

State-Space Models

Bayesian State Space Model

$$Y_t = g(X_t | \phi)$$
$$X_t = f(X_{t-1} | \theta)$$

Data Model

Process Model

- Y = observed data
- X = latent time series
- ε = process error
- ω = observation error

Random Walk State Space Model

- What are the conditional distributions?

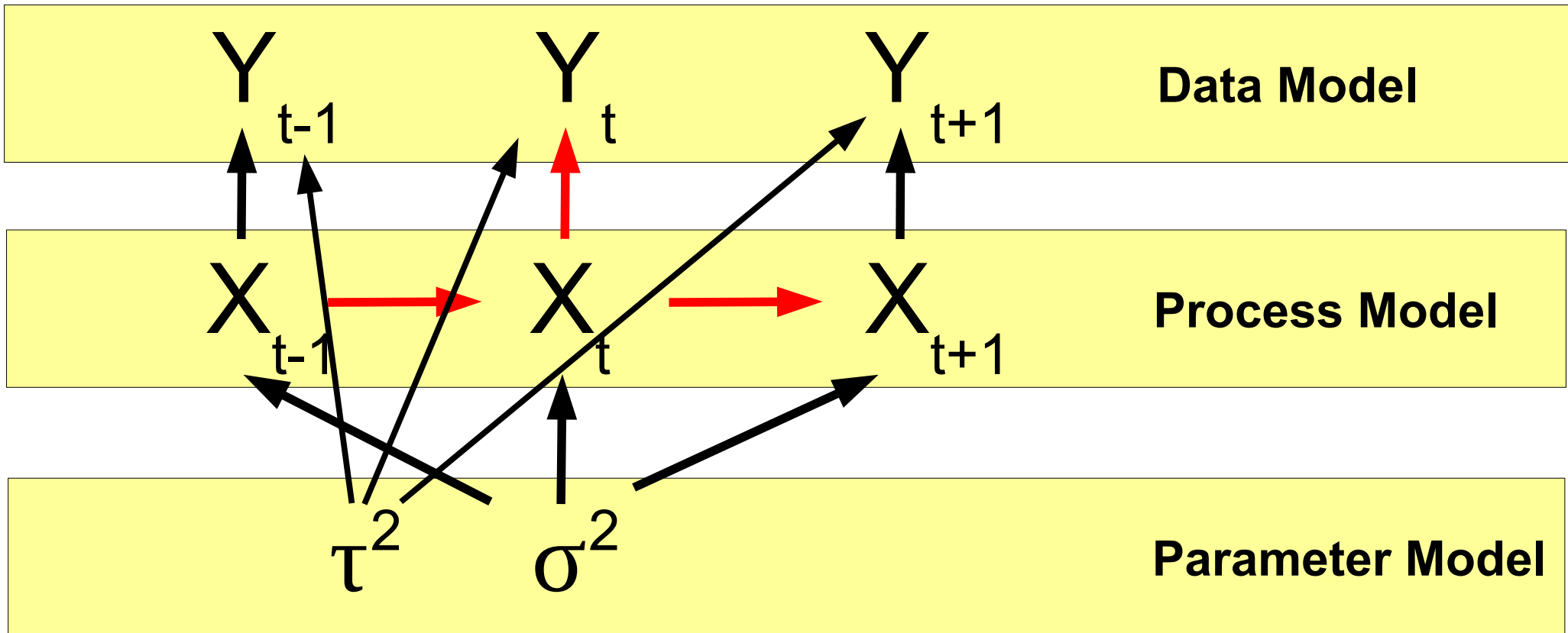
$$X_t \sim N(X_t | X_{t-1}, \tau_{add}^2) \times \\ N(X_{t+1} | X_t, \tau_{add}^2) \times \\ N(Y_t | X_t, \tau_{obs}^2)$$

- Three special cases

- First
- Last
- Missing Y

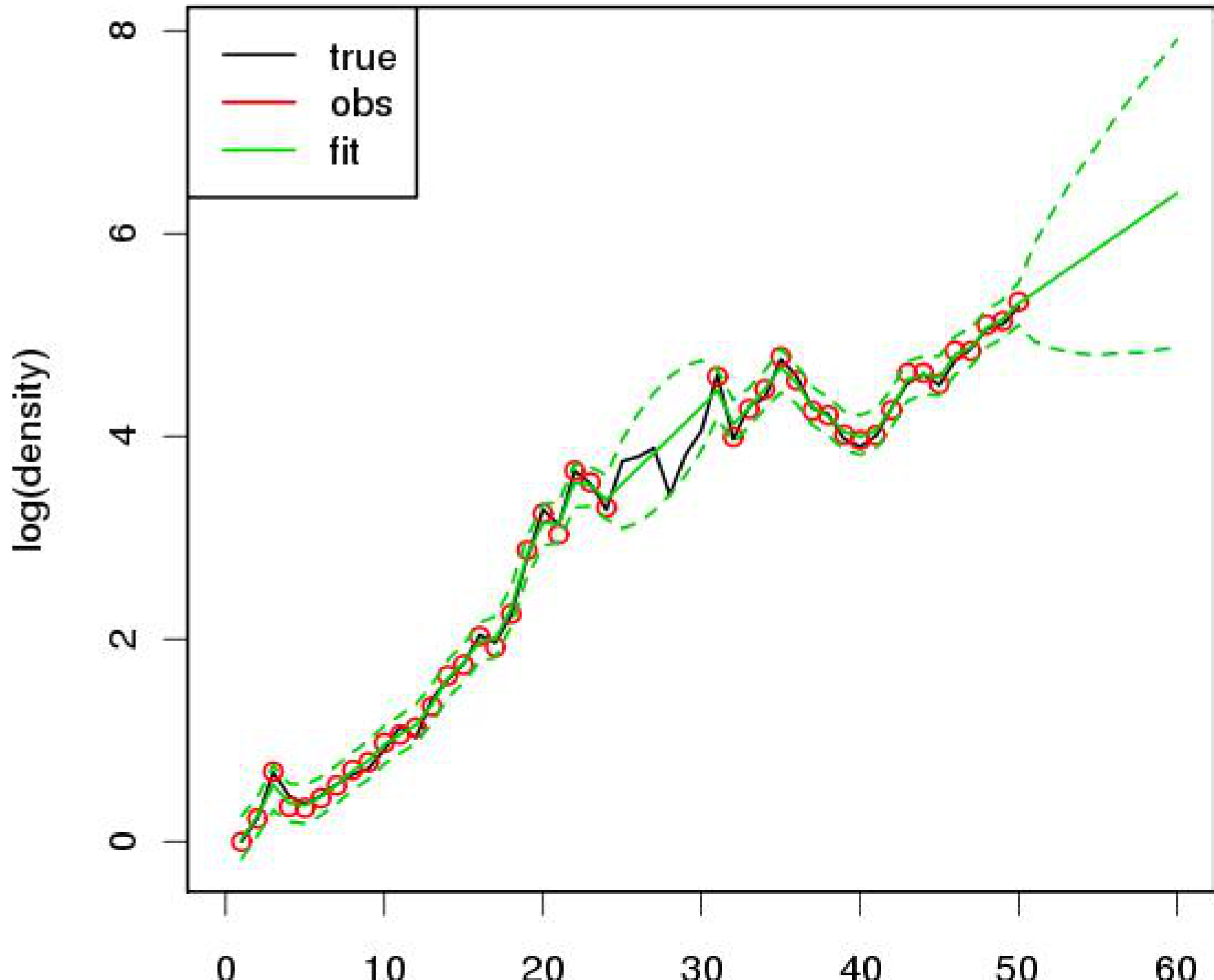
$$X_t \sim N(X_{t-1}, \tau_{add}^2) \\ Y_t \sim N(X_t, \tau_{obs}^2) \\ \tau_{obs}^2 \sim IG(a_{obs}, r_{obs}) \\ \tau_{add}^2 \sim IG(a_{add}, r_{add}) \\ X_0 \sim N(X_{ic}, \tau_{IC})$$

Random Walk State Space Model



Y's are conditionally independent given the X's

Prediction



Generality of the State Space Model

- Neither X nor Y need be Normal
- X and Y don't need to be the same type of data
- X and Y don't need to have the same time scale
- Easily handles missing data (gaps) and irregularly spaced data
- Easily handles multiple data sources (Y 's), which don't need to be the same type or synchronous
- Easily handles time-integrated observations

**Note: “easy” in concept not always equal to short code or fast runtime

Unequal Observation Errors

- Suppose we have an *a priori* reason to believe observation errors were different in different years
 - Different methodology
 - Different sample size
 - QA/QC error estimate

$$\tau_t^2 \sim IG\left(\alpha_t + \frac{1}{2}, \beta_t + \frac{1}{2}(y_t - x_t)^2\right)$$

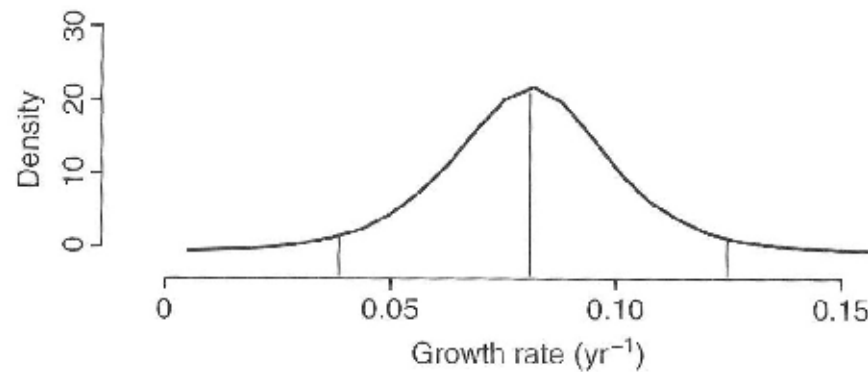
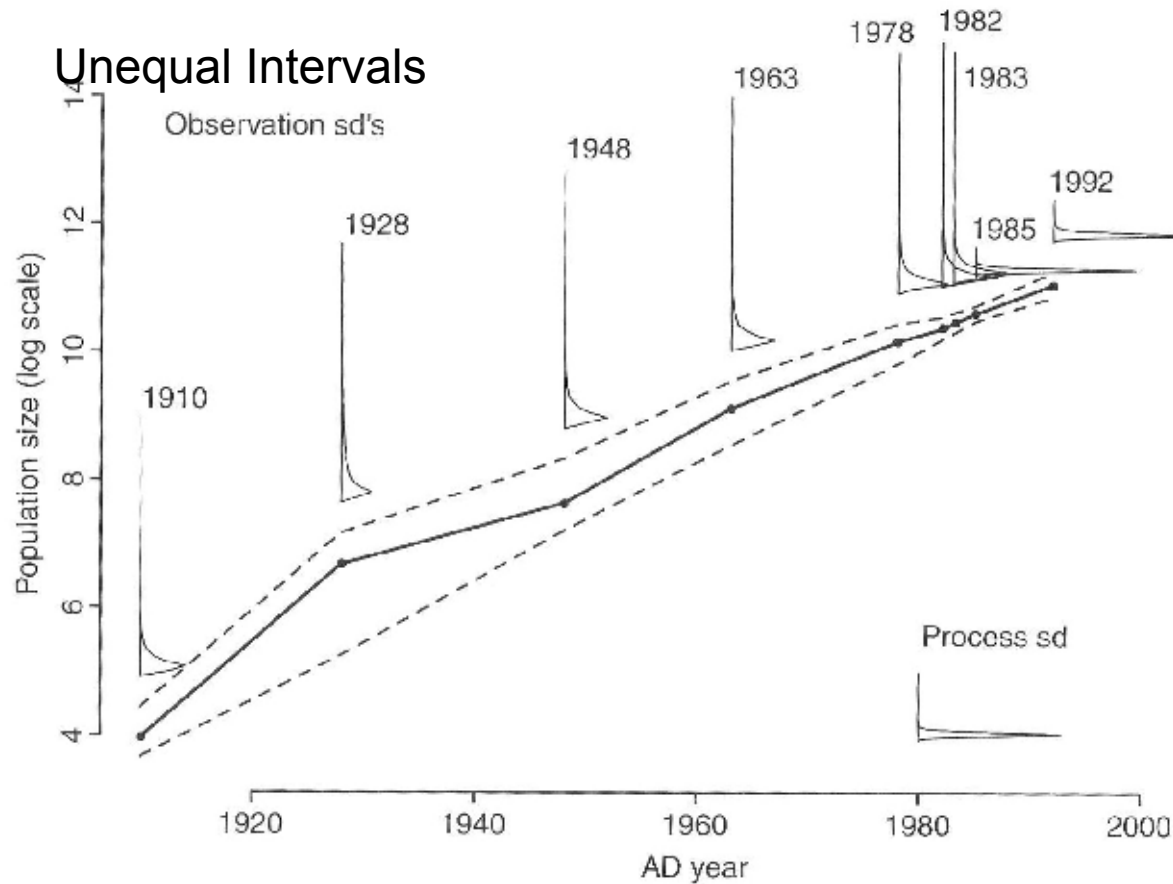
Defining prior differently by year

Unequal Sample Intervals

- Option 1: Treat as missing data
 - Generally applicable
- Option 2: Include time step in process model
 - Problem specific solution

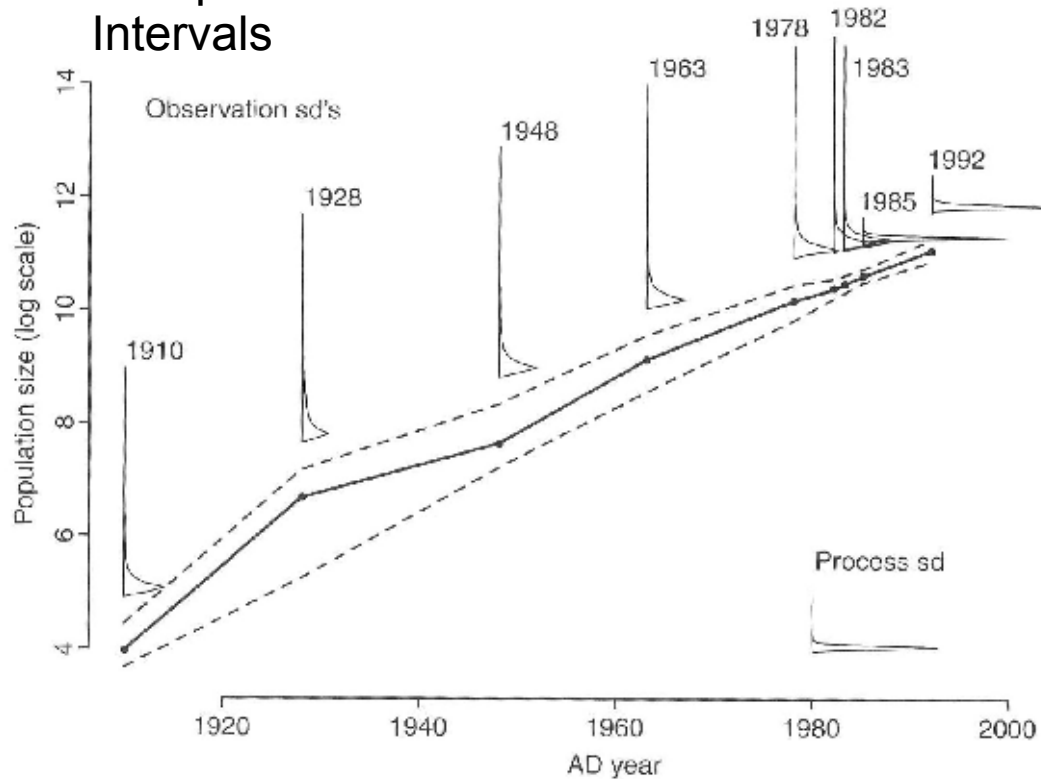
$$X_t = X_{t-\Delta t_i} + (r + \epsilon_t) \cdot \Delta t_i$$

Example: Black Noddy (*Anous minutus*)

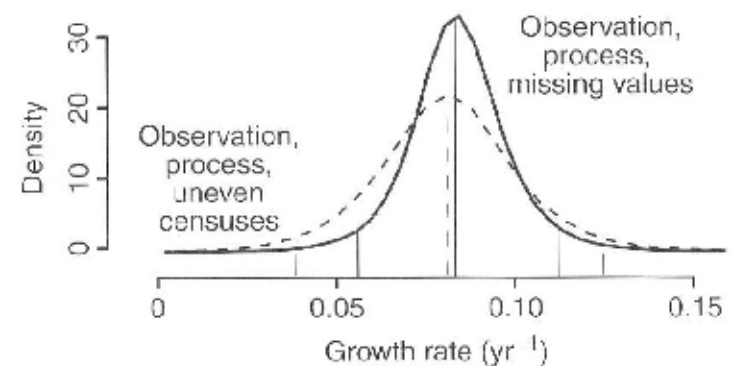
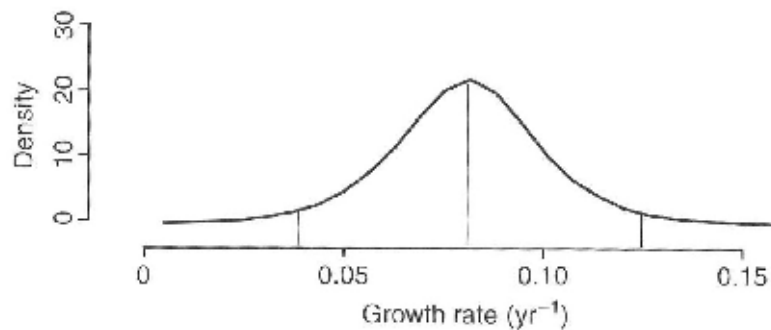
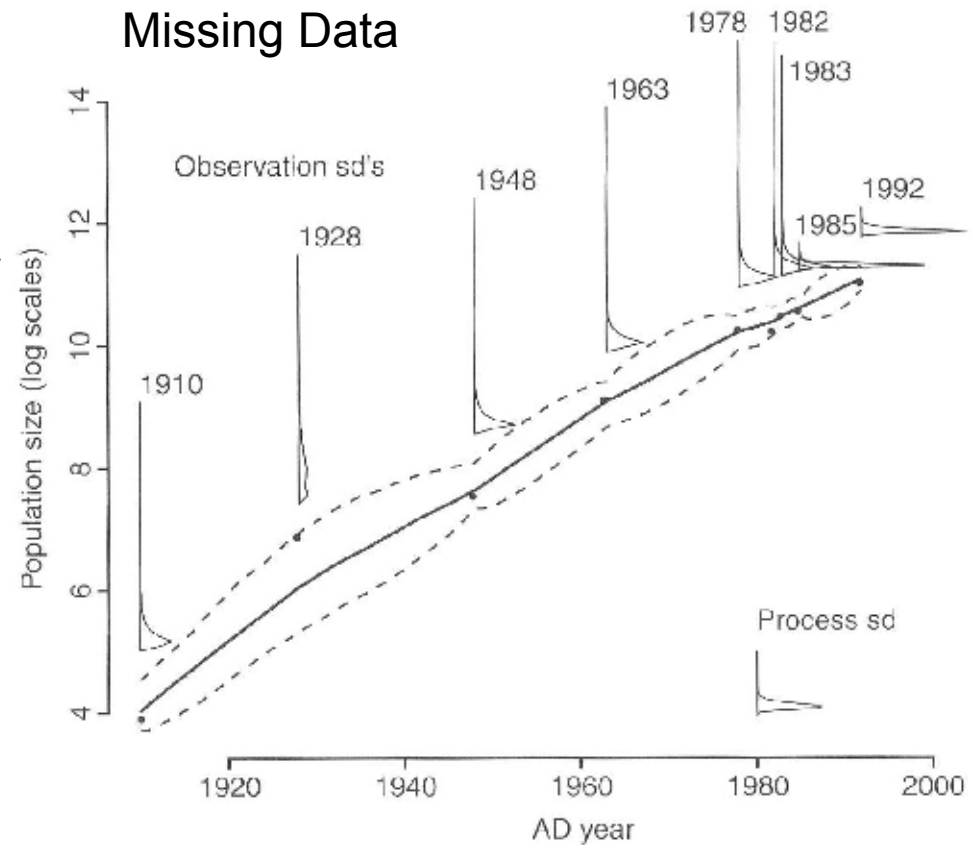


Example: Black Noddy (*Anous minutus*)

Unequal Intervals



Missing Data



Nonlinear State Space: Density Dependence

Exponential

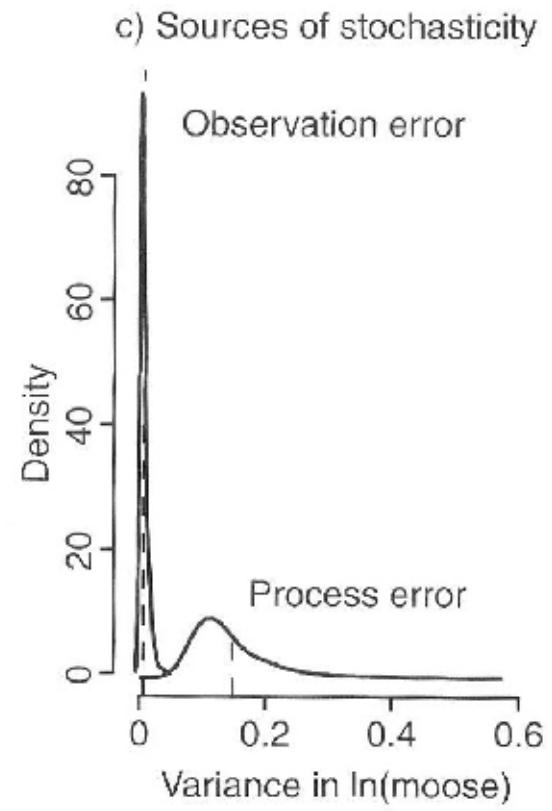
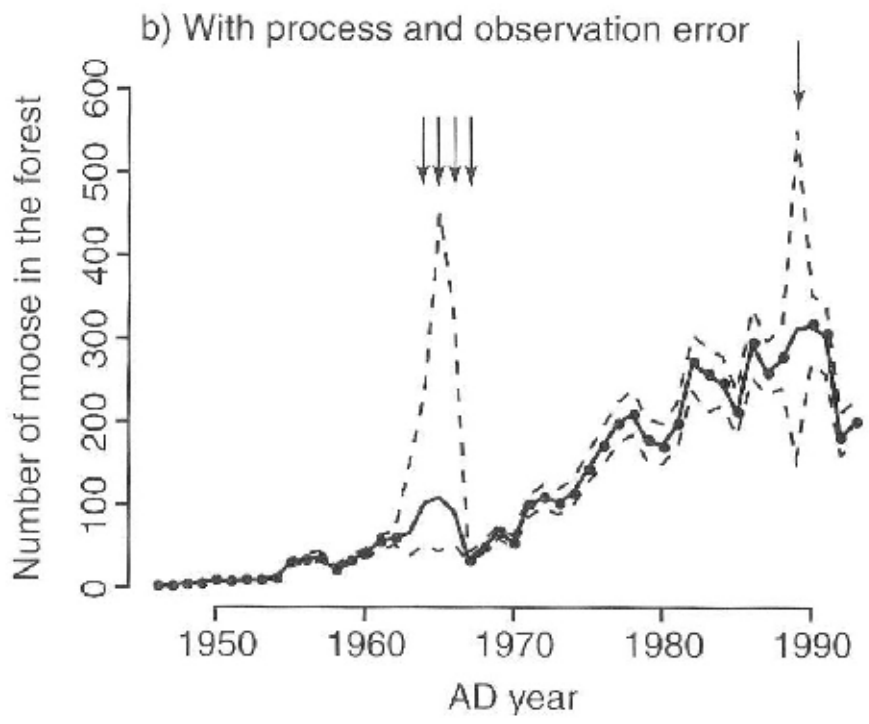
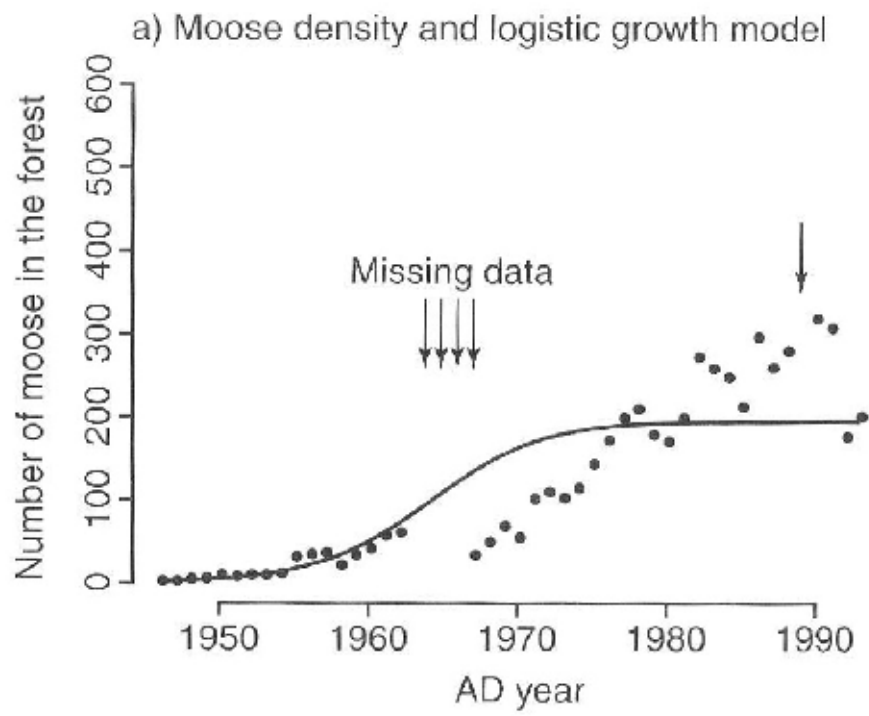
$$N_{t+1} = N_t \cdot e^{r + \epsilon_t} \quad X_{t+1} = X_t + r + \epsilon_t$$

Ricker Discrete Logistic

$$N_{t+1} = N_t \cdot e^{r(1 - N_t/K) + \epsilon_t} \quad X_{t+1} = X_t + r(1 - N_t/K) + \epsilon_t$$

- K needs a positive continuous prior (e.g. lognormal, gamma)
- Metropolis-Hastings sampling for K

Bialowieza Primeval Forest (BPF) Moose



Capture-Recapture

- Individuals captured, marked, and released
- Over repeated censuses will recapture some fraction of the population
- Recapture is random and $<100\%$
- Interested in demography (survival, reproduction, growth) and population size
- Very common with animal data

Missing Data

- Suppose an individual record consists of the following capture data

$$Y_i = [1, 0, 1, 0, 0]$$

- This is compatible with the following survival

$$Z_i = [1, 1, 1, 0, 0]$$

$$Z_i = [1, 1, 1, 1, 0]$$

$$Z_i = [1, 1, 1, 1, 1]$$

- Don't know the exact time of death
- DO know the second census was just a failure to recapture

Missing Data

- Suppose an individual record consists of the following capture data

$$Y_i = [1, 0, 1, 0, 0]$$

- This is compatible with the following survival

$$Z_i = [1, 1, 1, 0, 0]$$

$$Z_i = [1, 1, 1, 1, 0]$$

$$Z_i = [1, 1, 1, 1, 1]$$

Allow estimation of capture probability

- Don't know the exact time of death
- DO know the second census was just a failure to recapture

Basic Mark-Recapture State Space

- Process model

$$P(X_t = 1 | X_{t-1} = 1) = s_t$$

$$P(X_t = 1 | X_{t-1} = 0) = 0$$

$$P(X_t = 0 | X_{t-1} = 1) = 1 - s_t$$

$$P(X_t = 0 | X_{t-1} = 0) = 1$$

**Bernoulli Survival
Probability**

- Observation model

$$P(Y_t = 1 | X_t = 1) = p_t$$

$$P(Y_t = 1 | X_t = 0) = 0$$

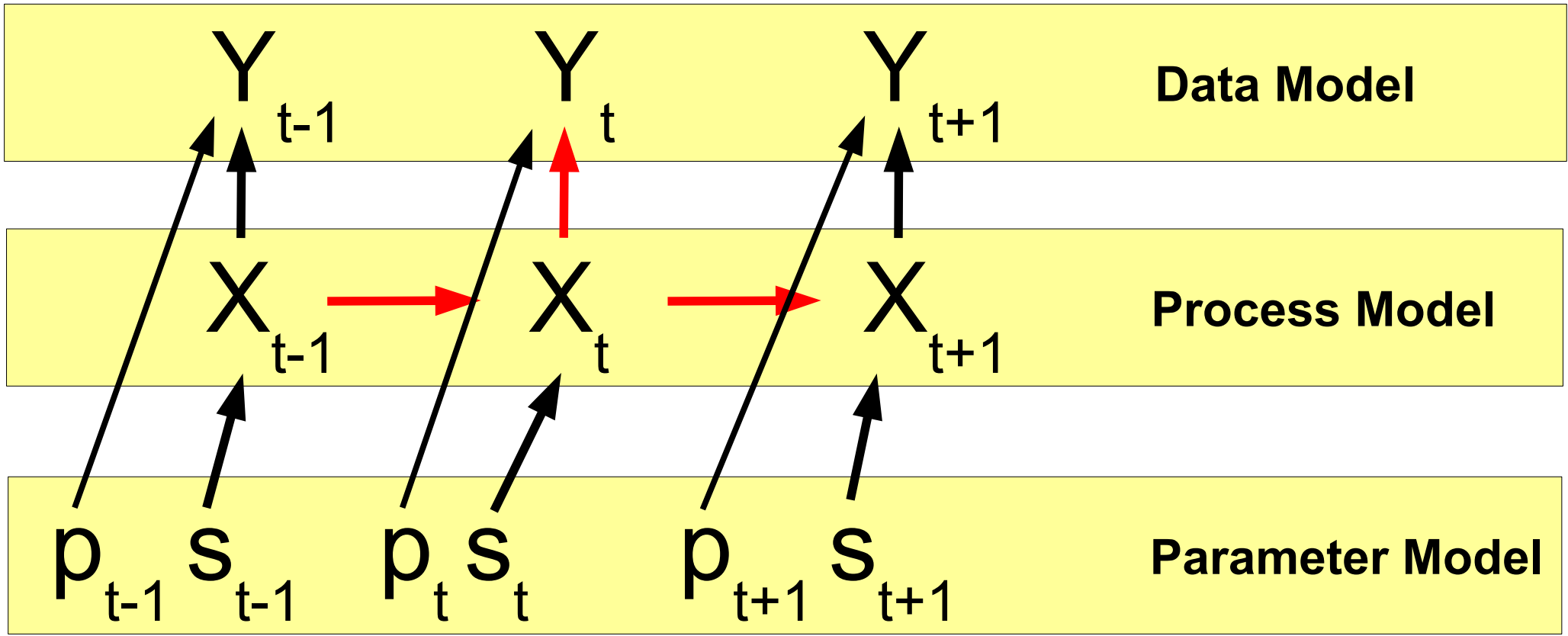
$$P(Y_t = 0 | X_t = 0) = 1$$

$$P(Y_t = 0 | X_t = 1) = 1 - p_t$$

**Bernoulli Detection
Probability**

- Priors on p and s (e.g. Beta)

Mark Recapture State Space



Sampling

- As with previous State-Space, update state variables sequentially based on previous X , next X , and current Y .
 - Don't need to update values if state is known
 - State is binomial $[0,1]$
- Survival and capture probabilities are Beta-Binomial (Gibbs)
 - Don't double count dead (only die once)

Extensions

- Current model assumes p and s vary with time
- Could assume a common, time-invariant p and/or s
- Could assume a hierarchical p and/or s
- Could make either a function of covariates (e.g. GLMM)

$$\textit{logit}(s) = Z\beta + \alpha$$

Note: no ε

Covariates

Fixed effects

Random effects

Mark Recapture State Space

