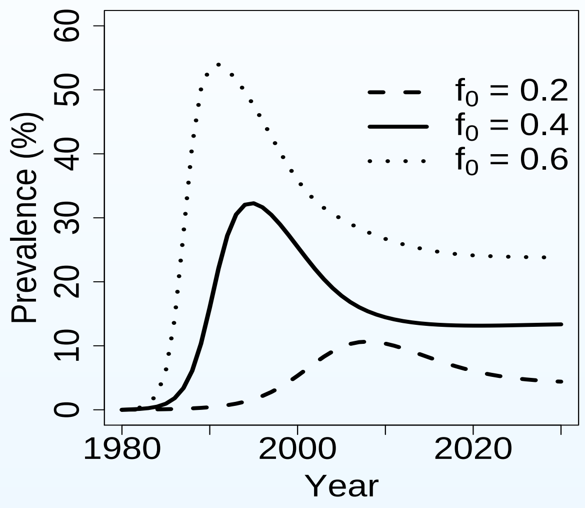
- The population 15-49 is divided into three groups: Not at risk of getting infected, at risk, infected
- Three differential equations describe the changes in those groups over time, thus in prevalence over time
- Four parameters determine the shape of the epidemic curve:
  - r = Growth rate of the epidemic
  - f_0 = Fraction of population initially at risk
  - t_0 = Start time epidemic
  - ϕ = Behavioral response

Input parameters (Solid curve: $r = 2, f_0 = 0.4, t_0 = 1980, \phi = 0$)

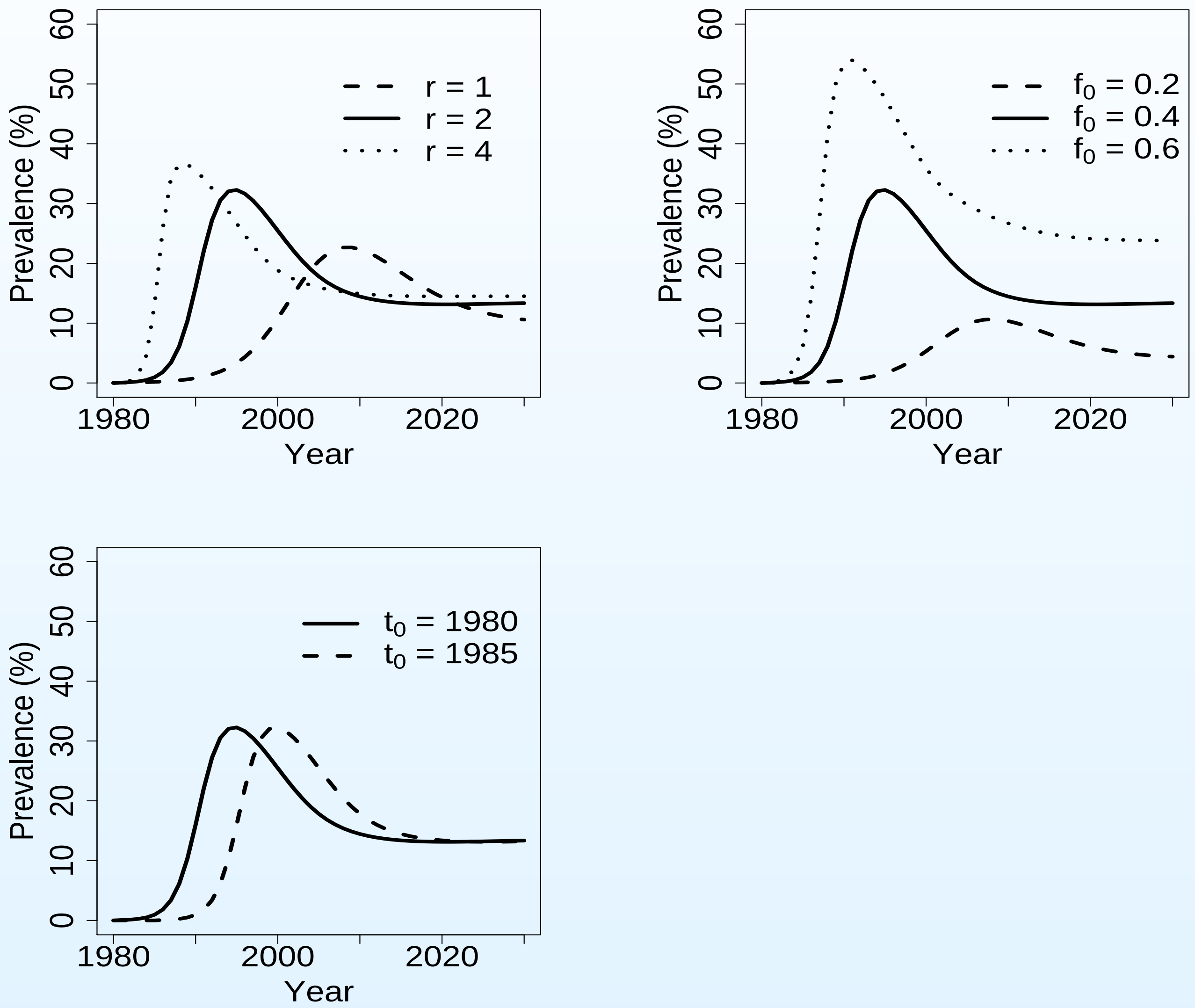


Input parameters (Solid curve: $r = 2, f_0 = 0.4, t_0 = 1980, \phi = 0$)

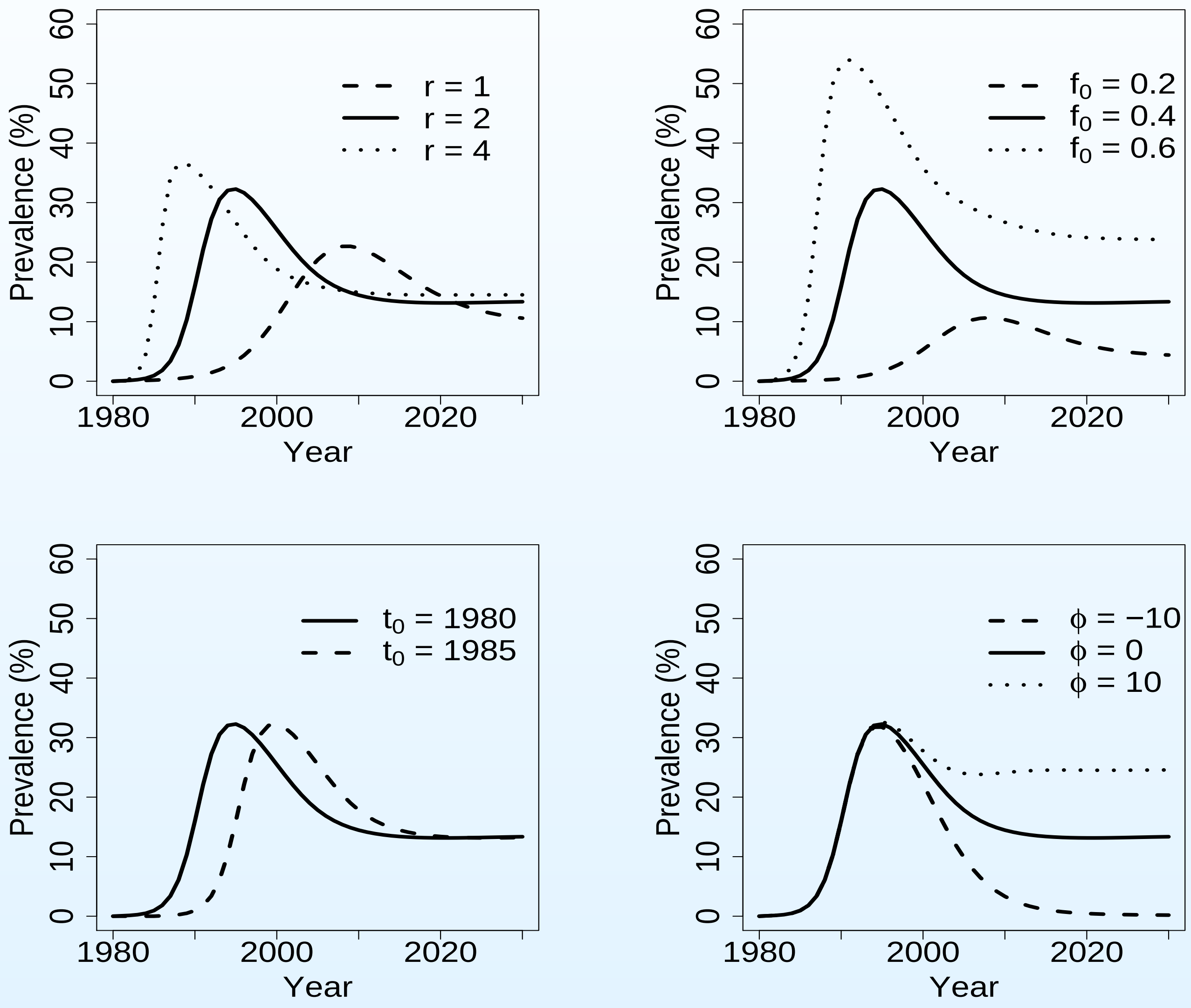




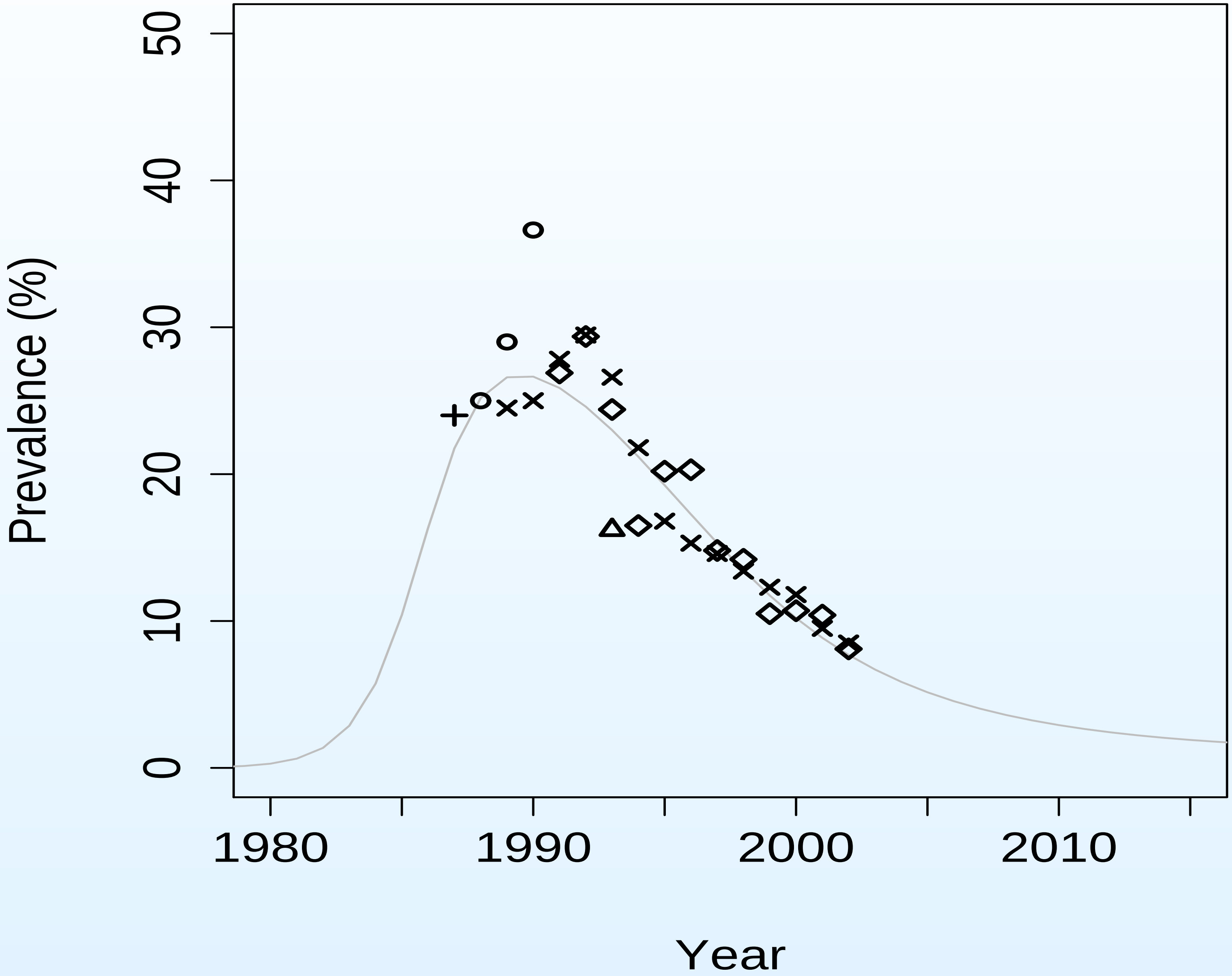
Input parameters (Solid curve: $r = 2, f_0 = 0.4, t_0 = 1980, \phi = 0$)



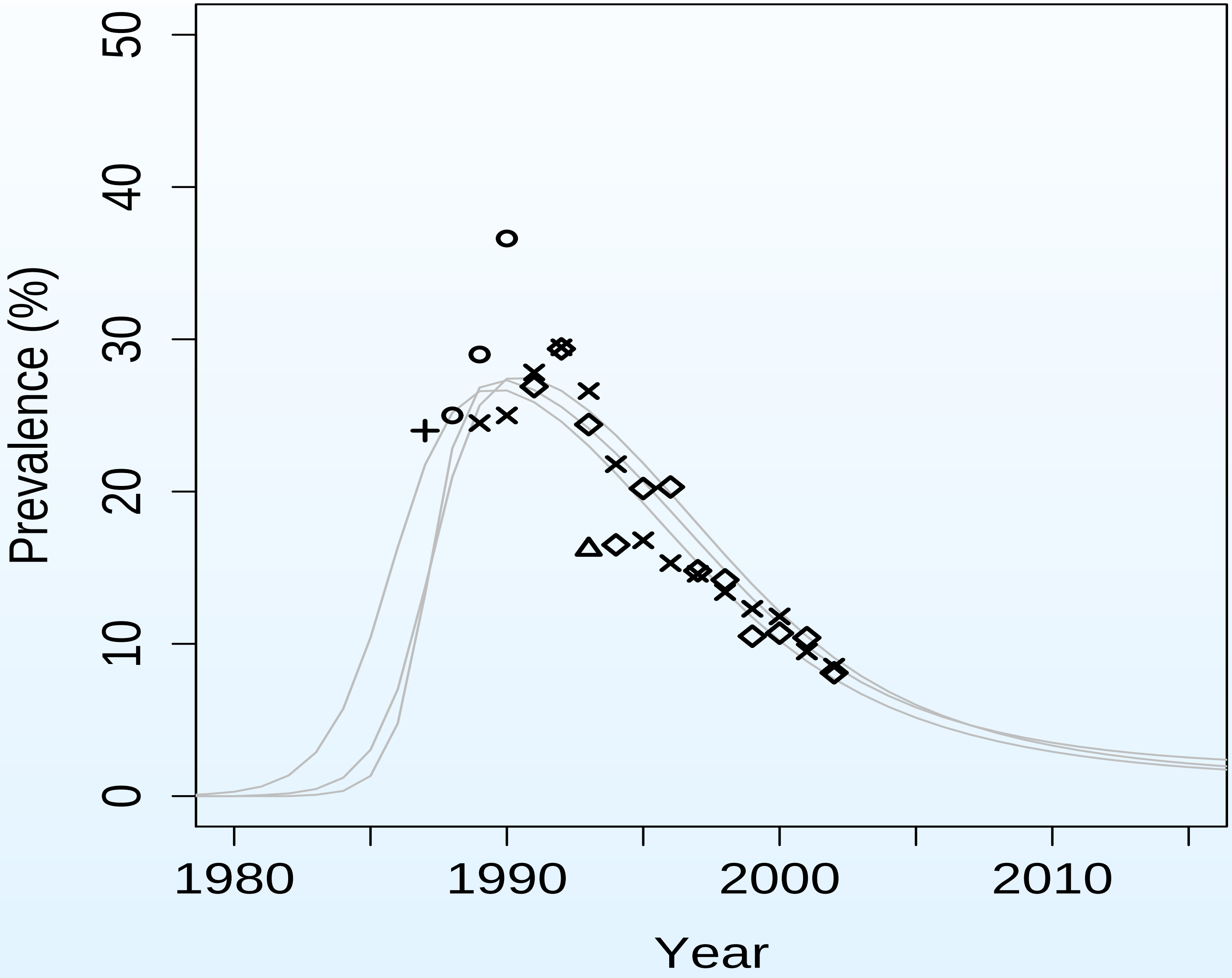
Input parameters (Solid curve: $r = 2, f_0 = 0.4, t_0 = 1980, \phi = 0$)



HIV prevalence prediction for Uganda urban



HIV prevalence prediction for Uganda urban



Outline

- HIV projections
- Bayesian melding
- Results for Uganda
- Summary

Bayesian melding

- Model M with inputs θ and output ρ; $M(\theta) = \rho$
  - Input $\theta = (r, f_0, t_0, \phi)$
  - Output ρ = Population prevalence

Bayesian melding

- Model M with inputs θ and output ρ; $M(\theta) = \rho$
  - Input $\theta = (r, f_0, t_0, \phi)$
  - Output ρ = Population prevalence
- Combine information on θ and ρ
  - Expert knowledge gives prior $q(\theta)$ for inputs
  - Observed prevalence gives likelihood $L(\rho)$ for output

Bayesian melding

- Model M with inputs θ and output ρ; $M(\theta) = \rho$
  - Input $\theta = (r, f_0, t_0, \phi)$
  - Output ρ = Population prevalence
- Combine information on θ and ρ
  - Expert knowledge gives prior $q(\theta)$ for inputs
  - Observed prevalence gives likelihood $L(\rho)$ for output
- Input posterior $\pi(\theta)$:

$$\pi(\theta) \;\; \propto \;\; q(\theta)L(\theta) = q(\theta)L(M(\theta))$$

Bayesian melding

- Model M with inputs θ and output ρ; $M(\theta) = \rho$
  - Input $\theta = (r, f_0, t_0, \phi)$
  - Output ρ = Population prevalence
- Combine information on θ and ρ
  - Expert knowledge gives prior $q(\theta)$ for inputs
  - Observed prevalence gives likelihood $L(\rho)$ for output
- Input posterior $\pi(\theta)$:

$$\pi(\theta) \;\propto\; q(\theta)L(\theta) = q(\theta)L(M(\theta))$$

- Output posterior $\pi(\rho)$:

$$\pi(\rho) \;\propto\; q^*(\rho)L(\rho)$$

with $q^*(\rho)$ the induced prior on the outputs

Sampling importance resampling algorithm

- Sample $\{\theta_1, ..., \theta_n\}$ from the input prior $q(\theta)$ on (r, f_0, t_0, ϕ)

Sampling importance resampling algorithm

- Sample $\{\theta_1, ..., \theta_n\}$ from the input prior $q(\theta)$ on (r, f_0, t_0, ϕ)
- For each θ_i, determine the prevalence rates $\rho_i = M(\theta_i)$, by running the model. This gives a sample of the induced prior $q^*(\rho)$ on the outputs

Sampling importance resampling algorithm

- Sample $\{\theta_1, ..., \theta_n\}$ from the input prior $q(\theta)$ on (r, f_0, t_0, ϕ)
- For each θ_i, determine the prevalence rates $\rho_i = M(\theta_i)$, by running the model. This gives a sample of the induced prior $q^*(\rho)$ on the outputs
- Form the sampling importance weights for each ρ_i, and thus for each θ_i:

$$w_i = \frac{L(\rho_i)}{\sum_{i=1}^{n} L(\rho_i)}$$

Sampling importance resampling algorithm

- Sample $\{\theta_1, ..., \theta_n\}$ from the input prior $q(\theta)$ on (r, f_0, t_0, ϕ)
- For each θ_i, determine the prevalence rates $\rho_i = M(\theta_i)$, by running the model. This gives a sample of the induced prior $q^*(\rho)$ on the outputs
- Form the sampling importance weights for each ρ_i, and thus for each θ_i:

$$w_i = \frac{L(\rho_i)}{\sum_{i=1}^{n} L(\rho_i)}$$

- Sample from the discrete distribution of $\{\theta_1, ..., \theta_n\}$ with probabilities w_i to get the posterior distribution for the inputs, same for the outputs

Outline

- HIV projections
- Bayesian melding
- Results for Uganda
- Summary

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0, 15] \\ f_0 & \sim U[0, 1] \\ t_0 & \sim \text{DiscreteU}(1970, 1990) \\ \phi & \sim \text{Logistic}(0, 10) \end{cases}$$

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0,15] \\ f_0 & \sim U[0,1] \\ t_0 & \sim \text{DiscreteU}(1970,1990) \\ \phi & \sim \text{Logistic}(0,10) \end{cases}$$

- Predict prevalence with EPP model

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0,15] \\ f_0 & \sim U[0,1] \\ t_0 & \sim \text{DiscreteU}(1970,1990) \\ \phi & \sim \text{Logistic}(0,10) \end{cases}$$

- Predict prevalence with EPP model
- Calculate the likelihood for each prevalence curve based on antenatal clinic data

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0, 15] \\ f_0 & \sim U[0, 1] \\ t_0 & \sim \text{DiscreteU}(1970, 1990) \\ \phi & \sim \text{Logistic}(0, 10) \end{cases}$$

- Predict prevalence with EPP model
- Calculate the likelihood for each prevalence curve based on antenatal clinic data
- Resample inputs and curves by likelihood weights

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0,15] \\ f_0 & \sim U[0,1] \\ t_0 & \sim \text{DiscreteU}(1970,1990) \\ \phi & \sim \text{Logistic}(0,10) \end{cases}$$

- Predict prevalence with EPP model
- Calculate the likelihood for each prevalence curve based on antenatal clinic data
- Resample inputs and curves by likelihood weights
- Sample sizes:

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0,15] \\ f_0 & \sim U[0,1] \\ t_0 & \sim \text{DiscreteU}(1970,1990) \\ \phi & \sim \text{Logistic}(0,10) \end{cases}$$

- Predict prevalence with EPP model
- Calculate the likelihood for each prevalence curve based on antenatal clinic data
- Resample inputs and curves by likelihood weights
- Sample sizes:
  - 200,000 initial sets of inputs

Bayesian melding results for Uganda

- Sample inputs from input prior

$$\begin{cases} r & \sim U[0,15] \\ f_0 & \sim U[0,1] \\ t_0 & \sim \text{DiscreteU}(1970,1990) \\ \phi & \sim \text{Logistic}(0,10) \end{cases}$$

- Predict prevalence with EPP model
- Calculate the likelihood for each prevalence curve based on antenatal clinic data
- Resample inputs and curves by likelihood weights
- Sample sizes:
  - 200,000 initial sets of inputs
  - resampling 3,000 sets

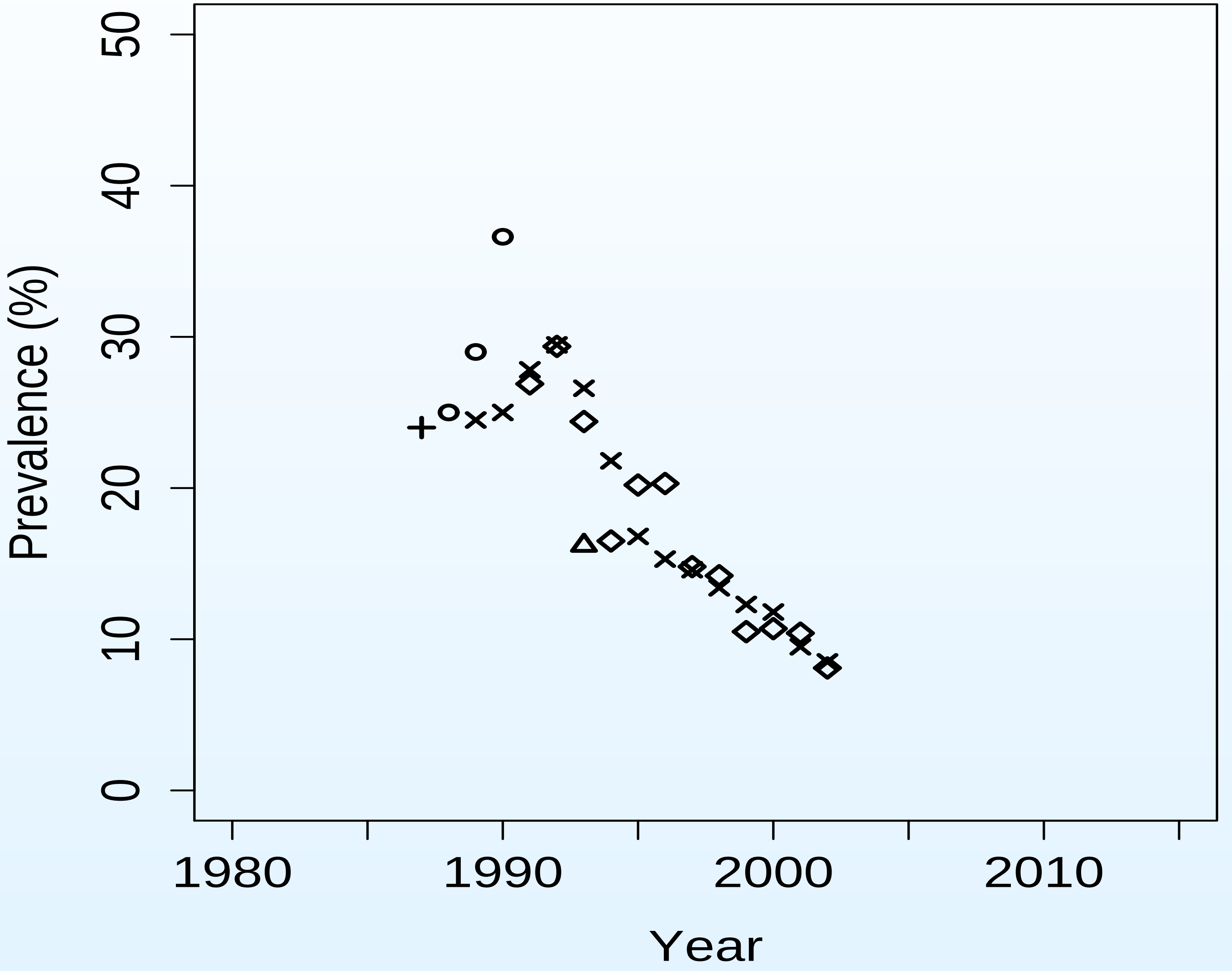
Bayesian melding results for Uganda

- Sample inputs from input prior

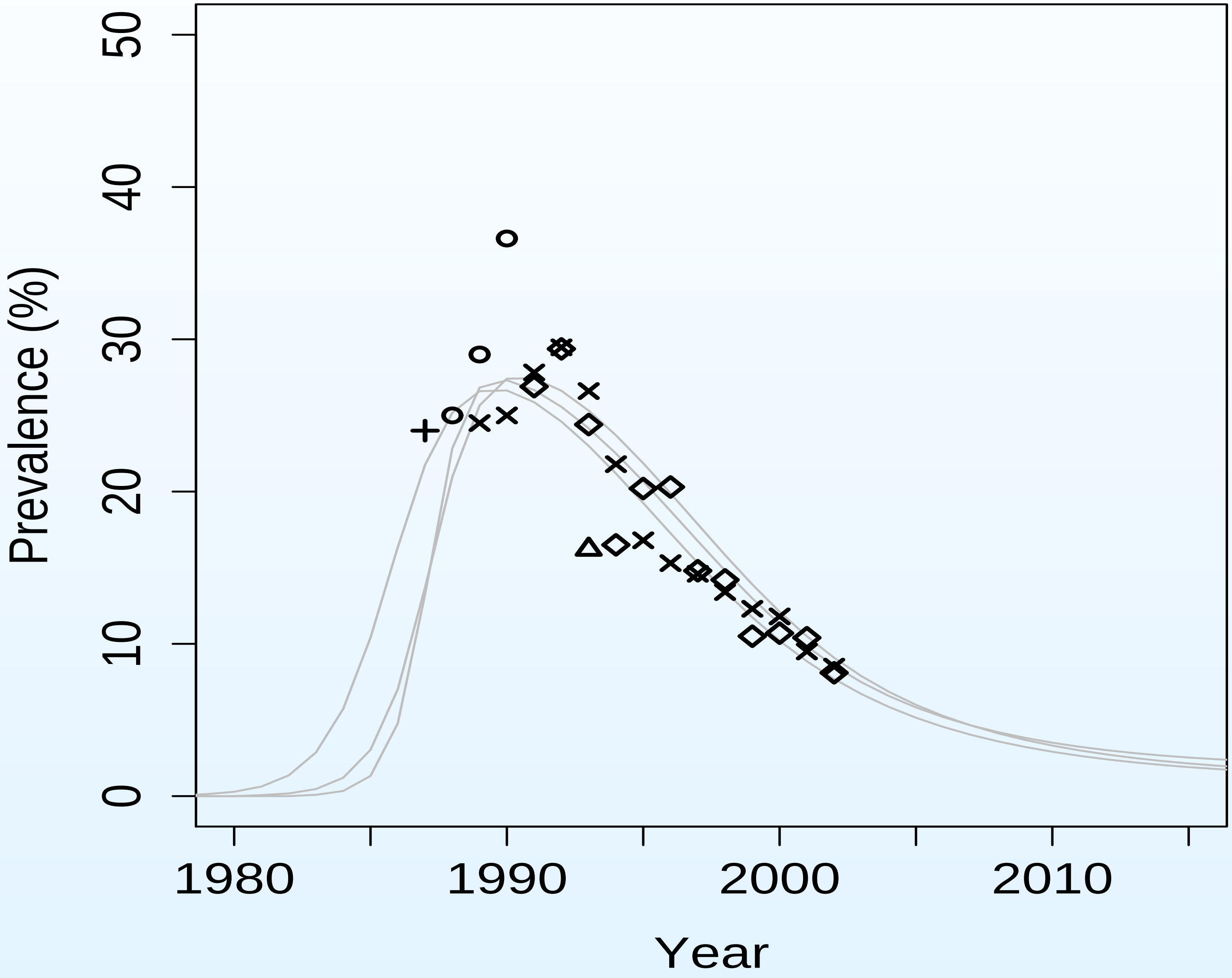
$$\begin{cases} r & \sim U[0,15] \\ f_0 & \sim U[0,1] \\ t_0 & \sim \text{DiscreteU}(1970,1990) \\ \phi & \sim \text{Logistic}(0,10) \end{cases}$$

- Predict prevalence with EPP model
- Calculate the likelihood for each prevalence curve based on antenatal clinic data
- Resample inputs and curves by likelihood weights
- Sample sizes:
  - 200,000 initial sets of inputs
  - resampling 3,000 sets
  - 508 unique combinations

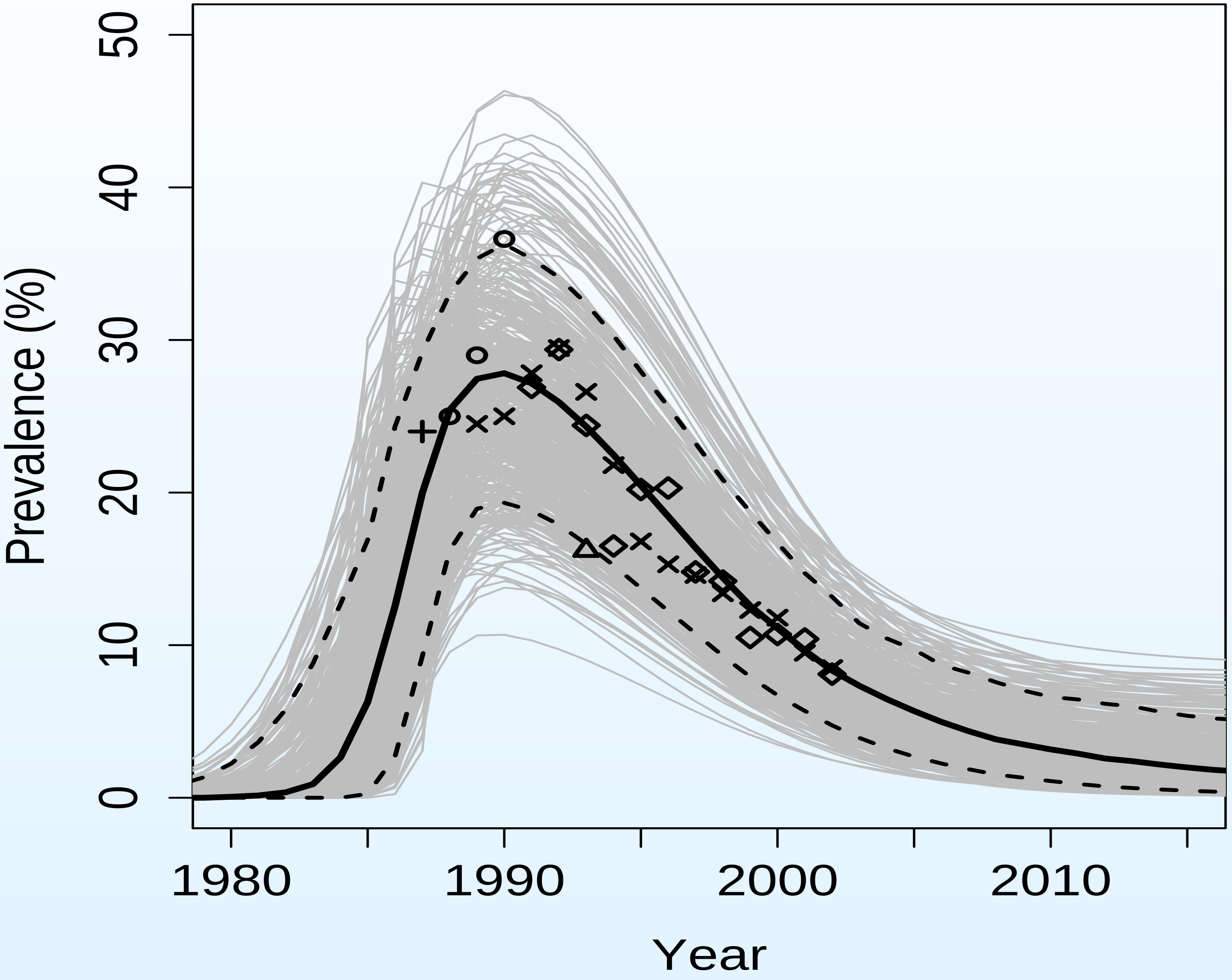
HIV prevalence Uganda



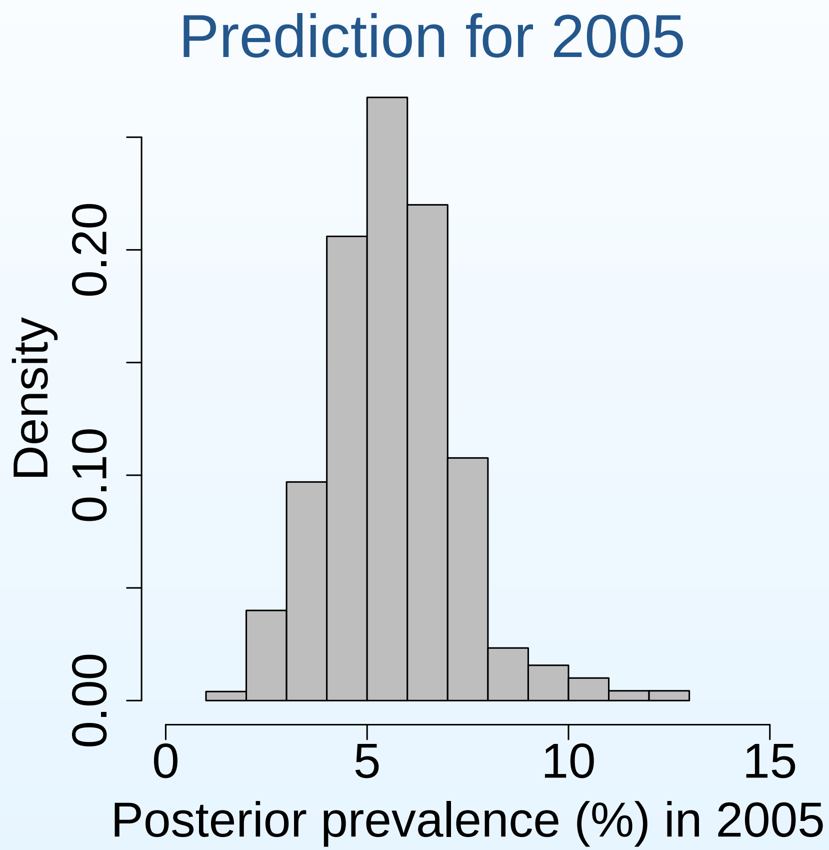
HIV prevalence Uganda



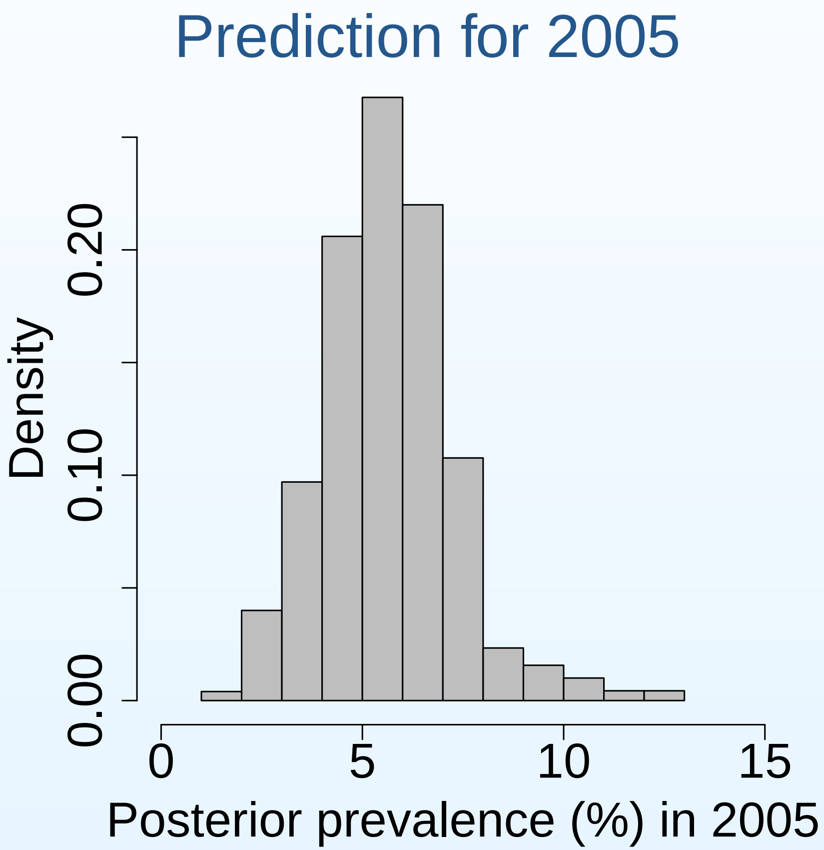
HIV prevalence Uganda

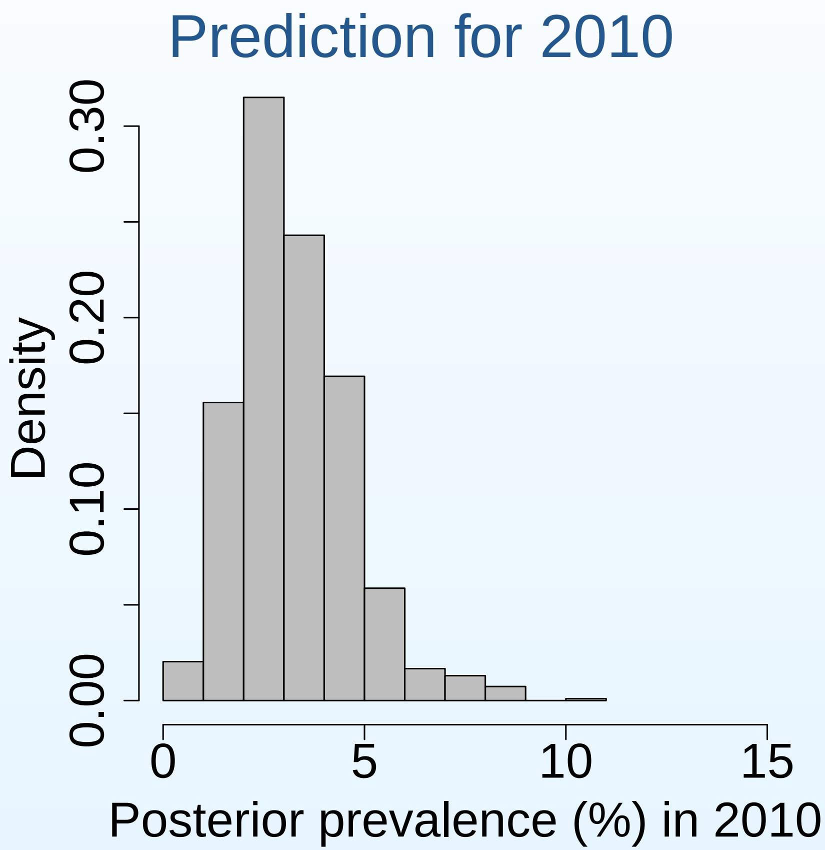


Posterior Predictive Distributions

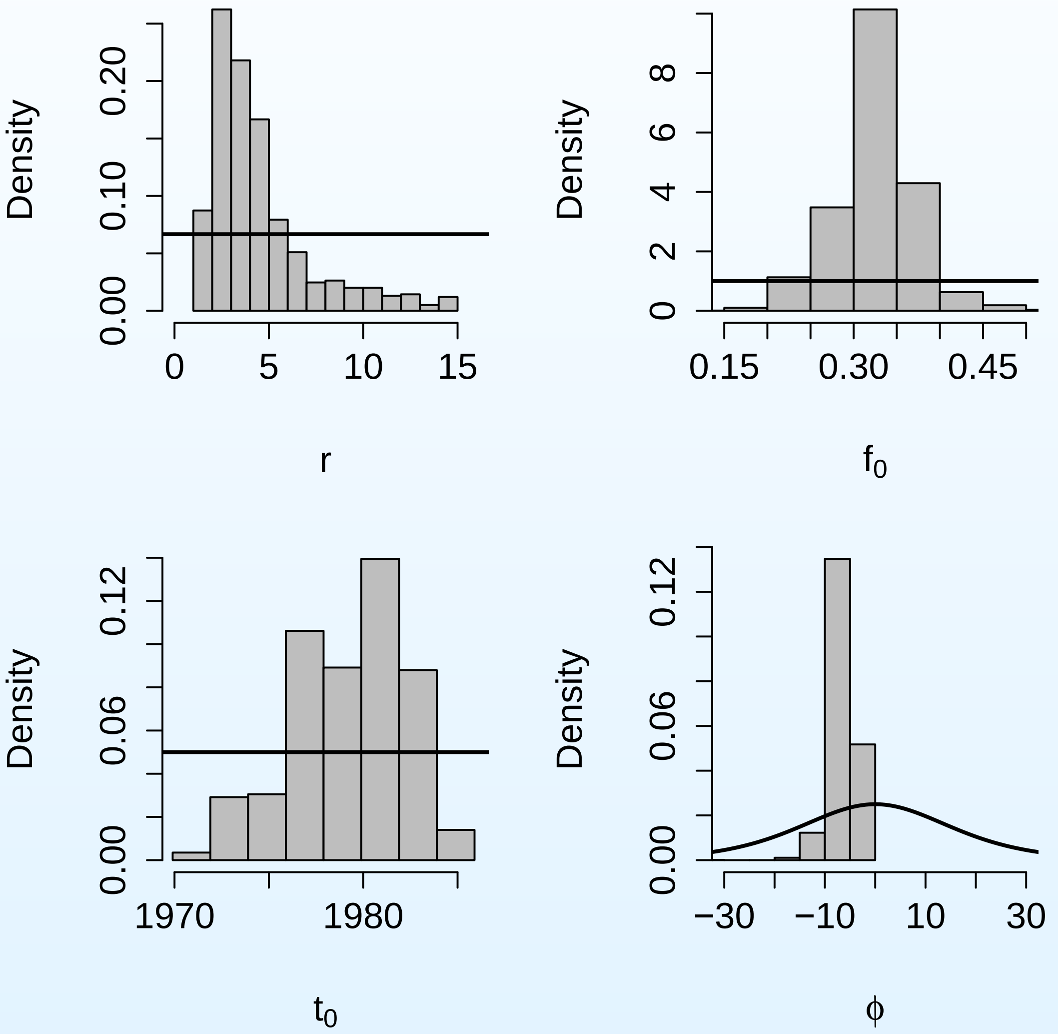


Posterior Predictive Distributions

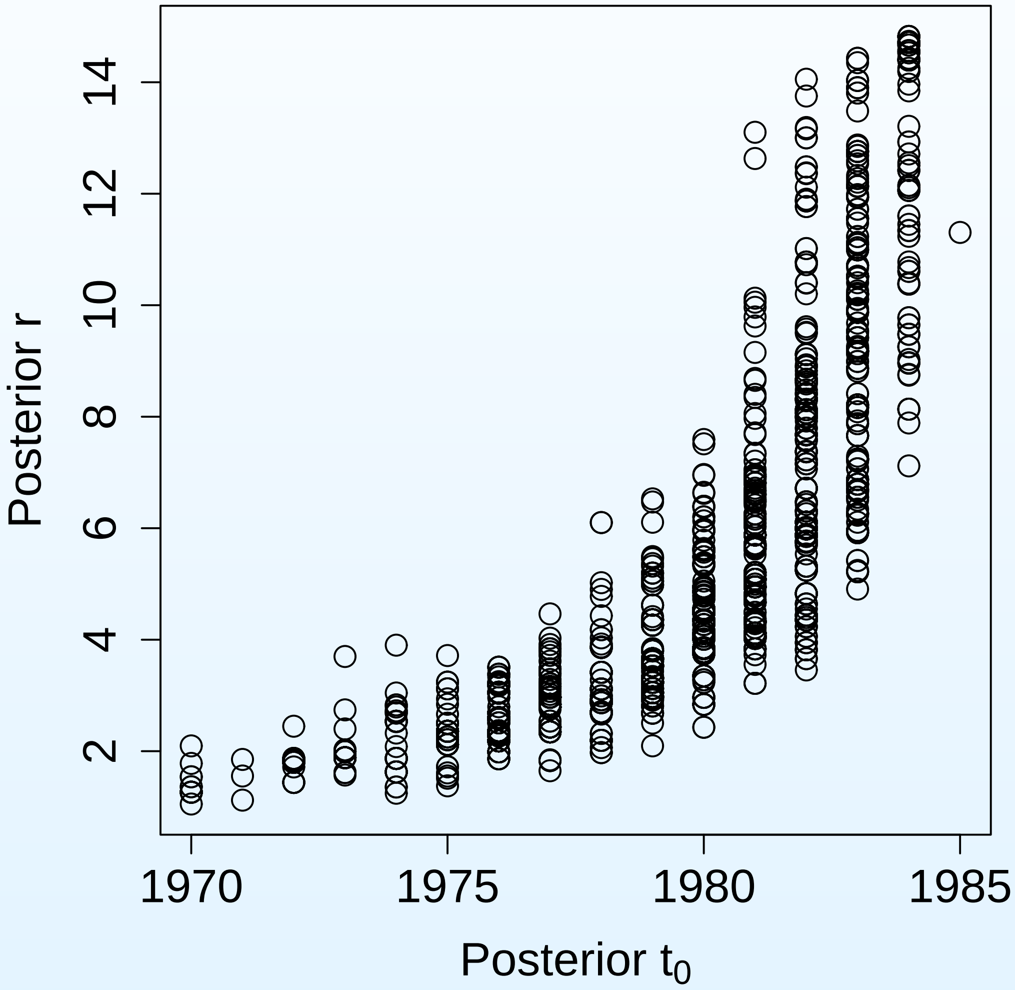




Prior and posterior distribution of inputs

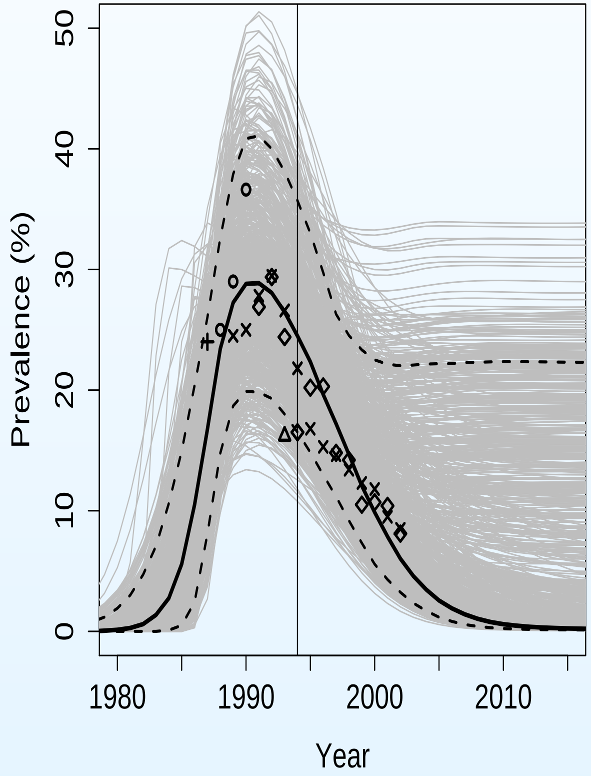


Joint posterior of input parameters r and t_0



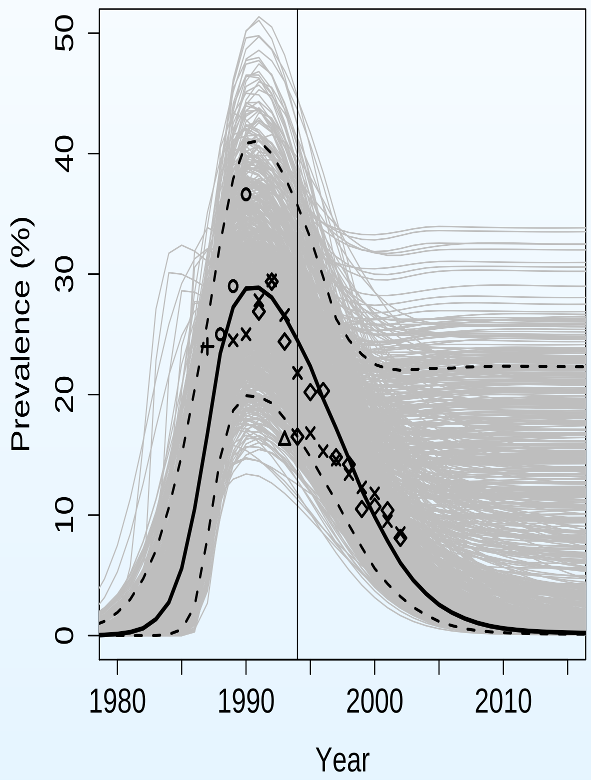
Changes in predictions over time

Data through 1994

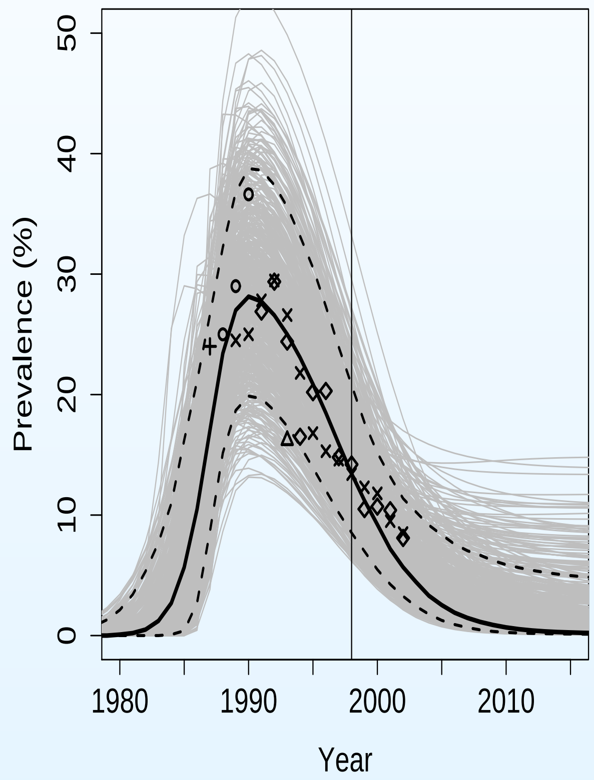


Changes in predictions over time

Data through 1994

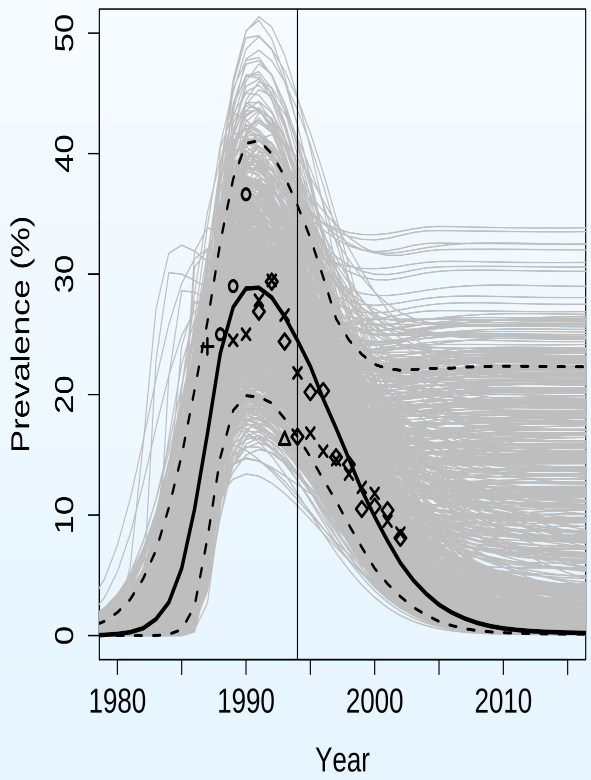


Data through 1998

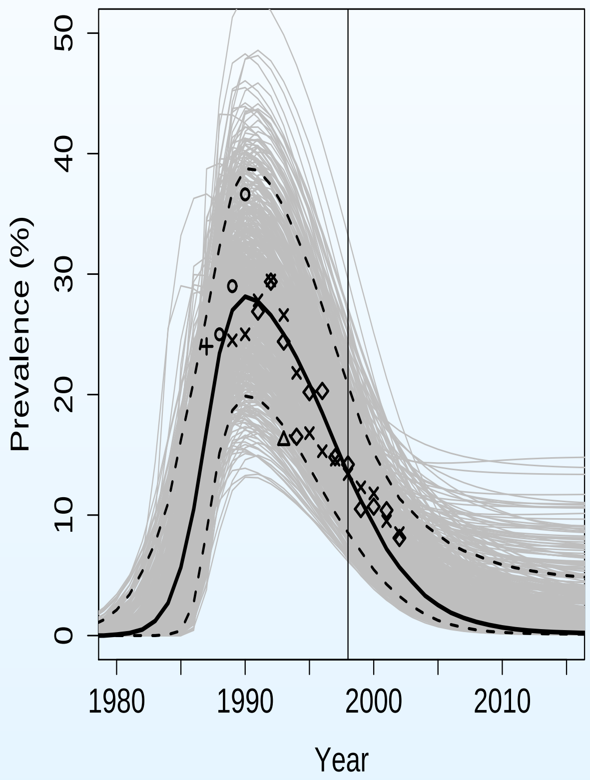


Changes in predictions over time

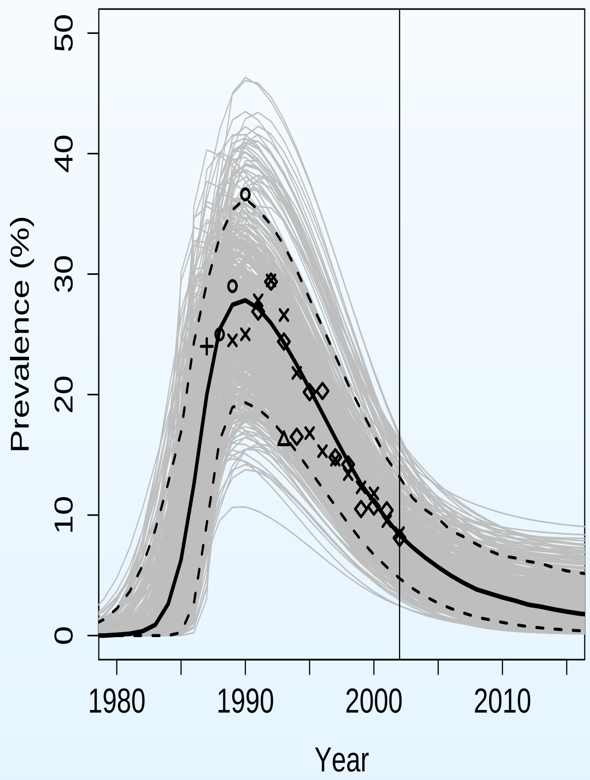
Data through 1994



Data through 1998



Data through 2002



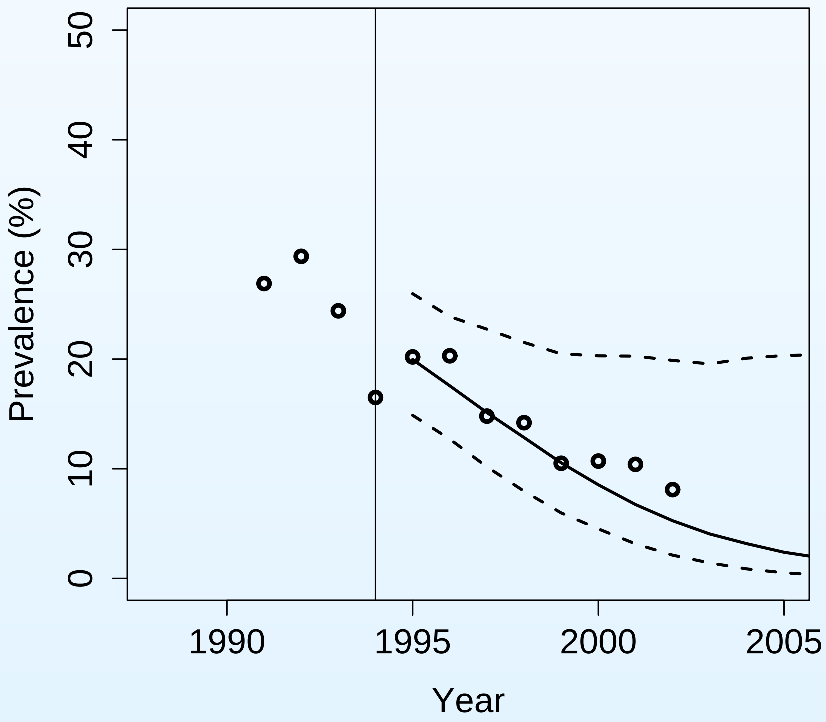
Predictive performance of EPP

- Compare predicted clinic prevalence to observed clinic prev.

Predictive performance of EPP

- Compare predicted clinic prevalence to observed clinic prev.
- Prediction intervals for one clinic in Uganda:

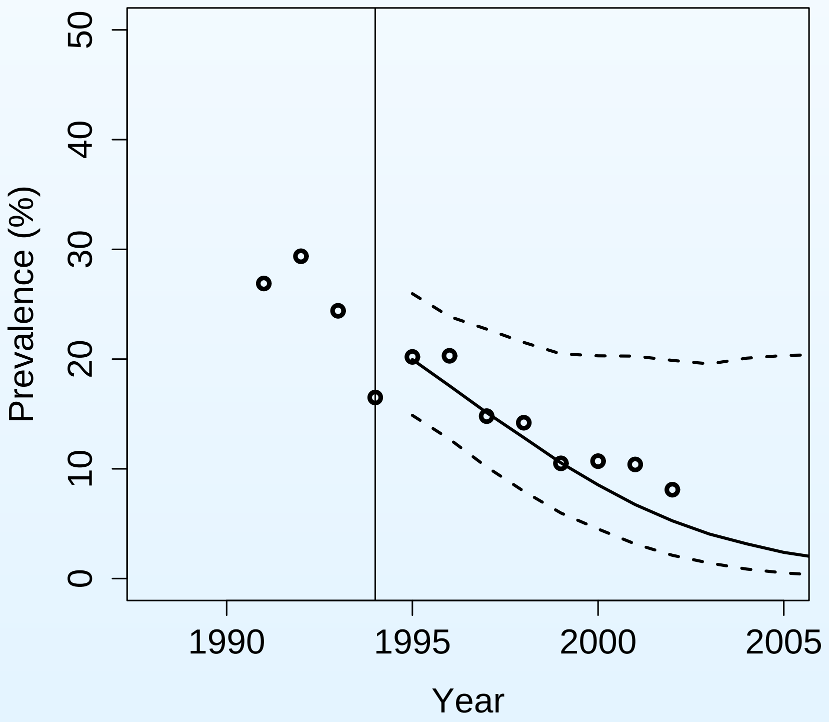
Data through 1994



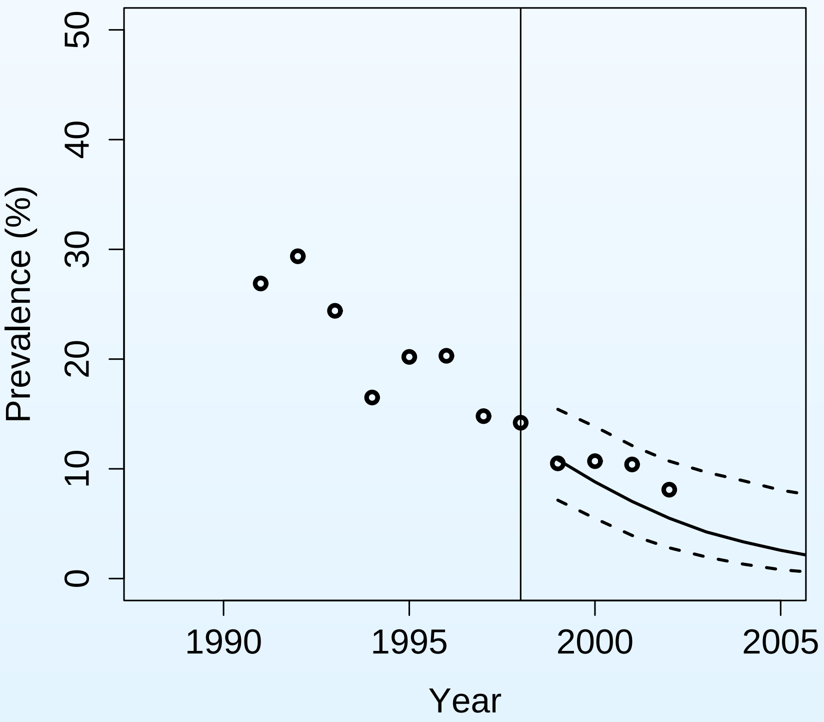
Predictive performance of EPP

- Compare predicted clinic prevalence to observed clinic prev.
- Prediction intervals for one clinic in Uganda:

Data through 1994



Data through 1998



Outline

- HIV projections
- Bayesian melding
- Results for Uganda
- Summary

Summary

- Projections of HIV prevalence

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding
  - Uncertainty assessment
    of inputs and outputs in a deterministic model

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding
  - Uncertainty assessment
    of inputs and outputs in a deterministic model
  - Combines prior information and data

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding
  - Uncertainty assessment
    of inputs and outputs in a deterministic model
  - Combines prior information and data
- Combining EPP and Bayesian melding

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding
  - Uncertainty assessment
    of inputs and outputs in a deterministic model
  - Combines prior information and data
- Combining EPP and Bayesian melding
  - Probabilistic projections of HIV prevalence

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding
  - Uncertainty assessment
    of inputs and outputs in a deterministic model
  - Combines prior information and data
- Combining EPP and Bayesian melding
  - Probabilistic projections of HIV prevalence
  - Predictive distributions for clinic prevalence

Summary

- Projections of HIV prevalence
  - UNAIDS EPP model
  - Simple but flexible model,
    fits various epidemics based on little data
- Bayesian melding
  - Uncertainty assessment
    of inputs and outputs in a deterministic model
  - Combines prior information and data
- Combining EPP and Bayesian melding
  - Probabilistic projections of HIV prevalence
  - Predictive distributions for clinic prevalence
- Bayesian melding will be available in EPP 2007