Availability Bias on Line Transect Surveys

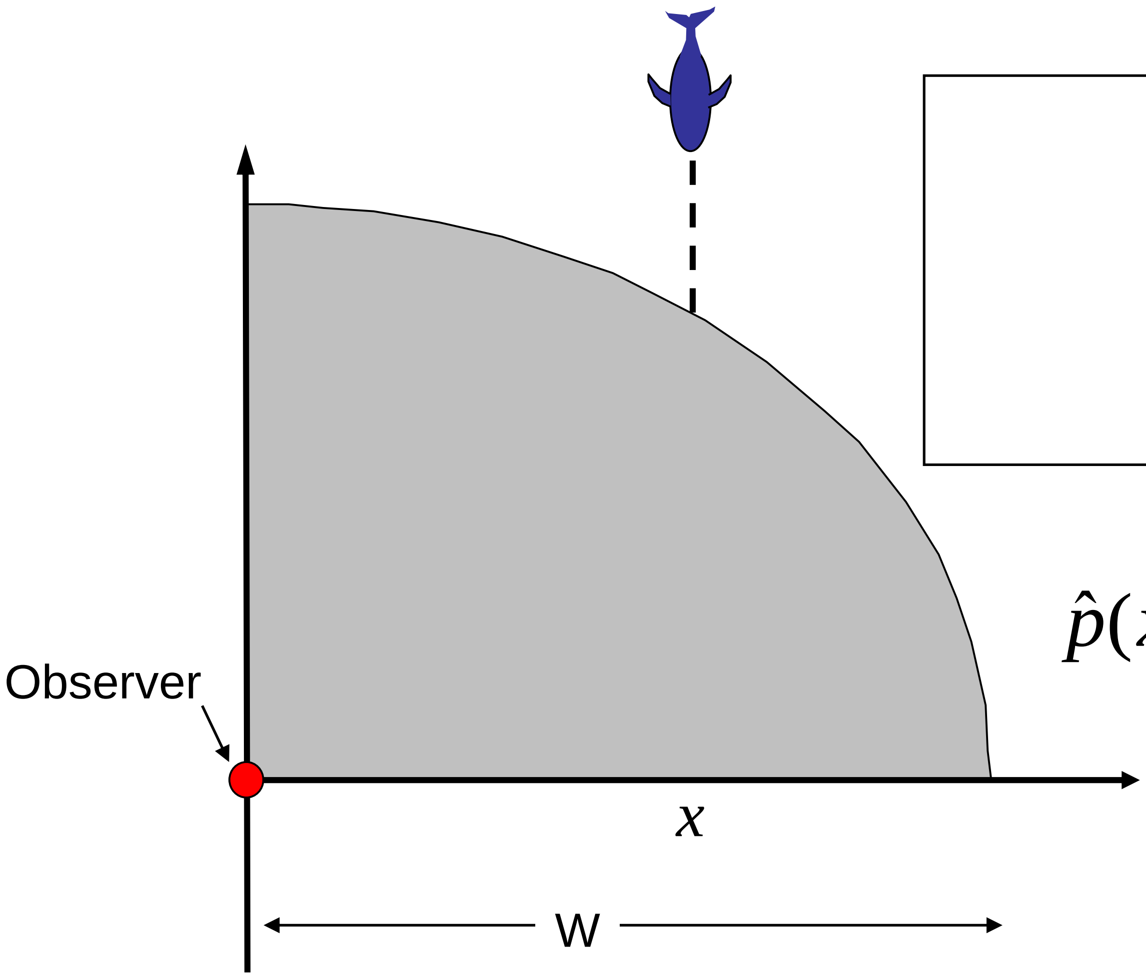
Availability is a kind of heterogeneity

- Animals that become available more often are more likely to be detected.

- Animals that become available closer to observers are more likely to be detected.

Simple Correction Factor

$$Prob(\text{animal is available}) = a$$



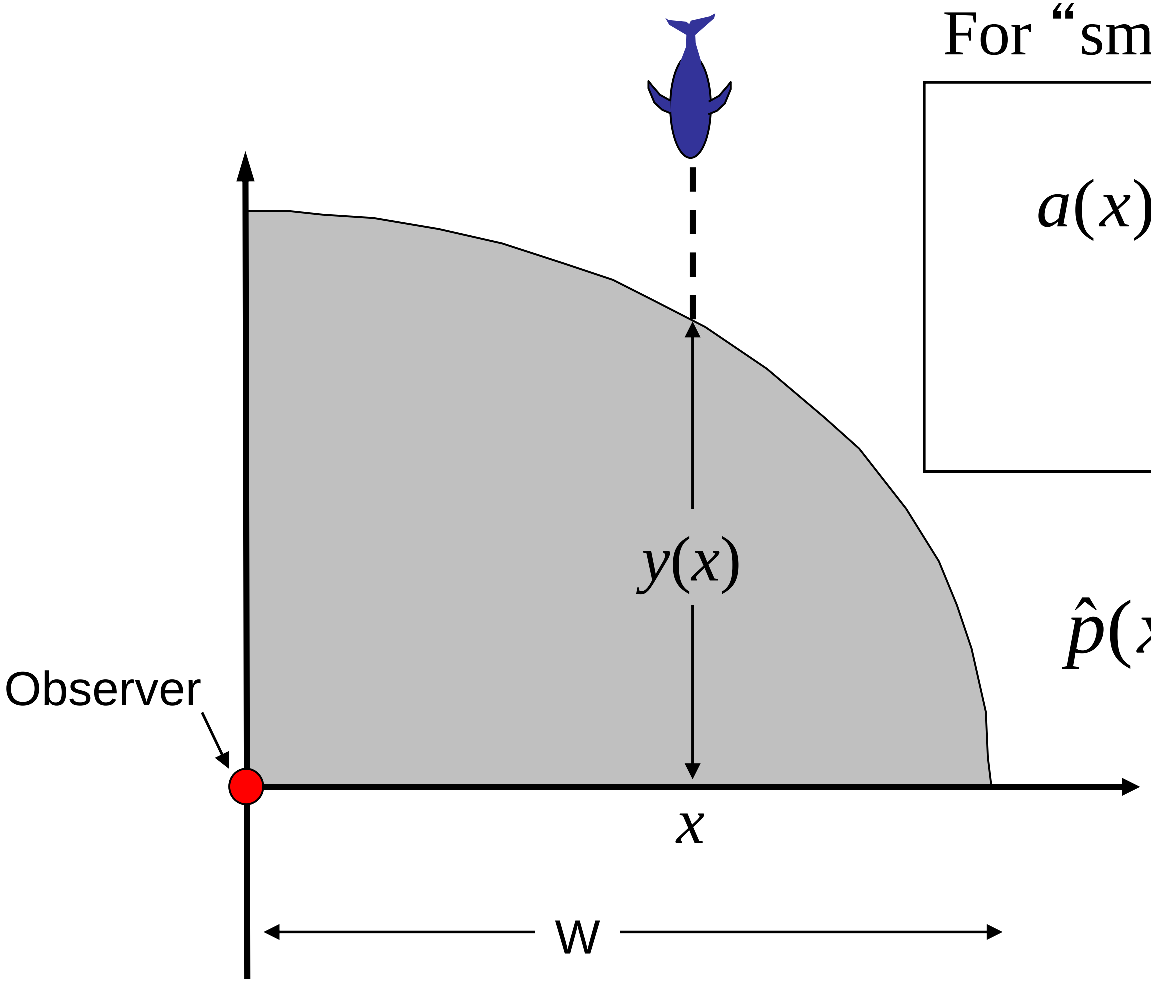
$$a = \frac{s}{s+d}$$

$$a = \frac{E[s]}{E[s]+E[d]}$$

$$\hat{p}(x) = \hat{p}_{available}(x) \times a$$

… But animals at small *x* have longer to be available

Prob(animal at *x* is available **at least once**)=*a*(*x*)



For "small" *y*(*x*):

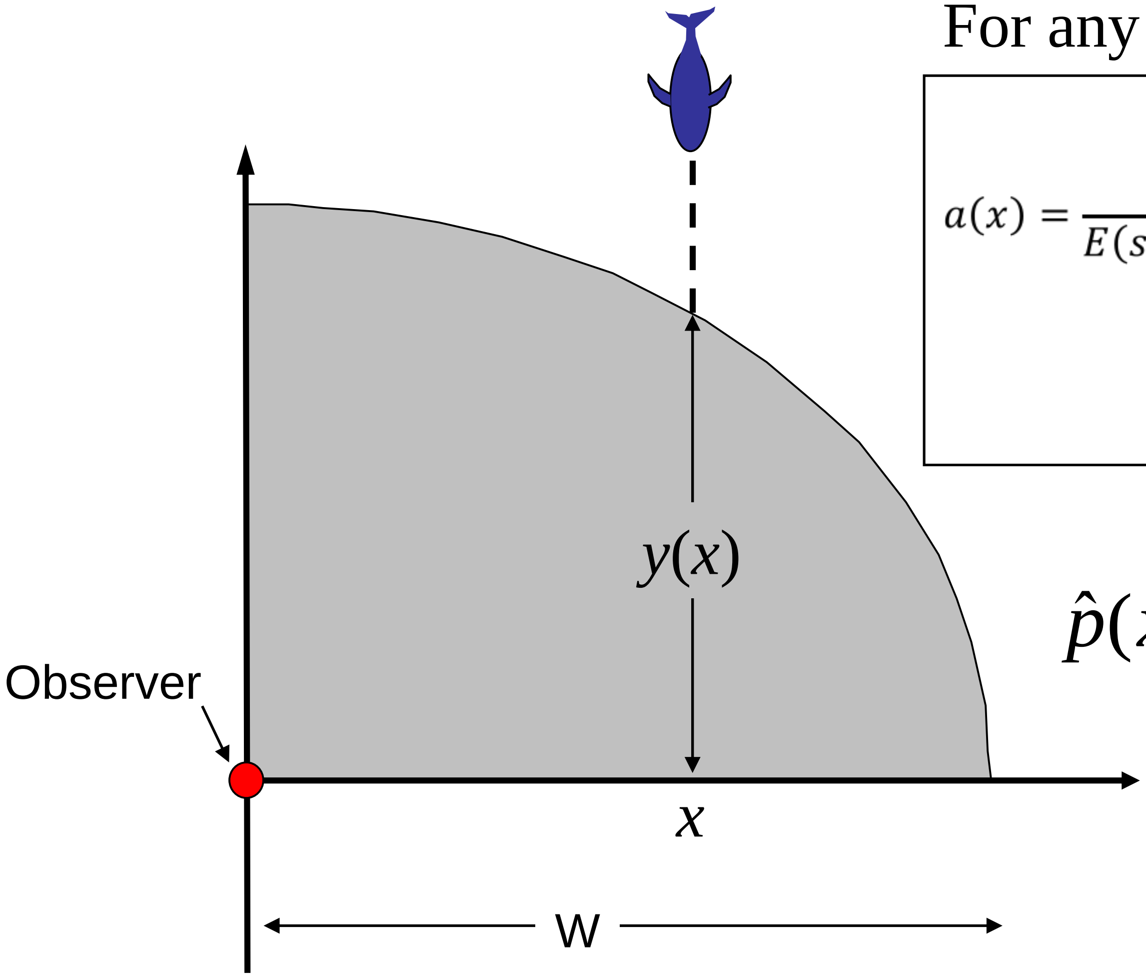
$$a(x) = \frac{s}{s+d} + \frac{y(x)}{s+d}$$

(McLaren, 1961)

$$\hat{p}(x) = \hat{p}_{available}(x) \times a(x)$$

… But McLaren's a(x) can be greater than 1

$$Prob(\text{animal at } x \text{ is available } \textbf{at least once}) = a(x)$$



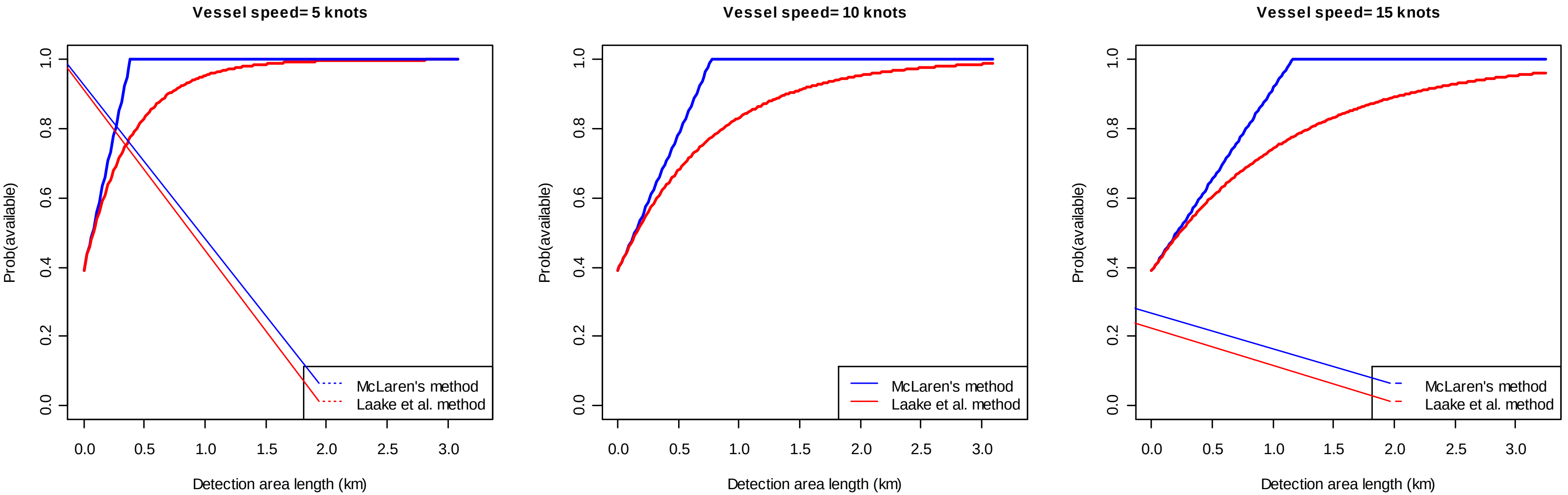
For any $y(x)$:

$$a(x) = \frac{E(s)}{E(s)+E(d)} + \frac{E(s)\left[1-\exp\left\{-\frac{y(x)}{E(d)}\right\}\right]}{E(s)+E(d)}$$

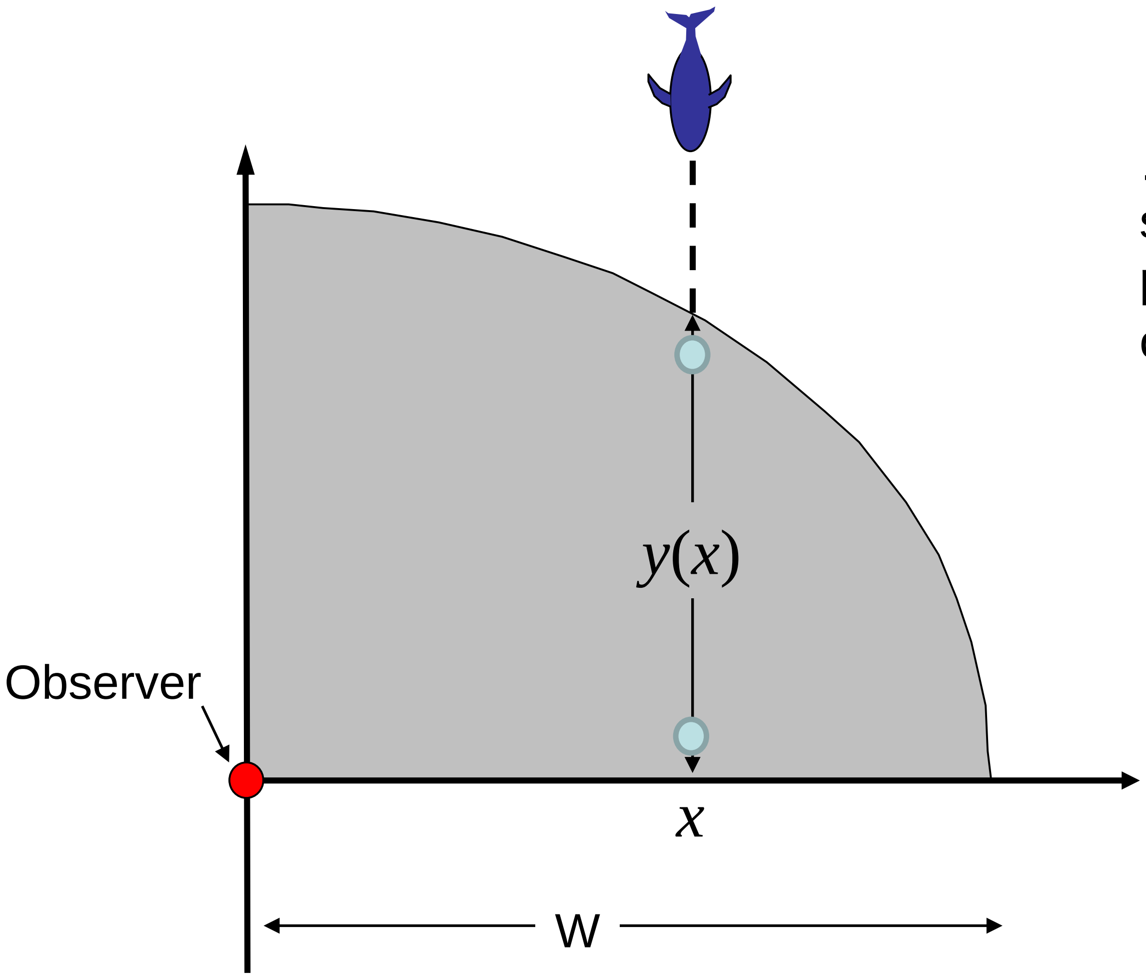
(Laake et al., 1997)

$$\hat{p}(x) = \hat{p}_{available}(x) \times a(x)$$

McLaren’s vs Laake’s Method



… But an animal available at big y is less detectable than an animal available at small y



… which means we should model detection probability as a function of x **and** y.

When there is availability bias

1. McLaren's method just a poorer version of Laake's method.

2. Laake's, McLarn's and similar methods potentially substantially biased. OK when animals in view for short time relative to availability cycle length.

3. Cue-counting an option. Not animal-based

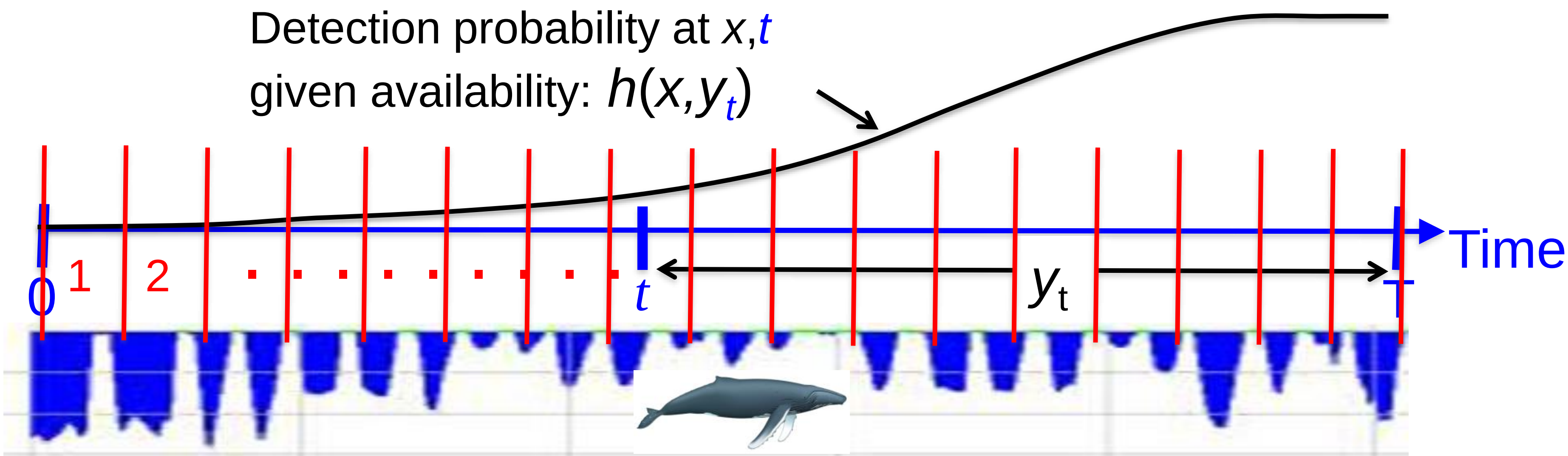
4. Alternatively, model p(x,y) and availability (animal-based):

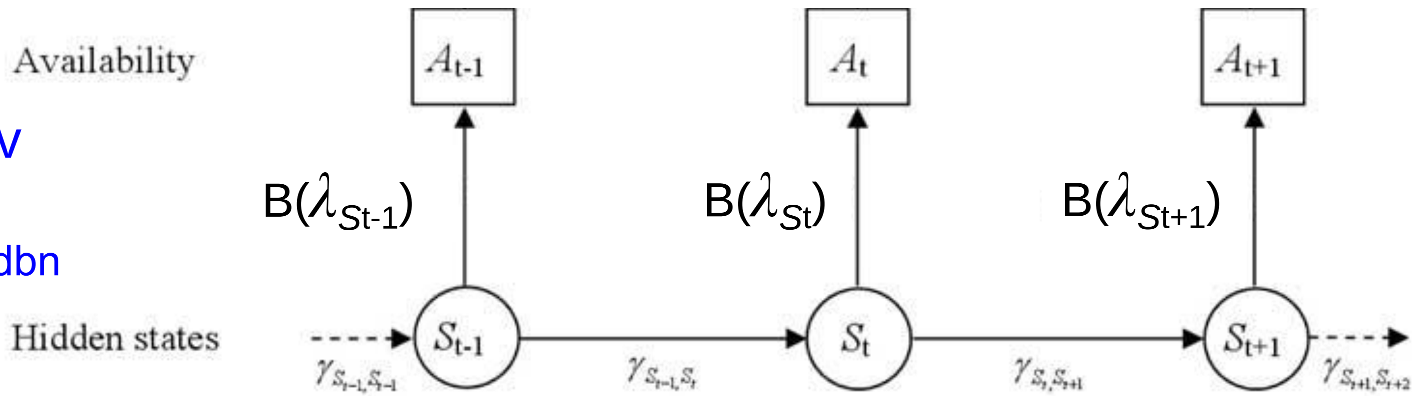
   – Borchers, Zucchini, Heide-Jorgensen & Canadas. Hidden Markov models to deal with availability bias on line transect surveys. *Bometrics* **69**:703-713.

   – Langrock, Borchers & Skaug. Markov-modulated nonhomogeneous Poisson processes for unbiased estimation of marine mammal abundance. *Journal of the American Statistical Association* **108**: 840-851.

   – Borchers, D.L. and Langrock, R. 2015. Double-Observer Line Transect Surveys with Markov-Modulated Poisson Process Models for Overdispersed Animal Availability. *Biometrics* **71**: 1060–1069.

Hidden Markov Model for Availability

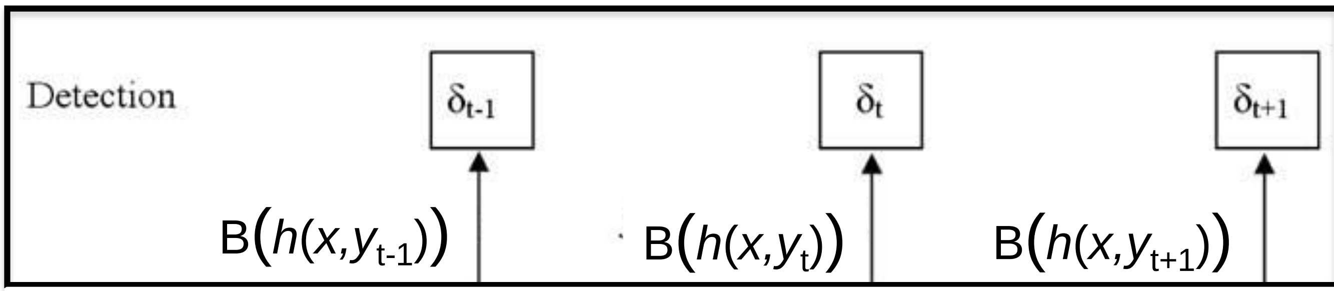




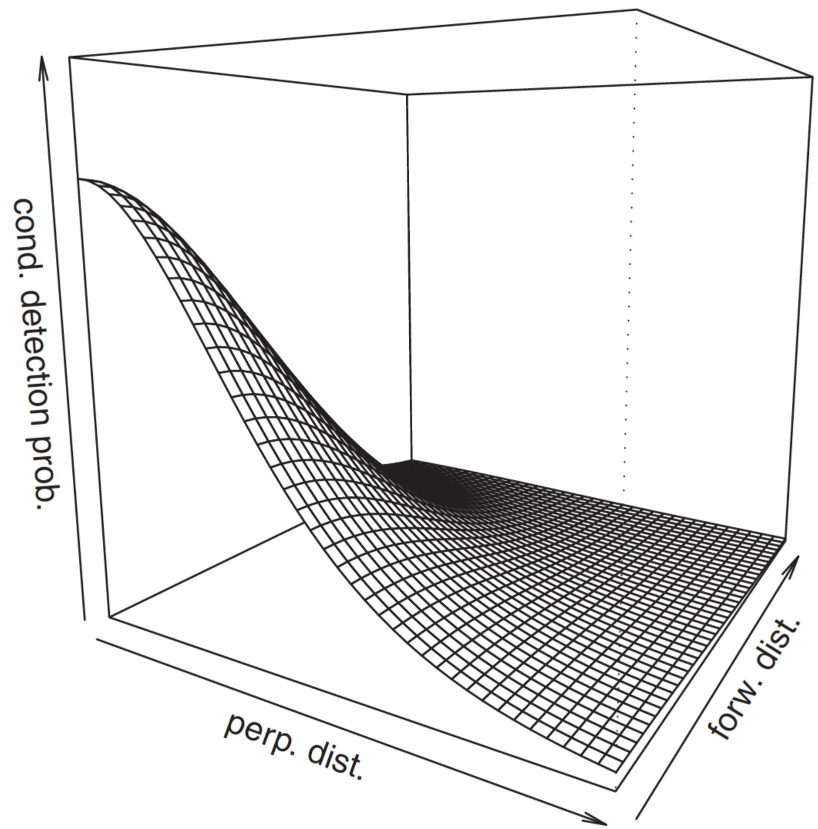
Hidden Markov Model (HMM)

"B( )" = Bernoulli dbn

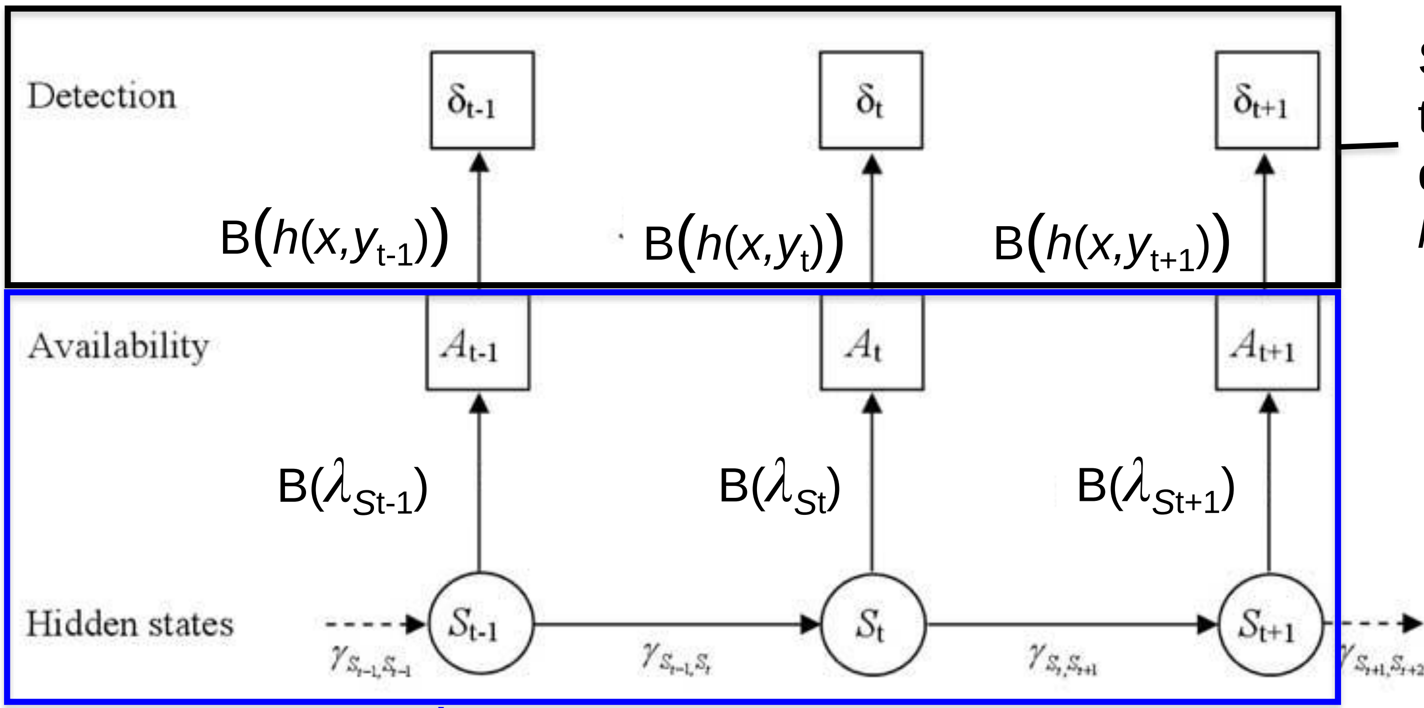
Hidden Markov Model for Estimator



Specify a form for the 2-dimensional detection hazard $h(x,y)$



Hidden Markov Model for Estimator



Specify a form for the 2-dimensional detection hazard $h(x,y_{t+1})$

Specify availability HMM parameters:

(1) State transition probabilities: γs

(2) Probability available given state: λs

Special case: 2-states, one of which is available (with $\lambda=1$), the other unavailable (with $\lambda=0$).

Can specify simply by giving mean times available and unavailable in a single availability cycle.

Summary

- When time in view very small relative to availability cycle, simple correction methods work OK – otherwise they do not.

- HMM (or MMPP) method better, but needs forward distances (so collect them – it is often not difficult).
  - In simplest case, HMM method needs no more data than Laake's method.
  - More general (generally better) case needs a HMM to have been fitted to availability time series data.

- R package hsltm implements HMM method (email/talk to me if interested).