Management performance evaluation of state-space models for Pacific pink salmon stock-recruitment analysis (Equations)

Zhenming Su
Ann Arbor, MI

Feb. 12, 2023

This document provides the full conditional posterior distributions used in the MCMC sampling algorithm for a state-space model proposed by Su (2023) "Management performance evaluation of state-space models for Pacific pink salmon stock-recruitment analysis" submitted for peer review and publication.

The model

The state-space form of Ricker stock-recruitment model incorporates time-varying productivity, observation errors in spawners and catch data, a model of harvest rates, and process variability in recruitment process. Time-varying productivity is modeled by representing the Ricker productivity 'a' parameter as a random walk.

Observation equation:

Let $y_t^s = \log\left(S_t^{obs}\right) = \log\,(E_t)$, $s_t = \log(S_t)$, $y_t^c = \log(C_t^{obs})$, $c_t = \log(C_t)$.

$$\begin{cases} y_t^s \sim N(s_t, \sigma_s^2), t = 1, 2, \dots, T \\ y_t^c \sim N(c_t, \sigma_c^2), t = k + 1, \dots, T \end{cases}$$

System equation:

Let log(recruitment) $r_t = \log(R_t)$, a_t = Ricker productivity, k = fixed maturity and return age.

$$\begin{cases} r_t \sim N(a_{t-k} + s_{t-k} - \beta S_{t-k}, \tau_r^2), t = k + 1, \dots, T \\ a_t \sim N(a_{t-1}, \tau_a^2), t = 2, \dots, T - k \\ \lambda_t \sim N(\lambda_{t-1}, \tau_\lambda^2), t = k + 2, \dots, T \end{cases}$$

$$\begin{cases} h_t = \exp(\lambda_t)/(1 + \exp(\lambda_t)) \\ C_t = h_t R_t \\ S_t = (1 - h_t) R_t \end{cases}$$

Initial conditions

$$\begin{cases} s_1, \dots, s_k \sim N(m_s, V_s) \\ \qquad m_s, V_s - prior\ mean\ and\ variance \\ a_{(1)} = a_1 \sim N(m_a, V_a) \\ \lambda_{(1)} = \lambda_{k+1} \sim N(m_{\lambda_1}, V_{\lambda_1}) \end{cases}$$

Parameters and states

$$\{\{a_t\}_{t=1}^{T-k}, \beta, \sigma_s^2, \sigma_c^2, \tau_a^2, \tau_r^2, \tau_\lambda^2, s_1, \dots, s_k, \{r_t\}_{t=k+1}^{T}, \{\lambda_t\}_{t=k+1}^{T}\}$$

Data: observed escapement E_t catch C_t

Data collection started with k escapement (spawner abundance) observations from t = 1 to k: E_1 to E_k. Returns and catch are available from $t = k + 1$ to T.

Models and variables

t	*year*	E_t	S_t	R_t	h_t	λ_t	C_t	C_t^{obs}	a_t
1	1960	E_1	$\boldsymbol{S_1}$						a_1
2	1961	E_2	$\boldsymbol{S_2}$						a_2
3	1962	E_3	S_3	R_3	h_3	λ_3	C_3	C_3^{obs}	a_3
4	1963	E_4	S_4	R_4	h_4	λ_4	C_4	C_4^{obs}	a_4
5	1964	E_5	S_5	$\boldsymbol{R_5}$	h_5	λ_5	C_5	C_5^{obs}	a_t
	...	...	...						
t		E_t	S_t	R_t	h_t	λ_t	C_t	C_t^{obs}	a_t
	...	...	...						
T-2		E_{T-2}	S_{T-2}						a_{T-2}
T-1		E_{T-1}	S_{T-1}	R_{T-1}	h_{T-1}	λ_{T-1}	C_{T-1}	C_{T-1}^{obs}	$\hat{a}_{T-1}$
T		E_T	S_T	R_T	h_T	λ_T	C_T	C_T^{obs}	$\hat{a}_T$
T+1				$\hat{R}_{T+1}$					

Note: $E_t = S_t^{obs}$

Bayesian estimation

Prior

$p(\theta) =$
$U(\sigma_s \mid 0, \infty)U(\sigma_c \mid 0, \infty)U(\tau_a \mid 0, \infty)$
$U(\tau_r \mid 0, \infty)U(\tau_\lambda \mid 0, \infty)$
$N(a_{(1)} \mid m_a, V_a)N(\beta \mid m_b, V_b)N(\lambda_{(1)} \mid m_\lambda, V_\lambda)$
$N(s_1 \mid m_s, V_s) \dots N(s_k \mid m_s, V_s)$

Posterior distribution

$y_t^s = \log(S_t^{obs})$, $y_t^c = \log(C_t^{obs})$, $r_t = \log(R_t)$, $s_t = \log(S_t) = \log((1 - h_t)R_t)$, and $c_t = \log(C_t) = \log(h_t R_t)$

Also note that:

$h_t = \exp(\lambda_t)/(1 + \exp(\lambda_t))$, $\log(h_t) = \lambda_t - \log(1 + \exp(\lambda_t))$, and $\log(1 - h_t) = -\log(1 + \exp(\lambda_t))$

$c_t = \log(C_t) = r_t + \log(h_t) = r_t + \lambda_t - \log(1 + \exp(\lambda_t))$

$s_t = \log(S_t) = r_t + \log(1 - h_t) = r_t - \log(1 + \exp(\lambda_t))$

Posterior for all unknowns:

$$
\begin{aligned}
&p(\theta \mid D) \\
&\propto p(\theta) \\
&\times \prod_{t=1}^{T} N(y_t^s \mid s_t, \sigma_s^2) \\
&\times \prod_{t=k+1}^{T} N(y_{c,t} \mid c_t, \sigma_c^2) \\
&\times \prod_{t=k+1}^{T} N(r_t \mid (a_{t-k} + s_{t-k} - \beta S_{t-k}), \tau_r^2) \\
&\times \prod_{t=2}^{T-k} N(a_t \mid a_{t-1}, \tau_a^2) \\
&\times \prod_{t=k+2}^{T} N(\lambda_t \mid \lambda_{t-1}, \tau_\lambda^2)
\end{aligned}
$$

Full conditionals distributions

Full conditional of σ_s^2

a) Inverse-Gamma prior on σ_s^2

Prior of σ_s^2

$$p(\sigma_s^2) = IG(c, d)$$

Full conditional density of σ_s^2 based on an inverse-Gamma prior for σ_s^2:

$$(\sigma_s^2 \mid .) \sim IG\left(c + T/2, d + \sum_{t=1}^{T} (y_t^s - s_t)^2 /2\right)$$

b) Uniform prior on σ_s

Prior of σ_s

$$p(\sigma_s) \propto 1, \text{with } \sigma_s > 0$$

Equivalently,

$$p(\sigma_s^2) \propto (\sigma_s^2)^{-\frac{1}{2}} = IG(-1/2, 0)$$

$$c = -\frac{1}{2}, d = 0$$

Full conditional density of σ_s^2:

$$(\sigma_s^2 \mid .) \sim IG\left((T-1)/2, \sum_{t=1}^{T} (y_t^s - s_t)^2 /2\right)$$

Full conditional of σ_c^2

a) Uniform prior on σ_c

$$p(\sigma_c) \propto 1, \text{with } \sigma_c > 0$$

Equivalently:

$$p(\sigma_c^2) \propto (\sigma_c^2)^{-1/2} = IG(-0.5,0)$$

Conditional posterior density of σ_c^2:

$$(\sigma_c^2 \mid .) \sim IG\left((T-(k+1))/2, \sum_{t=k+1}^{T} (y_t^c - c_t)^2 /2\right)$$

b) Inverse-Gamma prior on σ_c^2

$$p(\sigma_c^2) = IG(c, d)$$

$$(\sigma_c^2 \mid .) \sim IG\left(c + (T-k)/2, d + \sum_{t=k+1}^{T} (y_t^c - c_t)^2 /2\right)$$

Full conditional of τ_r^2

a) Uniform prior on τ_r

$$p(\tau_r) \propto 1, \text{with } \tau_r > 0$$

Equivalently:

$$p(\tau_r^2) \propto (\tau_r^2)^{-1/2} = IG(-0.5,0)$$

Conditional posterior density of τ_r^2:

$$(\tau_r^2 \mid .) \sim IG\left((T-k-1)/2, \sum_{t=k+1}^{T} (r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k}))^2 /2\right)$$

b) Inverse-Gamma prior on σ_r^2

$$p(\tau_r^2) = IG(c, d)$$

$$(\tau_r^2 \mid .) \sim IG\left(c + (T-k)/2, d + \sum_{t=k+1}^{T} (r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k}))^2 /2\right)$$

Full conditional of τ_λ^2 : independent $\{\lambda\}_{t=k+1}^{T}$

$$\tau_\lambda^2 = 0$$

Full conditional of τ_λ^2 for random walk $\boldsymbol{\lambda}$

a) Uniform prior on τ_λ

$$p(\tau_\lambda) \propto 1, \text{with } \tau_\lambda > 0$$

Equivalently:

$$p(\tau_\lambda^2) \propto (\tau_\lambda^2)^{-1/2} = IG(-0.5,0)$$

Conditional posterior density of τ_λ^2:

$$(\tau_\lambda^2 \mid .) \sim IG\left(\frac{T-k-2}{2}, \sum_{t=k+2}^{T} (\lambda_t - \lambda_{t-1})^2/2\right)$$

b) Inverse-Gamma prior on τ_λ^2

$$p(\tau_\lambda^2) = IG(c,d)$$

$$(\tau_\lambda^2 \mid .) \sim IG\left(c + \frac{T-k-1}{2}, d + \sum_{t=k+2}^{T} (\lambda_t - \lambda_{t-1})^2/2\right)$$

Full conditional of $\boldsymbol{\tau_\lambda^2}$ for hierarchical $\boldsymbol{\lambda}$

a) Uniform prior on τ_λ

$$p(\tau_\lambda) \propto 1, \text{with } \tau_\lambda > 0$$

Equivalently:

$$p(\tau_\lambda^2) \propto (\tau_\lambda^2)^{-1/2} = IG(-0.5,0)$$

Conditional posterior density of τ_λ^2:

$$(\tau_\lambda^2 \mid .) \sim IG\left((T-k-1)/2, \sum_{t=k+2}^{T} (\lambda_t - \mu_\lambda)^2/2\right)$$

b) Inverse-Gamma prior on τ_λ^2

$$p(\tau_\lambda^2) = IG(c,d)$$

$$(\tau_\lambda^2 \mid .) \sim IG\left(c + (T-k)/2, d + \sum_{t=k+2}^{T} (\lambda_t - \mu_\lambda)^2/2\right)$$

Full conditional of τ_a^2

a) Uniform prior on τ_a

$p(\tau_a) \propto 1$, with $\tau_a > 0$

Equivalently:

$$p(\tau_a^2) \propto (\tau_a^2)^{-1/2} = IG(-0.5,0)$$

Conditional posterior density of τ_a^2:

$$(\tau_a^2 \mid .) \sim IG\left((T-k-2)/2, \sum_{t=2}^{T-k} (a_t - a_{t-1})^2/2\right)$$

b) Inverse-Gamma prior on τ_a^2

$$p(\tau_a^2) = IG(c,d)$$

$$(\tau_a^2 \mid .) \sim IG\left(c + (T-k-1)/2, d + \sum_{t=2}^{T-k} (a_t - a_{t-1})^2/2\right)$$

Full conditional of r_t, $t = k+1, \ldots, T$

Conditional distribution of r_t

$$
\begin{aligned}
&p(r_t \mid r_{t-k}, r_{t+k}, y_t^s, y_t^c, \theta) \\
&= p(r_t \mid r_{t-k}) p(r_{t+k} \mid r_t) p(y_t^s \mid r_t, \sigma_s^2) p(y_t^c \mid r_t, \sigma_c^2)
\end{aligned}
$$

where:

$c_t = \log(C_t) = r_t + \log(h_t) = r_t + \lambda_t - \log(1 + \exp(\lambda_t))$

$s_t = \log(S_t) = r_t + \log(1 - h_t) = r_t - \log(1 + \exp(\lambda_t))$

$S_t = (1 - h_t)R_t = (1 - h_t)\exp(r_t)$

Log conditional distribution of r_t used in the Metropolis step

For $t = k + 1, \ldots, T - k$

$$
\begin{aligned}
&\log p(r_t \mid r_{t-k}, r_{t+k}, \theta, y_t^s, y_t^c) \\
&\propto \\
&-(y_t^s - s_t))^2/(2\sigma_s^2) \\
&-(y_t^c - c_t)^2/(2\sigma_c^2) \\
&-(r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k}))^2/(2\tau_r^2) \\
&-(r_{t+k} - (a_t + s_t - \beta S_t))^2/(2\tau_r^2) \\
&= \quad -(r_t + \log(1 - h_t) - y_t^s)^2/(2\sigma_s^2) \\
&\qquad -(r_t + \log(h_t) - y_t^c)^2/(2\sigma_c^2) \\
&\qquad -(r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k}))^2/(2\tau_r^2) \\
&\qquad -(a_t + r_t + \log(1 - h_t) - \beta(1 - h_t)\exp(r_t) - r_{t+k})^2/(2\tau_r^2)
\end{aligned}
$$

Gradient of $\log p(r_t \mid -)$:

$$
\begin{aligned}
\frac{d\log p(r_t \mid -)}{dr_t} = &\frac{1}{\sigma_s^2}(y_t^s - s_t) + \frac{1}{\sigma_c^2}(y_t^c - c_t) - \frac{1}{\tau_r^2}\big(r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k})\big) \\
&- \frac{1}{\tau_r^2}\big(r_{t+k} - (a_t + s_t - \beta S_t)\big)(1 - \beta S_t)
\end{aligned}
$$

For $t = (T - k + 1), \dots, T$

$$\log(p(r_t \mid .)) \\ \propto \\ -(y_t^s - s_t)^2/(2\sigma_s^2) \\ -(y_t^c - c_t)^2/(2\sigma_c^2) \\ -(r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k}))^2/(2\tau_r^2)$$

Gradient of $\log p(r_t \mid -)$ for $t = (T - k + 1), \dots, T$

$$\frac{d\log p(r_t \mid -)}{dr_t} = \frac{1}{\sigma_s^2}(y_t^s - s_t) + \frac{1}{\sigma_c^2}(y_t^c - c_t) - \frac{1}{\tau_r^2}\big(r_t - (a_{t-k} + s_{t-k} - \beta S_{t-k})\big)$$

Metropolis step

A Metropolis step with a normal proposal distribution $q(r_t^* \mid r_t^{(i)}) = N(r_t^* \mid r_t^{(i)}, sd_t^2)$, $t = k + 1, \dots, T$, can be used to update each $r_t = \log(R_t)$, where $r_t^{(i)}$ is the value of r_t at the current iteration i, and the r_t^* is a candidate value for r_t drawn from the proposal distribution with a specified standard deviation sd_t. In a Metropolis step, the r_t^* is accepted as an update for r_t with probability $\min\Big(1, p(r_t^* \mid r_t^{(i)}) p(r_t^{(t)} \mid r_t^*)\Big)$; otherwise, it is rejected and the chain will remain in place $r_t^{(i+1)} = r_t^{(i)}$.

The performance of the Metropolis-Hastings algorithm can be expressed by the acceptance rate of the candidate draws in the Metropolis-Hastings steps. Theoretical and empirical results show that the acceptance rate in the range 20% ~ 50% (depending on number of parameters) provides optimal performance (Gelman et al. 1995). For multilevel models, Browne and Draper (2000) proposed an acceptance rate of 40% ~ 60% for univariate updating.

To increase the efficiency of the Metropolis algorithm, we adopt an adaptive tuning step similar to that of Browne and Draper (2000) to tune the sd_t before generating sample draws for inference. The goal of the tuning is to obtain a target acceptance rate around 50%. The adaptive step is stopped after a fixed number of iterations, after which the burn-in period (the pre-convergence period) and main monitoring run (the post-convergence period) is started.

Independent chain

In this algorithm, the density function $p(r_t \mid r_{t-k})$ obtained from the transition equation is used as the proposal distribution for the general Metropolis-Hastings algorithm. In this case, $q(x^* \mid x) = q(x^*)$ does not depend on the current value $x^{(i)}$, so the algorithm is called the independent chain. The r_t can be updated using the general Metropolis-Hastings algorithm with the proposal $p(r_t \mid r_{t-k})$,

$$
\begin{aligned}
&q(r_t^*) \\
&= p(r_t^{(i)} \mid r_{t-k}^{(i)}) \\
&= N(r_t^{(i)} \mid f(r_{t-k}^{(i)}), (\tau_r^2)^{(i)}) \\
&= \frac{1}{\sqrt{2\pi(\tau_r^2)^{(i)}}} \exp(-\frac{1}{2(\tau_r^2)^{(i)}} (r_t^{(i)} - f(r_{t-k}^{(i)}))^2)
\end{aligned}
$$

where $t = k + 1, \dots, T$, $r_t^{(i)}$, $r_{t-k}^{(i)}$, and $(\tau_r^2)^{(i)}$ are the current values of r_t, r_{t-k} and τ_r^2 at iteration i, r_t^* denotes a candidate value for r_t. To update each $r_t = \ln(R_t)$, a r_t^* is drawn from $q(r_t^*)$. Then r_t^* is accepted as an update for r_t with probability:

$$
\begin{aligned}
&\min\left(1, \left(\frac{p(r_t^* \mid .)}{q(r_t^*)}\right) / \frac{p(r_t^{(i)} \mid .)}{q(r_t)}\right) \\
&= \min\left(1, \left(\frac{p(r_t^* \mid r_{t-k}^{(i)}) p(r_{t+k}^{(i)} \mid r_t^*) p(y_t \mid r_t^*)}{p(r_t^* \mid r_{t-k}^{(i)})}\right) / \left(\frac{p(r_t^{(i)} \mid r_{t-k}^{(i)}) p(r_{t+k}^{(i)} \mid r_t^{(i)}) p(y_t \mid r_t^{(i)})}{p(r_t^{(i)} \mid r_{t-k}^{(i)})}\right)\right); \\
&= \min\left(1, \frac{p(r_{t+k}^{(i)} \mid r_t^*) p(y_t \mid r_t^*)}{p(r_{t+k}^{(i)} \mid r_t^{(i)}) p(y_t \mid r_t^{(i)})}\right)
\end{aligned}
$$

otherwise, it is rejected and $r_t^{(i+1)} = r_t^{(i)}$.

Full conditional of the initial states s_1 to s_k

Log conditional distribution of r_t used in the Metropolis step

$$
\begin{aligned}
&(s_t \mid .) \\
&= p(s_t)p(r_{t+k} \mid s_t)p(y_t^s \mid s_t) \\
&= N(s_t \mid m_s, V_s)N(r_{t+k} \mid (a_t + s_t - \beta S_t), \tau_r^2)N(y_t^s \mid s_t, \sigma_s^2)
\end{aligned}
$$

$$
\begin{aligned}
\log p(s_t \mid .) \propto & \\
& - (s_t - m_s)^2/(2V_s) \\
& - (r_{t+k} - (a_t + s_t - \beta S_t))^2/(2\tau_r^2) \\
& - (y_t^s - s_t)^2/(2\sigma_s^2)
\end{aligned}
$$

Gradient of $\log p(s_t \mid -)$:

$$
\frac{d\log p(s_t \mid -)}{ds_t} = -\frac{1}{V_s}(s_t - m_s) - \frac{1}{\tau_r^2}\big(r_{t+k} - (a_t + s_t - \beta S_t)\big)(1 - \beta S_t) + \frac{1}{\sigma_s^2}(y_t^s - s_t)
$$

Metropolis step for updating s_1 to s_k

A Metropolis step with a normal proposal distribution $q(s_t^* \mid s_t^{(i)}) = N(s_t^* \mid s_t^{(i)}, sd_t^2)$, t = k+1, ..., T, is used to update each $s_t = \ln(S_t)$, where $s_t^{(i)}$ is the value of s_t at current iteration i, and the s_1^* is a candidate value for s_t drawn from the proposal distribution with a specified standard deviation sd_t. In a Metropolis step, the s_t^* is accepted as an update for x_t with probability $\min\big(1, p(s_t^* \mid s_t^{(i)})p(s_t^{(i)} \mid s_t^*)\big)$; otherwise, it is rejected, and $s_t^{(i+1)} = s_t^{(i)}$.

The performance of the Metropolis-Hastings algorithm can be expressed by the acceptance rate of the candidate draws in the Metropolis-Hastings steps. Theoretical and empirical results show that the acceptance rate in the range 20% ~ 50% (depending on number of parameters) provides optimal performance (Gelman et al. 1995). For multilevel models, Browne and Draper (2000) proposed an acceptance rate of 40% ~ 60% for univariate updating.

To increase the efficiency of the Metropolis algorithm, we adopt an adaptive tuning step similar to that of Browne and Draper (2000) to tune the sd_i before generating sample draws for inference. The goal of the tuning is to obtain a target acceptance rate around 50%. The adaptive step is stopped after a fixed number of iterations, after which the burn-in period (the pre-convergence period) and main monitoring run (the post-convergence period) is started.

Full conditional of $\{\lambda\}_{t=k+1}^{T}$: independent priors (fixed effects model)

Conditional distribution of λ_t

$$
\begin{aligned}
&p(\lambda_t \mid m_\lambda, V_\lambda, \theta, y_t^s, y_t^c) \\
&= p(\lambda_t \mid m_\lambda, V_\lambda) \\
&\quad p(y_t^s \mid \lambda_t, \sigma_s^2) \\
&\quad p(y_t^c \mid \lambda_t, \sigma_c^2)
\end{aligned}
$$

where:

$$h_t = \exp(\lambda_t)/(1 + \exp(\lambda_t))$$

$$c_t = \log(C_t) = r_t + \log(h_t) = r_t + \lambda_t - \log(1 + \exp(\lambda_t))$$

$$s_t = \log(S_t) = r_t + \log(1 - h_t) = r_t - \log(1 + \exp(\lambda_t))$$

$$\frac{ds_t}{d\lambda_t} = -\frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} = -h_t$$

$$\frac{dc_t}{d\lambda_t} = 1 - \frac{\exp(\lambda_t)}{1 + \exp(\lambda_t)} = 1 - h_t$$

Log conditional distribution of λ_t used in the Metropolis step

$$
\begin{aligned}
&\log p(\lambda_t \mid m_\lambda, V_\lambda, \theta, y_t^s, y_t^c) \\
&\propto \\
&-(\lambda_t - m_\lambda)^2/(2V_\lambda) \\
&-(y_t^s - s_t)^2/(2\sigma_s^2) \\
&-(y_t^c - c_t)^2/(2\sigma_c^2)
\end{aligned}
$$

Gradient of $\log p(\lambda_t \mid -)$:

$$\frac{d\log p(\lambda_t \mid -)}{d\lambda_t} = -\frac{1}{V_\lambda}(\lambda_t - m_\lambda) - \frac{1}{\sigma_s^2}(y_t^s - s_t)(h_t) + \frac{1}{\sigma_c^2}(y_t^c - c_t)(1 - h_t)$$

Full conditional of $\lambda_{t=k+1}^{T}$: hierarchical prior

Conditional distribution of λ_t

$$
\begin{aligned}
&p(\lambda_t \mid \mu_\lambda, \tau_\lambda^2, \sigma_s^2, \sigma_c^2, y_t^s, y_t^c) \\
&= N(\lambda_t \mid \mu, \tau_\lambda^2) N(y_t^s \mid \lambda_t, \sigma_s^2) N(y_t^c \mid \lambda_t, \sigma_c^2)
\end{aligned}
$$

Log conditional distribution of λ_t used in the Metropolis step

$$
\begin{aligned}
&\log p(\lambda_t \mid \mu_\lambda, \tau_\lambda^2, \sigma_s^2, \sigma_c^2, y_t^s, y_t^c) \\
&\propto \\
&-(\lambda_t - \mu_\lambda)^2/(2\tau_\lambda^2) \\
&-(y_t^s - s_t)^2/(2\sigma_s^2) \\
&-(y_t^c - c_t)^2/(2\sigma_c^2)
\end{aligned}
$$

Gradient of $\log p(\lambda_t \mid -)$:

$$\frac{d\log p(\lambda_t \mid -)}{d\lambda_t} = -\frac{1}{\tau_\lambda^2}(\lambda_t - \mu_\lambda) - \frac{1}{\sigma_s^2}(y_t^s - s_t)(h_t) + \frac{1}{\sigma_c^2}(y_t^c - c_t)(1 - h_t)$$

Full conditional of μ_λ: hierarchical prior

$$
\begin{aligned}
&p(\mu_\lambda \mid .) \\
&\sim N\left(\sum_{t=k+1}^{T} \lambda_t / (T-K), \tau_\lambda^2/(T-K)\right)
\end{aligned}
$$

Full conditional of $\lambda_{t=k+1}^T$: RW prior

Conditional distribution of λ_t used in the Metropolis step

$$
\begin{aligned}
&p(\lambda_t \mid \lambda_{t-1}, \lambda_{t+1}, \theta, y_t^s, y_t^c) \\
&= p(\lambda_t \mid \lambda_{t-1}) \\
&\times p(\lambda_{t+1} \mid \lambda_t) \\
&\times p(y_t^s \mid \lambda_t, \sigma_s^2) p(y_t^c \mid \lambda_t, \sigma_c^2)
\end{aligned}
$$

For the initial λ at $t = k + 1$: $\mathrm{p}(\lambda_{(1)} \mid .)$

$$
\begin{aligned}
&\log p(\lambda_t \mid \lambda_{t+1}, \theta, y_t^s, y_t^c) \\
&\propto \\
&-(\lambda_t - m_\lambda)^2/(2V_\lambda) \\
&-(y_t^s - s_t)^2/(2\sigma_s^2) \\
&-(y_t^c - c_t)^2/(2\sigma_c^2) \\
&-(\lambda_{t+1} - \lambda_t)^2/(2\tau_\lambda^2)
\end{aligned}
$$

Gradient of $\log p(\lambda_t \mid -)$:

$$\frac{d\log p(\lambda_t \mid -)}{d\lambda_t} = -\frac{1}{V_\lambda}(\lambda_t - m_\lambda) - \frac{1}{\sigma_s^2}(y_t^s - s_t)(h_t) + \frac{1}{\sigma_c^2}(y_t^c - c_t)(1 - h_t) + \frac{1}{\tau_\lambda^2}(\lambda_{t+1} - \lambda_t)$$

For $t = k + 2, \ldots, T - 1$: $p(\lambda_t \mid .)$

$$\log p(\lambda_t \mid \lambda_{t-1}, \lambda_{t+1}, \theta, y_t^s, y_t^c)$$
$$\propto$$
$$-(y_t^s - s_t)^2/(2\sigma_s^2)$$
$$-(y_t^c - c_t)^2/(2\sigma_c^2)$$
$$-(\lambda_t - \lambda_{t-1})^2/(2\tau_\lambda^2)$$
$$-(\lambda_{t+1} - \lambda_t)^2/(2\tau_\lambda^2)$$

Gradient of $\log p(\lambda_t \mid -)$:

$$\frac{d\log p(\lambda_t \mid -)}{d\lambda_t}$$
$$= -\frac{1}{\sigma_s^2}(y_t^s - s_t)(h_t) + \frac{1}{\sigma_c^2}(y_t^c - c_t)(1 - h_t) - \frac{1}{\tau_\lambda^2}(\lambda_t - \lambda_{t-1}) + \frac{1}{\tau_\lambda^2}(\lambda_{t+1} - \lambda_t)$$

For $t = T$: $p(\lambda_T \mid .)$

$$\log p(\lambda_T \mid .)$$
$$\propto$$
$$-(y_T^s - s_T)^2/(2\sigma_s^2)$$
$$-(y_T^c - c_T)^2/(2\sigma_c^2)$$
$$-(\lambda_T - \lambda_{T-1})^2/(2\tau_\lambda^2)$$

Gradient of $\log p(\lambda_t \mid -)$:

$$\frac{d\log p(\lambda_T \mid -)}{d\lambda_T} = -\frac{1}{\sigma_s^2}(y_T^s - s_T)(h_T) + \frac{1}{\sigma_c^2}(y_T^c - c_T)(1 - h_T) - \frac{1}{\tau_\lambda^2}(\lambda_T - \lambda_{T-1})$$

Full conditional of $a_t, t = 1, \ldots, T-k$

Conditional distribution of $\boldsymbol{a_t, t = 1, \ldots, T-k}$

$$p(a_t \mid a_{t-1}, a_{t+1}, \theta, r_t)$$
$$= N(a_t \mid a_{t-1}) N(a_{t+1} \mid a_t) p(r_t \mid r_{t-k})$$

Full conditional of $\boldsymbol{a_1}$

For $t = 1$: $p(a_{(1)} \mid .)$:

$$p(a_{(1)} \mid m_a, V_a, a_2, \beta, r_3)$$
$$= N(a_1 \mid m_a, V_a) N(a_2 \mid a_1, \tau_a^2) N(r_3 \mid a_1 + s_1 - \beta S_1)$$

$\log p(a_{(1)} \mid m_a, V_a, a_2, \beta, r_3)$
$\propto$
$-(a_1 - m_a)^2/(2V_a)$
$-(a_2 - a_1)^2/(2\tau_a^2)$
$-(r_3 - (a_1 + s_1 - \beta S_1))^2/(2\tau_a^2)$
$=$
$-(a_1^2 - 2m_a a_1 + m_a^2)/(2V_a)$
$-(a_1^2 - 2a_2 a_1 + a_2^2)/(2\tau_a^2)$
$-(a_1^2 - 2(r_3 - s_1 + \beta S_1)a_1 + .)/(2\tau_r^2)$
$=$
$-\left\{a_1^2 (\frac{1}{V_a} + \frac{1}{\tau_a^2} + \frac{1}{\tau_r^2})/2 - 2(\frac{m_a}{V_a} + \frac{a_2}{\tau_a^2} + (r_3 - s_1 + \beta S_1)/\tau_r^2)\frac{a_1}{2} + .\right\}$
$=$
$-\frac{\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2}{2V_a\tau_a^2\tau_r^2}\left\{a_1^2 - 2\frac{m_a\tau_r^2\tau_a^2 + a_2 V_a\tau_r^2 + (r_3 - s_1 + \beta S_1)V_a\tau_a^2}{(\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2)} a_1 + .\right\}$
$= \frac{1}{2\frac{V_a\tau_a^2\tau_r^2}{\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2}}\left\{a_1^2 - 2\frac{m_a\tau_r^2\tau_a^2 + a_2 V_a\tau_r^2 + (r_3 - s_1 + \beta S_1)V_a\tau_a^2}{(\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2)} a_1 + .\right\}$
$= N(a_1 \mid m_a\tau_r^2\tau_a^2 + a_2 V_a\tau_r^2 + (r_3 - s_1 + S_1)V_a\tau_a^2/(\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2), V_a\tau_a^2\tau_r^2/\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2)$

<u>Full conditional of a_1</u>:

$p(a_1 \mid .) \quad = N(\mu_{a_1}, \tau_{a_1})$

$\mu_{a_1} = m_a\tau_r^2\tau_a^2 + a_2 V_a\tau_r^2 + (r_3 - s_1 + S_1)V_a\tau_a^2/(\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2)$

$\tau_{a_1} = V_a\tau_a^2\tau_r^2/\tau_r^2\tau_a^2 + V_a\tau_r^2 + V_a\tau_a^2$

Full conditional of $\boldsymbol{a_t}$:

For $t = 2, t - k$ t = 2, …, T – k:

$\log p(a_t \mid a_{t-1}, a_{t+1}, \theta, r_t)$
$\propto$
$-(r_{t+k} - (a_t + s_t - \beta S_t))^2/(2\tau_r^2)$
$-(a_t - a_{t-1})^2/(2\tau_a^2)$
$-(a_{t+1} - a_t)^2/(2\tau_a^2)$
$=$
$-(a_t^2 - 2(r_{t+k} - s_t + \beta S_t)a_t + .)/(2\tau_r^2)$
$-(a_t^2 - 2a_{t-1}a_t + a^{2t-1})^2/(2\tau_a^2)$
$-(a_t^2 - 2a_{t+1}a_t + a^{2t+1})^2/(2\tau_a^2)$
$=$

$$-\left\{a_t^2 \frac{\tau_a^2 + 2\tau_r^2}{2\tau_r^2\tau_a^2} - 2\left(\frac{\tau_a^2(r_{t+k} - s_t + S_t) + \tau_r^2(a_{t-1} + a_{t+1})}{2\tau_a^2\tau_r^2}\right)a_t + .\right\}$$

$=$

$$\frac{1}{2\frac{\tau_r^2\tau_a^2}{\tau_a^2 + 2\tau_r^2}}\left(a_t^2 - 2\left(\frac{\tau_a^2(r_{t+k} - s_t + S_t) + \tau_r^2(a_{t-1} + a_{t+1})}{2\tau_a^2\tau_r^2}\right)\frac{2\tau_r^2\tau_a^2}{\tau_a^2 + 2\tau_r^2}a_t + .\right)$$

$$= \frac{1}{2V}\left(a_t^2 - 2\frac{\tau_a^2(r_{t+k} - s_t + \beta S_t) + \tau_r^2(a_{t-1} + a_{t+1})}{\tau_a^2 + 2\tau_r^2}a_t + .\right)$$

$$= N(a_t \mid \frac{\tau_a^2(r_{t+k} - s_t + \beta S_t) + \tau_r^2(a_{t-1} + a_{t+1})}{\tau_a^2 + 2\tau_r^2}, \frac{\tau_r^2\tau_a^2}{\tau_a^2 + 2\tau_r^2})$$

<u>Full conditional of a_t:</u>

$p(a_t \mid .) \quad = N(\widehat{m}_{a_t}, \widehat{V}_{a_t})$

$\widehat{m}_{a_t} = \tau_a^2(r_{t+k} - s_t + \beta S_t) + \tau_r^2(a_{t-1} + a_{t+1})/(\tau_a^2 + 2\tau_r^2)$

$\widehat{V}_{a_t} = \tau_r^2\tau_a^2/(\tau_a^2 + 2\tau_r^2)$

Full conditional of $\boldsymbol{\beta}$:

Prior $(\beta) \sim N(\beta \mid m_\beta, V_\beta)$

$(\beta \mid .) \propto N(\beta \mid m_\beta, V_\beta) N(\hat{m}_\beta, \hat{V}_\beta)$

$$\hat{m}_\beta = \sum_{t=k+1}^{T} S_{t-k}(a_{t-k} + s_{t-k} - r_t) / \sum_{t=k+1}^{T} S_{t-k}^2$$

$$\hat{V}_\beta = \frac{\tau_r^2}{\sum_{t=k+1}^{T} S_{t-k}^2}$$

<u>Full conditional of β:</u>

$p(\beta \mid .) \propto$

$$N\left(\frac{m_\beta \times \hat{V}_\beta + \hat{m}_\beta \times V_\beta}{V_\beta \times \hat{V}_\beta}, \frac{V_\beta \times \hat{V}_\beta}{V_\beta + \hat{V}_\beta}\right)$$