Adaptive nonparametric smoothing for capture-recapture models

James Faulkner [1,2]   Vladimir Minin [1]

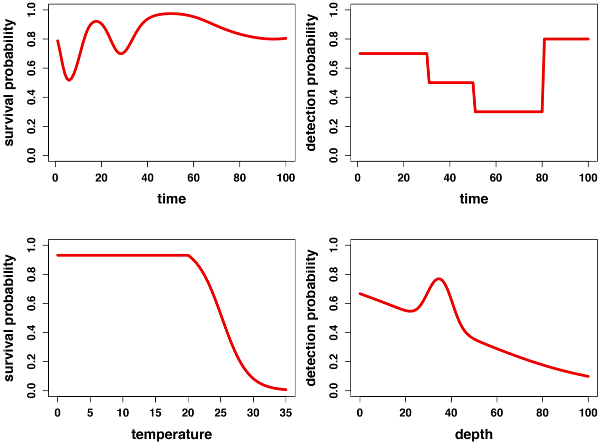
[1]University of Washington

[2]National Oceanic and Atmospheric Administration

IBC 2016

Example functions

Goal: develop nonparametric method able to fit functions with breakpoints and sharp features in a mark-recapture setting.



Current Methods

Some nonparametric smoothing methods used in mark-recapture models:

- P-splines (Gimenez 2006; Bonner and Schwarz 2011)
- Free-knot B-splines (Bonner et al. 2009)
- Gaussian processes (Royle and Dubovsky 2001)
- Conditional autoregressive (Saracco et al. 2010)

Background

Assume detection (survival) probability follows an unknown function $f(s)$, where s is a continuous index of time.

Let $\psi_j = f(j)$ be detection probability at discrete time $j \in \{1, \ldots, m\}$, and let $\theta_j = \text{logit}(\psi_j)$

Background

Assume detection (survival) probability follows an unknown function $f(s)$, where s is a continuous index of time.

Let $\psi_j = f(j)$ be detection probability at discrete time $j \in \{1, \ldots, m\}$, and let $\theta_j = \text{logit}(\psi_j)$

Then a simple kth-order GMRF prior for $\boldsymbol{\theta}$ is induced by letting:

$$\Delta^k \theta_i \sim \text{N}(0, \gamma^2), \qquad i = 1, \ldots, m - k$$

where $\Delta^k \theta_i$ is a kth-order forward difference operator.

Adaptive smoothing prior

We can allow locally-adaptive behavior and increase smoothing properties by putting a **shrinkage prior** on $\Delta^k \theta_i$:

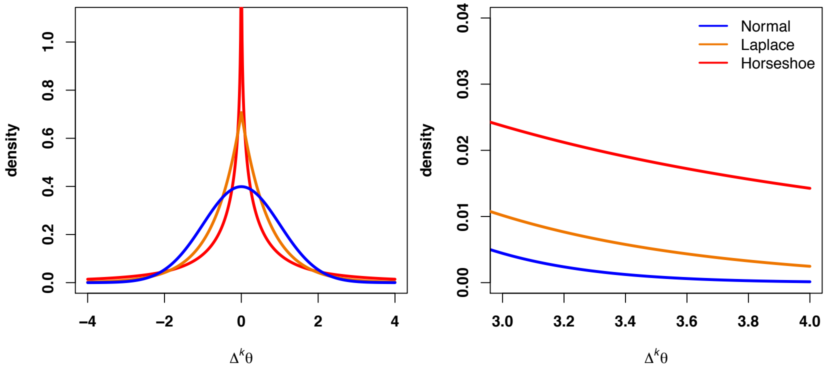
$$\Delta^k \theta_i \sim \text{Horseshoe}(0, \gamma)$$
$$\gamma \sim \text{C}^+(0, \zeta)$$

where γ is the global smoothing parameter.

The result is non-Gaussian (horseshoe) Markov random field prior for $\boldsymbol{\theta}$.

Prior comparisons

Good shrinkage prior has high density at zero and fat tails



Cormack-Jolly-Seber model

Survival process:

Let $s_{i,t} \in \{0,1\}$ be the latent survival state, and $\phi_{i,t}$ be survival probability for individual i at time t, where

$$s_{i,t} \sim \text{Bernoulli}(\phi_{i,t} s_{i,t-1})$$

Cormack-Jolly-Seber model

Survival process:
Let $s_{i,t} \in \{0, 1\}$ be the latent survival state, and $\phi_{i,t}$ be survival probability for individual i at time t, where

$$s_{i,t} \sim \text{Bernoulli}(\phi_{i,t} s_{i,t-1})$$

Observation process:
Let $y_{i,t} \in \{0, 1\}$ be the observation variable, and $\psi_{i,t}$ be detection probability, where

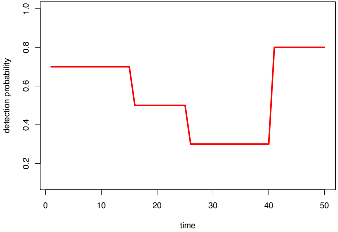
$$y_{i,t} \sim \text{Bernoulli}(\psi_t s_{i,t})$$

Simulations

Simulated mark-recapture data

Scenario 1:

- 50 recapture times
- 100 individuals released, 20 new enter each time
- Constant survival across time ($\phi_t = 0.98$ for all t)
- Piece-wise constant detection probability over time

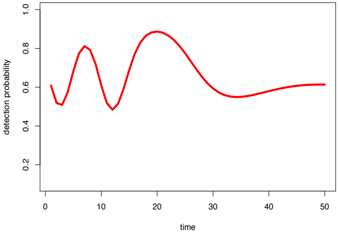


Simulations

Simulated mark-recapture data

Scenario 2:

- 50 recapture times
- 50 individuals released, 10 new enter each time
- Constant survival across time ($\phi_t = 0.90$ for all t)
- Smooth varying detection probability over time



Simulation results

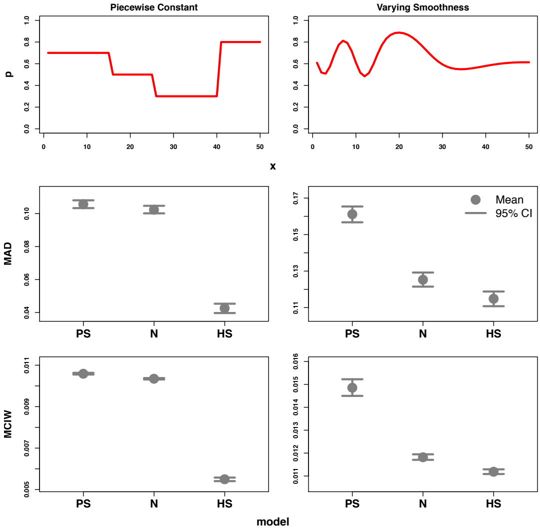
100 Simulations
Hamiltonian Monte Carlo

Models:

- P-spline (PS)
- Normal (N)
- Horseshoe (HS)

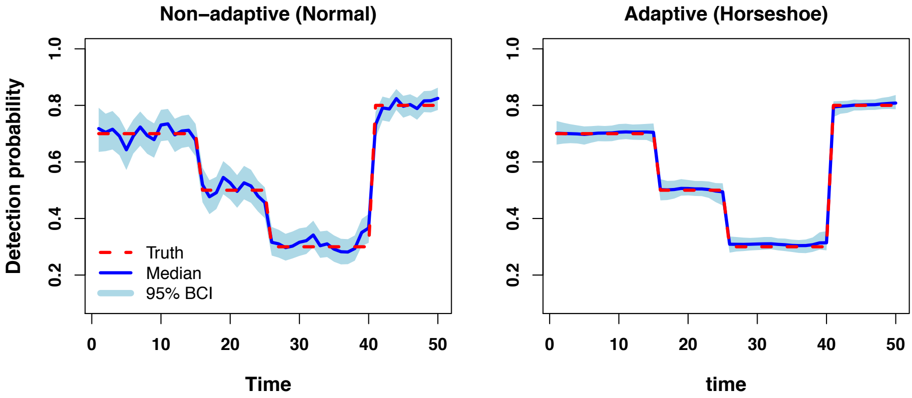
Metrics:

- Mean absolute deviation (MAD)
- Mean credible interval width (MCIW)



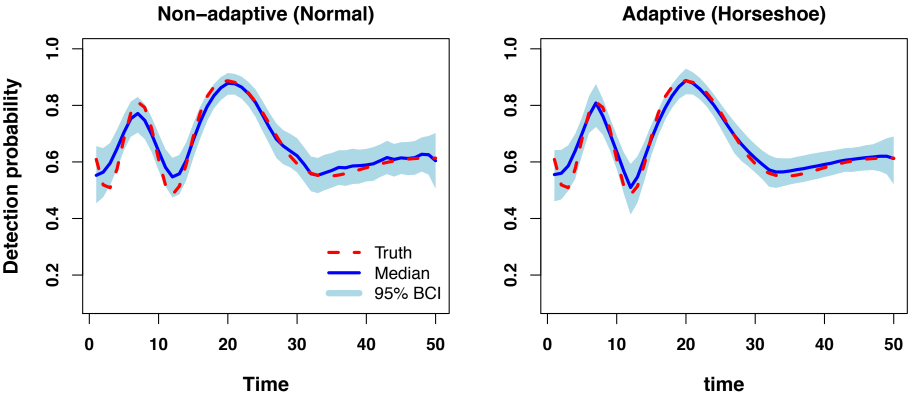
Simulations: example fits

Scenario 1:



Simulations: example fits

Scenario 2:

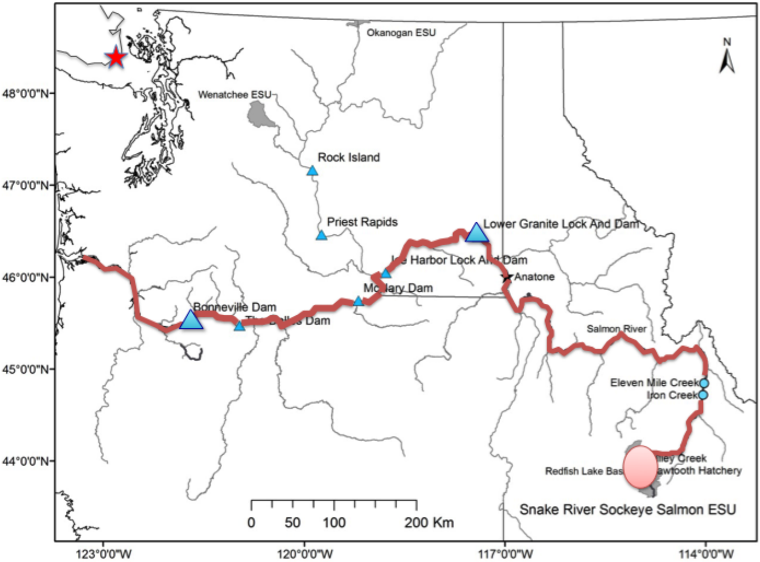


Data Example: Adult Sockeye Survival

- Endangered Snake-River sockeye salmon return in low numbers as adults
- Pass through series of hydroelectric dams where tags are detected
- Mark-recapture models can be used to estimate detection and survival probabilities
- Susceptible to heat stress caused by high water temperatures

Study area

From Crozier et al. (2014)



Methods

Objective: estimate effect of water temperature on individual survival between Bonneville and Lower Granite Dams

Data

- 1,942 individuals from 2008-2015
- All detected at Bonneville Dam
- Average daily temperature over 10 days prior to detection at Bonneville
- 281 unique values of average temperature

Models

Let

- x denote water temperature where x_j is a unique temperature value.
- $\delta_j = x_{j+1} - x_j$ be the difference between adjacent temperature readings.
- $\text{logit}(\phi_j) = \theta_j, \quad$ where ϕ is survival probability
- $\Delta^2\theta_j = \theta_{j+2} - \left(1 + \frac{\delta_{j+1}}{\delta_j}\right)\theta_{j+1} + \frac{\delta_{j+1}}{\delta_j}\theta_j$

Models

1. Non-adaptive GMRF:

$$\Delta^2\theta_j \sim \mathrm{N}(0, d\gamma^2)$$

2. Adaptive MRF:

$$\Delta^2\theta_j \sim \mathrm{HS}(d^{1/2}\gamma)$$

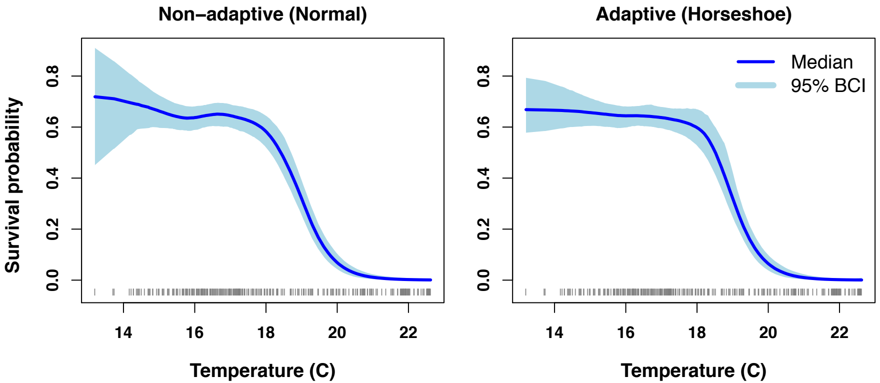
where global smoothing parameter

$$\gamma \sim \mathrm{C}^+(0, 0.1)$$

and

$$d = \frac{\delta_{j+1}^2(\delta_j + \delta_{j+1})}{2}$$

Results


Non–adaptive (Normal)
Adaptive (Horseshoe)
Survival probability
0.0
0.2
0.4
0.6
0.8
14
16
18
20
22
Temperature (C)
Median
95% BCI

Summary

1. Nonparametric smoothing achieved by placing a shrinkage prior on kth-order discrete derivatives
2. The method results in local adaptivity with global control
3. We intend to extend the method to spatial and semi-parametric models

Acknowledgments

- IBC committee members
- University of Washington
- NOAA
- Bonneville Power Adminstration
- QERM at UW
- FE Quant. Ecology Team
- Vladimir Minin
- Minin Lab
- Mark Scheuerell
- Ben Sandford
- Dan Widener

Contact: jfaulknr@uw.edu