Supplementary 1: Sequential Monte Carlo methods for data assimilation problems

Kwaku Peprah Adjei[1,2], Robert B. O'Hara[1,2], Nick Isaac[3], Daina Bowler[3], Rob Cooke[3]

[1]Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

[2]Centre of Biodiversity Dynamics, Norwegian University of Science and Technology, Trondheim, Norway

[3]Center of Ecology and Hydrology, Wallingford, UK

1 Metropolis Hastings Acceptance ratio (MHAR) for proposed model

1.1 MHAR for particle MCMC defined by Andrieu *et al.* (2010)

Particle Markov chain Monte Carlo (pMCMC hereafter) (Andrieu *et al.* 2010) proposes a value for the top-level parameter ($\theta^\star$) from a proposal distribution ($\pi(\theta^\star|\theta)$). A seqiential monte carlo (SMC) method is fitted with the proposed value $\theta^\star$ and this proposed value is accepted or rejected based on a Metropolis Hasting acceptance ratio (MHAR) defined by Andrieu *et al.* (2010):

$$MHAR = \frac{p_\theta(y_{1:t})p(\theta^\star)\pi(\theta|\theta^\star)}{p_{\theta^\star}(y_{1:t})p(\theta)\pi(\theta^\star|\theta)}; \tag{1}$$

where $p_\theta(y_{1:t})$ is the marginal distribution of the observed data given the parameter θ, $p(\theta^\star)$ and $p(\theta)$ is the prior distribution of $\theta^\star$ and θ respectively and $\pi(.|.)$ is the proposal distribution for the parameters θ.

We use the same definitions and descriptions of state space models (SSMs) described in the main paper (equation 1 and Table 1).

1.1.1 Importance weights for proposed model $w^i_{t|\theta}$

As we defined in equation (3) of the main paper (which is repeated here for emphasis in the formulation of the MHAR), the importance weights for the proposed model was defined as:

$$\begin{cases} w^{(i)}_{t|\theta} = 1 & \text{for } 1 \leq t \leq t_r \\ w_{t|\theta} \text{ defined by equations (2) - (4) in main paper} & \text{for } t_{r+1} \leq t \leq T \end{cases} \tag{2}$$

where t_r is the last time step we have fitted the reduced SSM.

1.1.2 Marginal likelihood $p_\theta(y_{1:T})$

Sequential Monte Carlo approaches provide estimates of the marginal likelihood of the observed data. From equation (9) in Andrieu *et al.* (2010), the estimate of the marginal likelihood $p_\theta(y_{1:T})$ is given by:

$$
\begin{aligned}
\widehat{p}_\theta(y_{1:t}) &= \widehat{p}_\theta(y_1)\prod_{t=2}^{T}\widehat{p}_\theta(y_t|y_{t-1}) \\
&= \widehat{p}_\theta(y_1)\prod_{t=2}^{T}\frac{1}{M}\sum_{i=1}^{M} w^i_{t|\theta}; \\
\text{with}\quad \widehat{p}_\theta(y_t|y_{t-1}) &= \frac{1}{M}\sum_{i=1}^{M} w^i_{t|\theta}
\end{aligned}
\tag{3}
$$

$w^i_{t|\theta}$ is the i^{th} importance weight at time t given the value of θ.

For our proposed model, the estimated marginal likelihood in equation (3) is modified with our proposed model importance weights in equation (2) and obtain the an updated marginal likelihood $\widetilde{p}_\theta(y_{1:t})$ as:

$$
\begin{aligned}
\widetilde{p}_\theta(y_{1:t}) &= \widetilde{p}_\theta(y_1)\frac{1}{M}\left[\prod_{t=2}^{t_r}\sum_{i=1}^{M} w^i_{t|\theta}\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta}\right]; \\
&= \widetilde{p}_\theta(y_1)\frac{1}{M}\left[M(t_r-1)\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta}\right]; \quad \text{since } w^i_{t|\theta}=1 \quad \forall t\le t_r \\
&= (t_r-1)\widetilde{p}_\theta(y_1)\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta}
\end{aligned}
\tag{4}
$$

1.2 Updated MHAR for our proposed framework

Substituting equation (4) instead of (3) into the MHAR proposed by Andrieu *et al.* (2010) defined by equation (1), we obtain the updated MHAR ($MHAR_{upd}$):

$$
\begin{aligned}
MHAR_{upd} &= \frac{(t_r-1)\widetilde{p}_\theta(y_1)\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta}p(\theta^\star)\pi(\theta|\theta^\star)}{(t_r-1)\widetilde{p}_\theta(y_1)\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta^\star}p(\theta)\pi(\theta^\star|\theta)} \\
&= \frac{\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta}p(\theta^\star)\pi(\theta|\theta^\star)}{\prod_{t=t_r+1}^{T}\sum_{i=1}^{M} w^i_{t|\theta^\star}p(\theta)\pi(\theta^\star|\theta)};
\end{aligned}
\tag{5}
$$

where $w^i_{t|\theta}$ is the i^{th} importance weight at time t given the value of θ, $p(\theta^\star)$ and $p(\theta)$ is the prior distribution of $\theta^\star$ and θ respectively and $\pi(.|.)$ is the proposal distribution for the parameters θ.

2 References

Andrieu, C., Doucet, A. & Holenstein, R. (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **72**, 269–342.