

HAUL-OUT SITE ESTIMATION USING A TRUNCATION APPROXIMATION OF A DIRICHLET PROCESS MIXTURE MODEL

Brian M. Brost

24 August 2015

Description

A Dirichlet process mixture model for haul-out site estimation. Unlike the model implemented in `haulout.dp.mixture.2.mcmc.R`, this version models the haul-out site locations ($\mu_{0,h}$) as a function of covariates. As such, this model allows inference on factors influencing harbor seal choice of haul-out sites.

Implementation

The file `haulout.dp.mixture.3.sim.R` simulates data according to the model statement presented below, and `haulout.dp.mixture.3.mcmc.R` contains the MCMC algorithm for model fitting. Dirichlet process mixture model implementation follows the blocked Gibbs sampler truncation approximation of Ishwaran and James (2001) and Gelman et al. (2014).

Model statement

Denote the support of the harbor seal movement process as \mathcal{S} , and the support of haul-out sites as $\tilde{\mathcal{S}}$. Note that \mathcal{S} and $\tilde{\mathcal{S}}$ overlap, i.e., $\tilde{\mathcal{S}} \subset \mathcal{S}$. The domain defined by \mathcal{S} therefore represents at-sea locations or locations of the individual while milling adjacent to the haul-out site. Also note that $\tilde{\mathcal{S}}$, the complement of \mathcal{S} , represents locations inaccessible to the individual (i.e., terrestrial sites that are not haul-outs).

Let $\mathbf{s}(t) = (s_1(t), s_2(t))'$ and $\boldsymbol{\mu}(t) = (\mu_1(t), \mu_2(t))'$ be observed and true locations of a single individual at some time t , respectively. Also let $\boldsymbol{\mu}_{0,h} = (\mu_{0,1,h}, \mu_{0,2,h})'$, for $h = 1, \dots, H$, be the locations of haul-out sites (i.e., cluster centroids), where H is the maximum number of haul-outs allowed per the truncation approximation of the Dirichlet process mixture model. The latent indicator variable $z(t)$ denotes when location $\mathbf{s}(t)$ is on a haul-out site ($z(t) = 1$) or not ($z(t) = 0$). When an individual is hauled-out, note that $\boldsymbol{\mu}(t) = \boldsymbol{\mu}_{0,h_t}$, where h_t acts as a classification variable that identifies the $\boldsymbol{\mu}_{0,h}$ associated with each $\boldsymbol{\mu}(t)$. Furthermore, let $\boldsymbol{\gamma}$ be a vector of coefficients that describe the influence of spatially-referenced covariates contained in the design matrix \mathbf{U} that characterize the environment at all possible haul-out sites in $\tilde{\mathcal{S}}$.

Information pertaining to the wet/dry status of the individual is available from a second data source, $y(\tilde{t})$ (e.g., records from a *.SEA file). Records in y are observed at times \tilde{t} , which may not be the same as the times t at which locations are collected. The binary data in $y(\tilde{t})$ are modeled using an auxiliary variable \mathbf{v} , which itself is modeled semiparametrically as a function of $\boldsymbol{\beta}$, the coefficients that provide inference on covariates of biological interest in the matrix \mathbf{X} , and the parameters $\boldsymbol{\alpha}$ that describe non-linear trend or dependence in wet/dry status through the basis expansion \mathbf{W} . The auxiliary variable \mathbf{v} describes a continuous (in time)

haul-out process, and links the wet/dry status of the individual with its telemetry location.

$$\begin{aligned}
\mathbf{s}(t) &\sim \begin{cases} \mathcal{N}(\boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I}), & z(t) = 1 \\ \mathcal{N}(\boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I}), & z(t) = 0 \end{cases} \\
z(t) &\sim \begin{cases} 0, & v(t) \leq 0 \\ 1, & v(t) > 0 \end{cases} \\
y(\tilde{t}) &\sim \begin{cases} 0, & v(\tilde{t}) \leq 0 \\ 1, & v(\tilde{t}) > 0 \end{cases} \\
\mathbf{v} &\sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha}, \mathbf{I}) \\
h_t &\sim \text{Cat}\left(\frac{\pi_h}{\sum_{\tilde{h}=1}^H \pi_{\tilde{h}}}\right) \\
\pi_h &\sim \text{Stick}(\theta) \\
\boldsymbol{\mu}_{0,h} &\sim \frac{\exp\{\mathbf{u}'(\boldsymbol{\mu}_{0,h})\boldsymbol{\gamma}\}}{\sum_{\tilde{h}=1}^H \exp\{\mathbf{u}'(\boldsymbol{\mu}_{0,\tilde{h}})\boldsymbol{\gamma}\}} \\
\theta &\sim \text{Gamma}(r, q) \\
\boldsymbol{\beta} &\sim \mathcal{N}(\boldsymbol{\mu}_\beta, \sigma_\beta^2 \mathbf{I}) \\
\boldsymbol{\alpha} &\sim \mathcal{N}(\mathbf{0}, \sigma_\alpha^2 \mathbf{I}) \\
\boldsymbol{\gamma} &\sim \mathcal{N}(\boldsymbol{\mu}_\gamma, \sigma_\gamma^2 \mathbf{I}) \\
\sigma &\sim \text{Unif}(l_\sigma, u_\sigma) \\
\sigma_\mu &\sim \text{Unif}(l_{\sigma_\mu}, u_{\sigma_\mu})
\end{aligned}$$

The concentration parameter θ affects the clustering in the Dirichlet process mixture: smaller values yield fewer clusters with more observations per cluster, whereas larger values yield more clusters with fewer observations per cluster. Note that the lines in this model statement pertaining to h_t , π_h , and $\boldsymbol{\mu}_{0,h}$ comprise the stick-breaking representation of the Dirichlet process mixture model, i.e.,

$$\begin{aligned}
\boldsymbol{\mu}_{0,h} &\sim \mathbf{P} \\
\mathbf{P} &\sim \text{DP}(\theta, \mathbf{P}_0) \\
\mathbf{P}_0 &\sim \frac{\exp\{\mathbf{u}'(\boldsymbol{\mu}_{0,h})\boldsymbol{\gamma}\}}{\int_{\mathcal{S}} \exp\{\mathbf{u}'(\boldsymbol{\mu}_0)\boldsymbol{\gamma}\} d\boldsymbol{\mu}_0}
\end{aligned}$$

Full conditional distributions

Haul-out site locations ($\boldsymbol{\mu}_{0,h}$):

$$\begin{aligned}
[\boldsymbol{\mu}_{0,h} | \cdot] &\propto \prod_{\{t: h_t=h\}} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h}, z(t), \sigma^2, \sigma_\mu^2] [\boldsymbol{\mu}_{0,h} | \gamma, \tilde{\mathcal{S}}] \\
&\propto \prod_{\{t: h_t=h\}} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h}, \sigma^2]^{z(t)} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h}, \sigma^2, \sigma_\mu^2]^{1-z(t)} [\boldsymbol{\mu}_{0,h} | \gamma, \tilde{\mathcal{S}}] \\
&\propto \prod_{\{t: h_t=h\}} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h}, \sigma^2 \mathbf{I})^{z(t)} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})^{1-z(t)} \left(\frac{\exp \{ \mathbf{u}'(\boldsymbol{\mu}_{0,h}) \gamma \}}{\sum_{\tilde{h}=1}^H \exp \{ \mathbf{u}'(\boldsymbol{\mu}_{0,\tilde{h}}) \gamma \}} \right) 1_{\{\boldsymbol{\mu}_{0,h} \in \tilde{\mathcal{S}}\}}
\end{aligned}$$

The update for $\boldsymbol{\mu}_{0,h}$ proceeds using Metropolis-Hastings. Note that the product is over all $\mathbf{s}(t)$ that are members of $\boldsymbol{\mu}_{0,h}$, and proposals for $\boldsymbol{\mu}_{0,h}$ not in $\tilde{\mathcal{S}}$ are rejected outright. For each of the unoccupied haul-out locations (i.e., clusters with no members), $\boldsymbol{\mu}_{0,h}$ is sampled directly from the prior, i.e., with probability proportional to $\mathbf{u}'(\boldsymbol{\mu}_{0,h}) \gamma$.

Stick-breaking process (π_h):

The stick-breaking representation of a Dirichlet process mixture consists of two components, namely a cluster weight and a cluster probability. Let η_h denote the weight assigned to cluster h , where $\eta_h \sim \text{Beta}(1, \theta)$. The associated full-conditional is

$$[\eta_h | \cdot] \sim \text{Beta} \left(1 + n_h, \theta + \sum_{\tilde{h}=h+1}^H n_{\tilde{h}} \right), \text{ for } h = 1, \dots, H-1,$$

and $\eta_H = 1$. The parameter n_h denotes the number of observations allocated to cluster h . Note that η_h is sampled in order of decreasing n_h , i.e., n_h is sorted largest to smallest and η_h is sampled in sequence. The cluster probabilities (π_h) are deterministic and calculated as

$$\pi_h = \eta_h \prod_{\tilde{h} < h} (1 - \eta_{\tilde{h}}).$$

The probabilities π_h are also calculated in order of decreasing n_h . See page 553 in Gelman et al. (2014) and Section 5.2 in Ishwaran and James (2001).

Dirichlet process concentration parameter (θ):

$$[\theta | \cdot] \propto \text{Gamma}(r + H - 1, q - \sum_{h=1}^{H-1} \log(1 - \eta_h)).$$

See page 553 in Gelman et al. (2014).

Haul-out classification variable (h_t):

$$\begin{aligned}
[h_t | \cdot] &\sim [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, z(t), \sigma^2, \sigma_\mu^2] [h_t | \pi_h] \\
&\sim \text{Cat} \left(\frac{\pi_h [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, z(t), \sigma^2, \sigma_\mu^2]}{\sum_{\tilde{h}=1}^H \pi_{\tilde{h}} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,\tilde{h}}, z(t), \sigma^2, \sigma_\mu^2]} \right) \\
&\sim \text{Cat} \left(\frac{\pi_h [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2]^{z(t)} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma_\mu^2]^{1-z(t)}}{\sum_{\tilde{h}=1}^H \pi_{\tilde{h}} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,\tilde{h}}, \sigma^2]^{z(t)} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,\tilde{h}}, \sigma^2, \sigma_\mu^2]^{1-z(t)}} \right) \\
&\sim \text{Cat} \left(\frac{\pi_h \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2)^{z(t)} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})^{1-z(t)}}{\sum_{\tilde{h}=1}^H \pi_{\tilde{h}} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,\tilde{h}}, \sigma^2)^{z(t)} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,\tilde{h}}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})^{1-z(t)}} \right).
\end{aligned}$$

This update proceeds just as in multinomial sampling; see page 552 in Gelman et al. (2014).

Latent haul-out indicator variable for locations ($z(t)$):

$$\begin{aligned}
[z(t) | \cdot] &\propto [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h}, z(t), \sigma^2, \sigma_\mu^2] [z(t) | v(t)] \\
&\propto [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2]^{z(t)} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma_\mu^2]^{1-z(t)} [z(t) | v(t)] \\
&\propto \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I})^{z(t)} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})^{1-z(t)} \text{Bern}(z(t) | v(t)) \\
&\propto \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I})^{z(t)} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})^{1-z(t)} p(t)^{z(t)} (1-p(t))^{1-z(t)} \\
&\propto (p(t) \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I}))^{z(t)} ((1-p(t)) \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma_\mu^2))^{1-z(t)} \\
&= \text{Bern}(\widetilde{p(t)}),
\end{aligned}$$

where $v(t) = \mathbf{x}'(t)\boldsymbol{\beta} + \mathbf{w}'(t)\boldsymbol{\alpha}$, $p(t) = \Phi(v(t))$, and

$$\widetilde{p(t)} = \frac{p(t) \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I})}{p(t) \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I}) + (1-p(t)) \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})}.$$

Auxiliary variable for continuous haul-out process (\mathbf{v}):

$$\begin{aligned}
[\mathbf{v} | \cdot] &\propto \prod_{\{t, \tilde{t}\}} [z(t), y(\tilde{t}) | \mathbf{v}] [\mathbf{v} | \boldsymbol{\alpha}, \boldsymbol{\beta}] \\
&\propto \prod_t \{[z(t) | v(t)] [v(t) | \boldsymbol{\alpha}, \boldsymbol{\beta}]\} \prod_{\tilde{t}} \{[y(\tilde{t}) | v(\tilde{t})] [v(\tilde{t}) | \boldsymbol{\alpha}, \boldsymbol{\beta}]\} \\
&\propto \prod_t \{ (1_{\{z(t)=0\}} 1_{\{v(t) \leq 0\}} + 1_{\{z(t)=1\}} 1_{\{v(t) > 0\}}) \times \mathcal{N}(v(t) | \mathbf{x}(t)' \boldsymbol{\beta} + \mathbf{w}(t)' \boldsymbol{\alpha}, 1) \} \times \\
&\quad \prod_{\tilde{t}} \{ (1_{\{y(\tilde{t})=0\}} 1_{\{v(\tilde{t}) \leq 0\}} + 1_{\{y(\tilde{t})=1\}} 1_{\{v(\tilde{t}) > 0\}}) \times \mathcal{N}(v(\tilde{t}) | \mathbf{x}(\tilde{t})' \boldsymbol{\beta} + \mathbf{w}(\tilde{t})' \boldsymbol{\alpha}, 1) \}
\end{aligned}$$

For wet/dry data (y) observed at times \tilde{t} ,

$$\begin{aligned}
[v(\tilde{t}) | \cdot] &\propto [y(\tilde{t}) | v(\tilde{t})] [v(\tilde{t}) | \boldsymbol{\alpha}, \boldsymbol{\beta}] \\
&\propto (1_{\{y(\tilde{t})=0\}} 1_{\{v(\tilde{t}) \leq 0\}} + 1_{\{y(\tilde{t})=1\}} 1_{\{v(\tilde{t}) > 0\}}) \times \mathcal{N}(v(\tilde{t}) | \mathbf{x}(\tilde{t})' \boldsymbol{\beta} + \mathbf{w}(\tilde{t})' \boldsymbol{\alpha}, 1) \\
&= \begin{cases} \mathcal{TN}(\mathbf{x}(\tilde{t})' \boldsymbol{\beta} + \mathbf{w}(\tilde{t})' \boldsymbol{\alpha}, 1)_{-\infty}^0, & y(\tilde{t}) = 0 \\ \mathcal{TN}(\mathbf{x}(\tilde{t})' \boldsymbol{\beta} + \mathbf{w}(\tilde{t})' \boldsymbol{\alpha}, 1)_{0}^{\infty}, & y(\tilde{t}) = 1 \end{cases}
\end{aligned}$$

For the latent wet/dry status (z) of telemetry locations at times t ,

$$\begin{aligned}
[v(t) | \cdot] &\propto [z(t) | v(t)] [v(t) | \boldsymbol{\alpha}, \boldsymbol{\beta}] \\
&\propto (1_{\{z(t)=0\}} 1_{\{v(t) \leq 0\}} + 1_{\{z(t)=1\}} 1_{\{v(t) > 0\}}) \times \mathcal{N}(v(t) | \mathbf{x}(t)' \boldsymbol{\beta} + \mathbf{w}(t)' \boldsymbol{\alpha}, 1) \\
&= \begin{cases} \mathcal{TN}(\mathbf{x}(t)' \boldsymbol{\beta} + \mathbf{w}(t)' \boldsymbol{\alpha}, 1)_{-\infty}^0, & z(t) = 0 \\ \mathcal{TN}(\mathbf{x}(t)' \boldsymbol{\beta} + \mathbf{w}(t)' \boldsymbol{\alpha}, 1)_{0}^{\infty}, & z(t) = 1 \end{cases}
\end{aligned}$$

Haul-out process coefficients (β):

$$\begin{aligned}
[\beta|\cdot] &\propto [\mathbf{v}|\beta, \alpha][\beta | \mu_\beta, \sigma_\beta^2] \\
&\propto \mathcal{N}(\mathbf{v} | \mathbf{X}\beta + \mathbf{W}\alpha, \mathbf{1}) \mathcal{N}(\mu_\beta | \mathbf{0}, \sigma_\beta^2 \mathbf{I}) \\
&\propto \exp \left\{ -\frac{1}{2} (\mathbf{v} - (\mathbf{X}\beta + \mathbf{W}\alpha))' (\mathbf{v} - (\mathbf{X}\beta + \mathbf{W}\alpha)) \right\} \times \\
&\quad \exp \left\{ -\frac{1}{2} (\beta - \mu_\beta)' (\sigma_\beta^2 \mathbf{I})^{-1} (\beta - \mu_\beta) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} ((\mathbf{v} - \mathbf{W}\alpha) - \mathbf{X}\beta)' ((\mathbf{v} - \mathbf{W}\alpha) - \mathbf{X}\beta) \right\} \times \\
&\quad \exp \left\{ -\frac{1}{2} (\beta - \mu_\beta)' (\sigma_\beta^2 \mathbf{I})^{-1} (\beta - \mu_\beta) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left(-2 \left((\mathbf{v} - \mathbf{W}\alpha) \mathbf{X} + \mu_\beta (\sigma_\beta^2 \mathbf{I})^{-1} \right) \beta + \beta' \left(\mathbf{X}' \mathbf{X} + (\sigma_\beta^2 \mathbf{I})^{-1} \right) \beta \right) \right\} \\
&= \mathcal{N}(\mathbf{A}^{-1} \mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}$$

where $\mathbf{A} = \mathbf{X}' \mathbf{X} + (\sigma_\beta^2 \mathbf{I})^{-1}$ and $\mathbf{b}' = (\mathbf{v} - \mathbf{W}\alpha)' \mathbf{X} + \mu_\beta' (\sigma_\beta^2 \mathbf{I})^{-1}$. Note that the matrices \mathbf{X} and \mathbf{W} contain covariates pertaining to times associated with both $\mathbf{s}(t)$ and $y(\tilde{t})$.

Coefficients for non-linear trend/dependence (α):

$$\begin{aligned}
[\alpha|\cdot] &\propto [\mathbf{v}|\beta, \alpha][\alpha | \mathbf{0}, \sigma_\alpha^2] \\
&\propto \mathcal{N}(\mathbf{v} | \mathbf{X}\beta + \mathbf{W}\alpha, \mathbf{1}) \mathcal{N}(\alpha | \mathbf{0}, \sigma_\alpha^{-1} \mathbf{I}) \\
&\propto \exp \left\{ -\frac{1}{2} (\mathbf{v} - (\mathbf{X}\beta + \mathbf{W}\alpha))' (\mathbf{v} - (\mathbf{X}\beta + \mathbf{W}\alpha)) \right\} \times \\
&\quad \exp \left\{ -\frac{1}{2} (\alpha - \mathbf{0})' (\sigma_\alpha^2 \mathbf{I})^{-1} (\alpha - \mathbf{0}) \right\} \\
&= \mathcal{N}(\mathbf{A}^{-1} \mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}$$

where $\mathbf{A} = \mathbf{W}' \mathbf{W} + (\sigma_\alpha^2 \mathbf{I})^{-1}$ and $\mathbf{b}' = (\mathbf{v}' - \mathbf{X}\beta)' \mathbf{W}$. Note that the matrices \mathbf{X} and \mathbf{W} contain covariates pertaining to times associated with both $\mathbf{s}(t)$ and $y(\tilde{t})$.

Haul-out location coefficients (γ):

$$\begin{aligned}
[\gamma|\cdot] &\propto \prod_{h=1}^{\tilde{H}} [\mu_{0,h} | \gamma][\gamma] \\
&\propto \prod_{h=1}^{\tilde{H}} \left(\frac{\exp \{ \mathbf{u}'(\mu_{0,h}) \gamma \}}{\int_{\tilde{\mathcal{S}}} \exp \{ \mathbf{u}'(\mu_0) \gamma \} d\mu_0} \right) \mathcal{N}(\mu_\gamma, \sigma_\gamma^2 \mathbf{I})
\end{aligned}$$

The update for γ proceeds using Metropolis-Hastings. Note that the product is over all occupied clusters (i.e., with non-zero membership) $h = 1, \dots, \tilde{H}$.

Error in the observation process (σ):

$$\begin{aligned}
[\sigma|\cdot] &\propto \prod_t [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, z(t), \sigma^2, \sigma_\mu^2][\sigma] \\
&\propto \prod_t [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2]^{z(t)} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma_\mu^2]^{1-z(t)}[\sigma] \\
&\propto \prod_t \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I})^{z(t)} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})^{1-z(t)}
\end{aligned}$$

The update for σ proceeds using Metropolis-Hastings.

Homerange dispersion parameter (σ_μ):

$$\begin{aligned}
[\sigma_\mu|\cdot] &\propto \prod_{t\in\{t:z(t)=0\}} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, z(t), \sigma^2, \sigma_\mu^2][\sigma_\mu] \\
&\propto \prod_{t\in\{t:z(t)=0\}} [\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma_\mu^2]^{1-z(t)}[\sigma_\mu] \\
&\propto \prod_{t\in\{t:z(t)=0\}} \mathcal{N}(\mathbf{s}(t) | \boldsymbol{\mu}_{0,h_t}, \sigma^2 \mathbf{I} + \sigma_\mu^2 \mathbf{I})
\end{aligned}$$

The update for σ_μ proceeds using Metropolis-Hastings. Note that the product is over all t for which $z(t) = 0$.

References

- Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. 2014. Bayesian data analysis. CRC Press.
- Ishwaran, H., and L.F. James. 2001. Gibbs sampling methods for stick-breaking priors. Journal of the American Statistical Association 96: 161–173.