Research Article

Statistics
in Medicine

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000

Supplementary Material for 'Time Series Analysis of fMRI Data: Spatial Modelling and Bayesian Computation'

Ming Teng[a], Timothy D. Johnson[a], Farouk S. Nathoo[b]*

[a] *University of Michigan, Department of Biostatistics, 1415 Washington Heights, Ann Arbor, MI 48109*
[b] *University of Victoria, Department of Mathematics and Statistics, Victoria, BC, Canada, V8W 3P4*
* *Correspondence to: University of Victoria, Department of Mathematics and Statistics, Victoria, BC, Canada, V8W 3P4. E-mail: nathoo@uvic.edu*

Appendices

Appendix A: Re-expression of the log-likelihood

By elaborating the vector multiplication in Equation (3) from the paper we have

$$\begin{aligned} l_n &= -\frac{\lambda_n}{2}\sum_{t=P+1}^{T}\left[y_{tn} - \sum_k x_{tk}w_{kn} - \sum_{p=1}^{P}(y_{t-p,n} - \sum_k x_{t-p,k}w_{kn})a_{pn}\right]^2 \\ &\quad + \frac{T-P}{2}\log\lambda_n + const \end{aligned} \tag{1}$$

Let $\mathbf{a}_n^* = (-1, \mathbf{a}_n)$, so $a_{pn}^* = a_{pn}$ if $p \geq 1$ and $a_{pn}^* = -1$ if $p = 0$, then equation (1) can be written as

$$\begin{aligned} l_n &= -\frac{\lambda_n}{2}\sum_{t=P+1}^{T}\left[\sum_{p=0}^{P} y_{t-p,n}a_{pn}^* - \sum_{p=0}^{P}\sum_k x_{t-p,k}w_{kn}a_{pn}^*\right]^2 + \frac{T-P}{2}\log\lambda_n + const \\ &= -\frac{\lambda_n}{2}\sum_{t=P+1}^{T}\Big(\sum_{p_1=0}^{P}\sum_{p_2=0}^{P} y_{t-p_1,n}y_{t-p_2,n}a_{p_1n}^*a_{p_2n}^* - 2\sum_{p_1=0}^{P}\sum_{p_2=0}^{P}\sum_{k=1}^{K} y_{t-p_1,n}x_{t-p_2,k} \\ &\quad w_{kn}a_{p_1n}a_{p_2n} + \sum_{p_1=1}^{P}\sum_{p_2=1}^{P}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K} x_{t-p_1,k_1}x_{t-p_2,k_2}w_{k_1n}w_{k_2n}a_{p_1n}a_{p_2n}\Big) \\ &\quad + \frac{T-P}{2}\log\lambda_n + const \\ &= -\frac{\lambda_n}{2}\Big(\sum_{p_1=0}^{P}\sum_{p_2=0}^{P} yy_{p_1p_2n}a_{p_1n}a_{p_2n} - 2\sum_{p_1=0}^{P}\sum_{p_2=0}^{P}\sum_{k=1}^{K} yx_{p_1np_2k}w_{kn}a_{p_1n}a_{p_2n} \\ &\quad + \sum_{p_1=0}^{P}\sum_{p_2=0}^{P}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K} xx_{p_1k_1p_2k_2}w_{k_1n}w_{k_2n}a_{p_1n}a_{p_2n}\Big) + \frac{T-P}{2}\log\lambda_n + const \end{aligned} \tag{2}$$

where

$$\begin{aligned} yy_{p_1p_2n} &= \sum_{t=P+1}^{T} y_{t-p_1,n}y_{t-p_2,n}, \\ yx_{p_1np_2k} &= \sum_{t=P+1}^{T} y_{t-p_1,n}x_{t-p_2,k}, \\ xx_{p_1k_1p_2k_2} &= \sum_{t=P+1}^{T} x_{t-p_1,k_1}x_{t-p_2,k_2}. \end{aligned}$$

In this way, the sum across t can be pre-computed instead of computing at every iteration in the algorithm.

Define $\mathbf{F}$ to be a $P \times P$ matrix with (p_1, p_2) entry $f_{p_1p_2} = yy_{p_1p_2n} - 2\sum_{k=1}^{K} yx_{p_1np_2k}w_{kn} + \sum_{k_1=1}^{K}\sum_{k_2=1}^{K} xx_{p_1k_1p_2k_2}w_{k_1n}w_{k_2n}$. Then the derivation above is just

$$l_n = -\frac{\lambda_n}{2}\mathbf{a}_n^{*T}\mathbf{F}\mathbf{a}_n^* + \frac{T-P}{2}\log\lambda_n + const. \tag{3}$$

which is Equation (19).

Appendix B: Derivation of the gradients

Based on Equations (13) and (31), the gradients are derived as follows:

$$\nabla w_{kn} \log p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{X}) = \lambda_n \Big(\sum_{p_1=0}^{P} \sum_{p_2=0}^{P} yx_{p_1 n p_2 k} a_{p_1 n} a_{p_2 n} - \sum_{p_1=0}^{P} \sum_{p_2=0}^{P} \sum_{k_2=1}^{K}$$
$$xx_{p_1 k p_2 k_2} w_{k_2 n} a_{p_1 n} a_{p_2 n} \Big) - \alpha_k (\mathbf{S}^T \mathbf{S})_n \mathbf{w}_k^T$$
$$= \lambda_n \mathbf{a}_n^* \mathbf{G} \mathbf{a}_n^{*T} - \alpha_k (\mathbf{S}^T \mathbf{S})_n \mathbf{w}_k^T \tag{4}$$

where $\mathbf{G}$ is a $P \times P$ matrix with (p_1, p_2) entry $g_{p_1 p_2} = yx_{p_1 n p_2 k} - \sum_{k_2=1}^{K} xx_{p_1 k p_2 k_2} w_{k_2 n}$.

$$\nabla a_{pn} \log p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{X}) = \lambda_n \Big(- \sum_{p_2=0}^{P} yy_{p p_2 n} a_{p_2 n} + \sum_{p_2=0}^{P} \sum_{k=1}^{K} yx_{p n p_2 k} w_{kn} a_{p_2 n}$$
$$+ \sum_{p_1=0}^{P} \sum_{k=1}^{K} yx_{p_1 n p k} w_{kn} a_{p_1 n} - \sum_{p_2=0}^{P} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} xx_{p k_1 p_2 k_2} w_{k_1 n} w_{k_2 n} a_{p_2 n} \Big) - \beta_p (\mathbf{D}^T \mathbf{D})_n \mathbf{a}_p^T$$
$$= \lambda_n \mathbf{f}_p \mathbf{a}_n^* - \beta_p (\mathbf{D}^T \mathbf{D})_n \mathbf{a}_p^T \tag{5}$$

where $\mathbf{f}_p$ is just the p^{th} row of $\mathbf{F}$.

Based on Equations (13) and (19), it is easy to show that

$$\nabla \alpha_k \log p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{X}) = -\frac{1}{2} \mathbf{w}_k (\mathbf{S}^T \mathbf{S}) \mathbf{w}_k^T + (\frac{N}{2} + q_1 - 1)/\alpha_k - \frac{1}{q_2} \tag{6}$$
$$\nabla \beta_p \log p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{X}) = -\frac{1}{2} \mathbf{a}_p (\mathbf{D}^T \mathbf{D}) \mathbf{a}_p^T + (\frac{N}{2} + r_1 - 1)/\beta_p - \frac{1}{r_2} \tag{7}$$
$$\nabla \lambda_n \log p(\boldsymbol{\theta} \mid \mathbf{Y}, \mathbf{X}) = -\frac{1}{2} \mathbf{a}_n^{*T} \mathbf{F} \mathbf{a}_n^* + \frac{(T-P)/2 + u_1 - 1}{\lambda_n} - \frac{1}{u_2} \tag{8}$$

1. Supplementary Figures

1.1. Simulation One

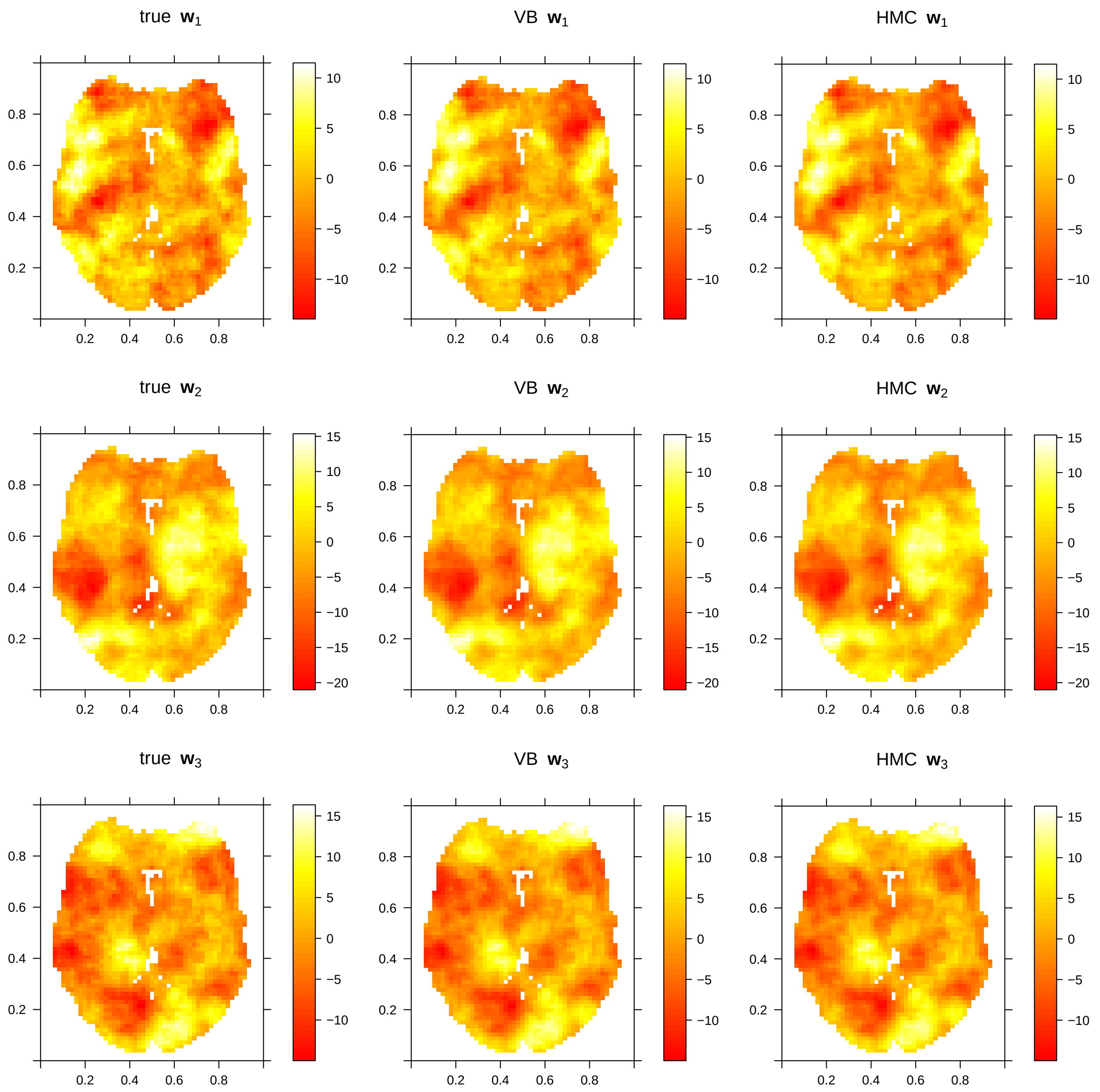

Figure 1. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$ from HMC and VB for Simulation One. The estimates are compared with true image in each row.

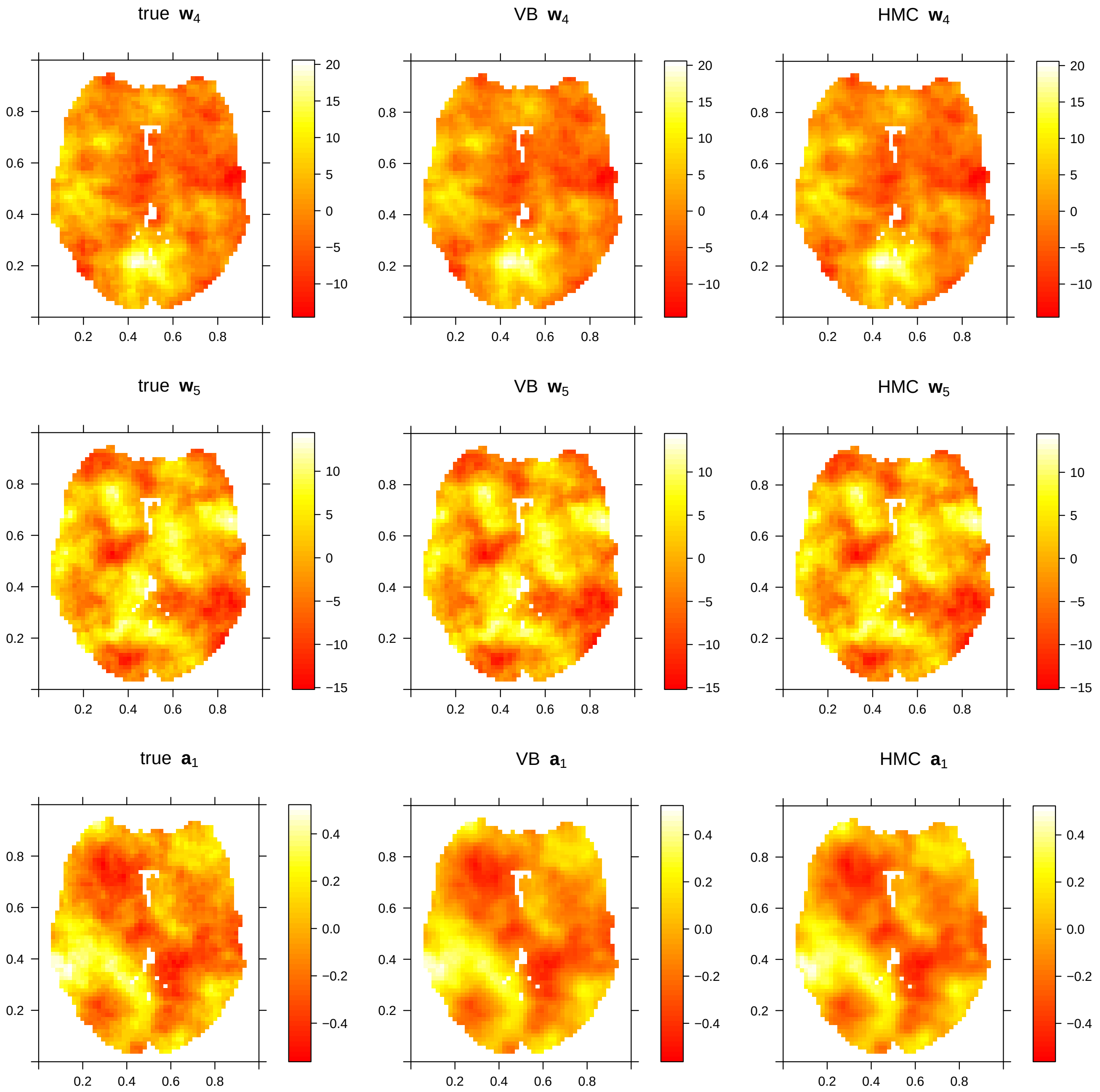

Figure 2. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_4$, $\mathbf{w}_5$, $\mathbf{a}_1$ from HMC and VB for Simulation One. The estimates are compared with true image in each row.

1.2. Simulation Two

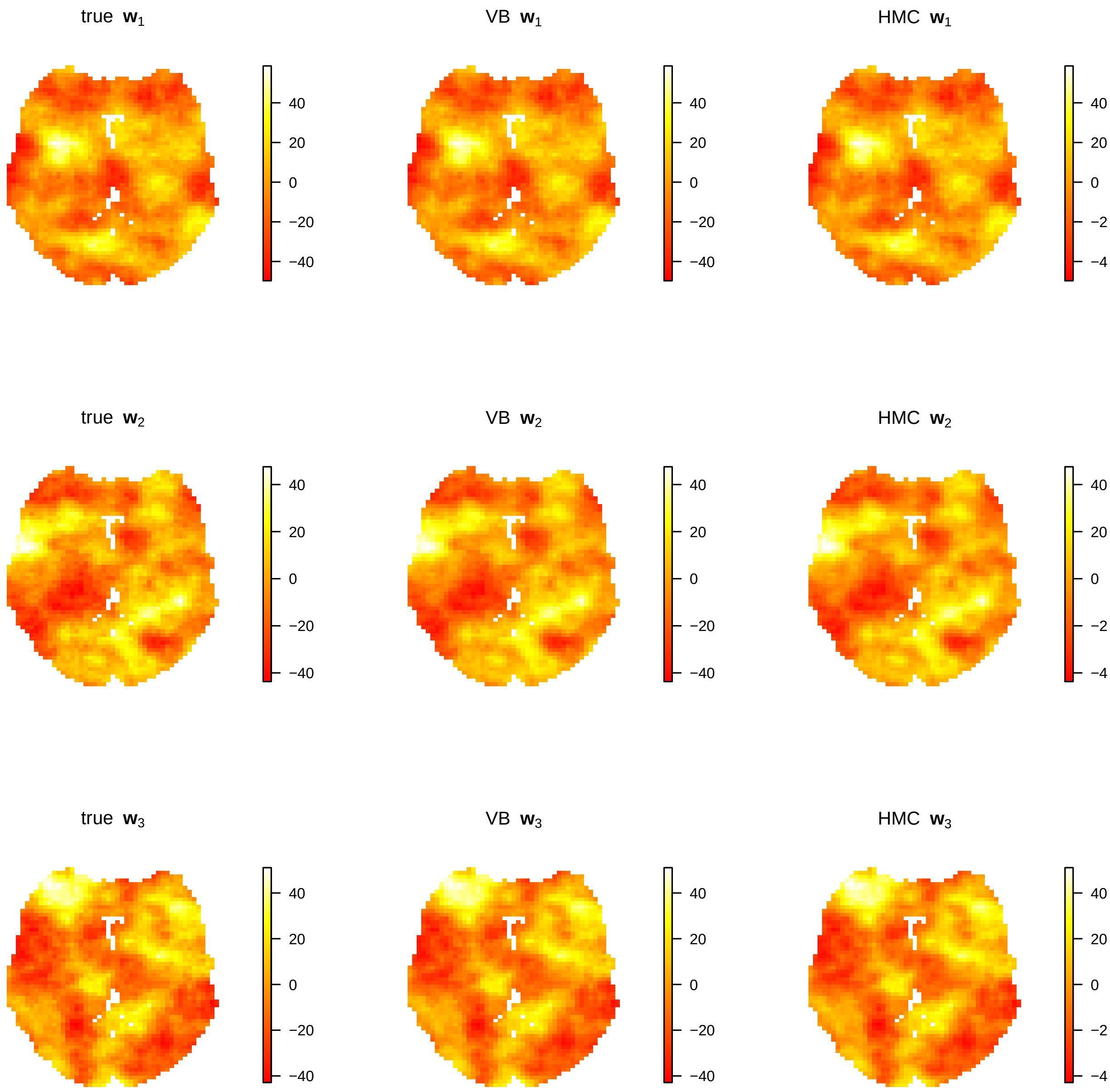

Figure 3. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$ from HMC and VB for Simulation Two. The estimates are compared with true image in each row.

1.3. Real applicatioin

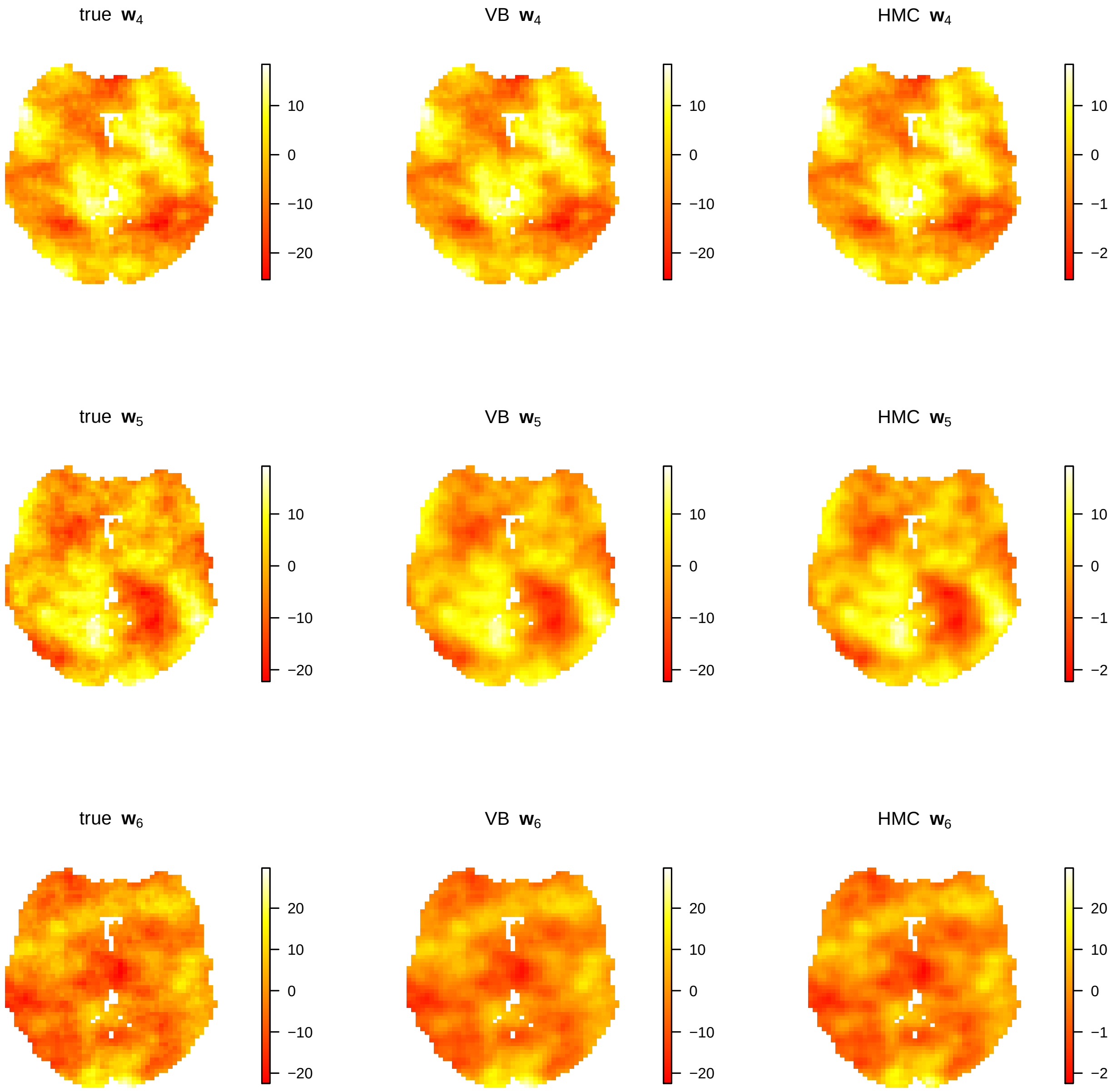

Figure 4. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_4$, $\mathbf{w}_5$, $\mathbf{w}_6$ from HMC and VB for Simulation Two. The estimates are compared with true image in each row.

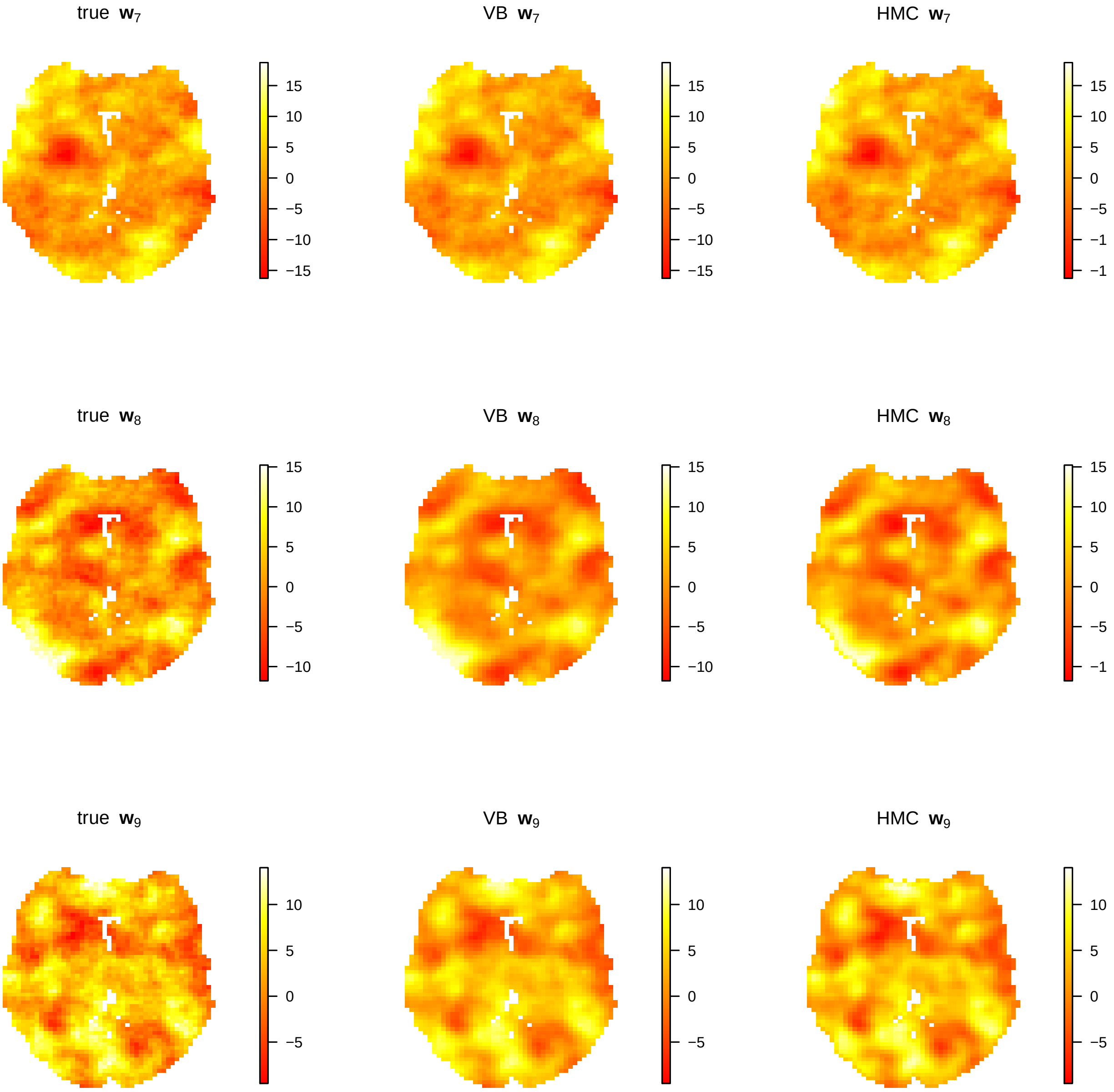

Figure 5. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_7$, $\mathbf{w}_8$, $\mathbf{w}_9$ from HMC and VB for Simulation Two. The estimates are compared with true image in each row.

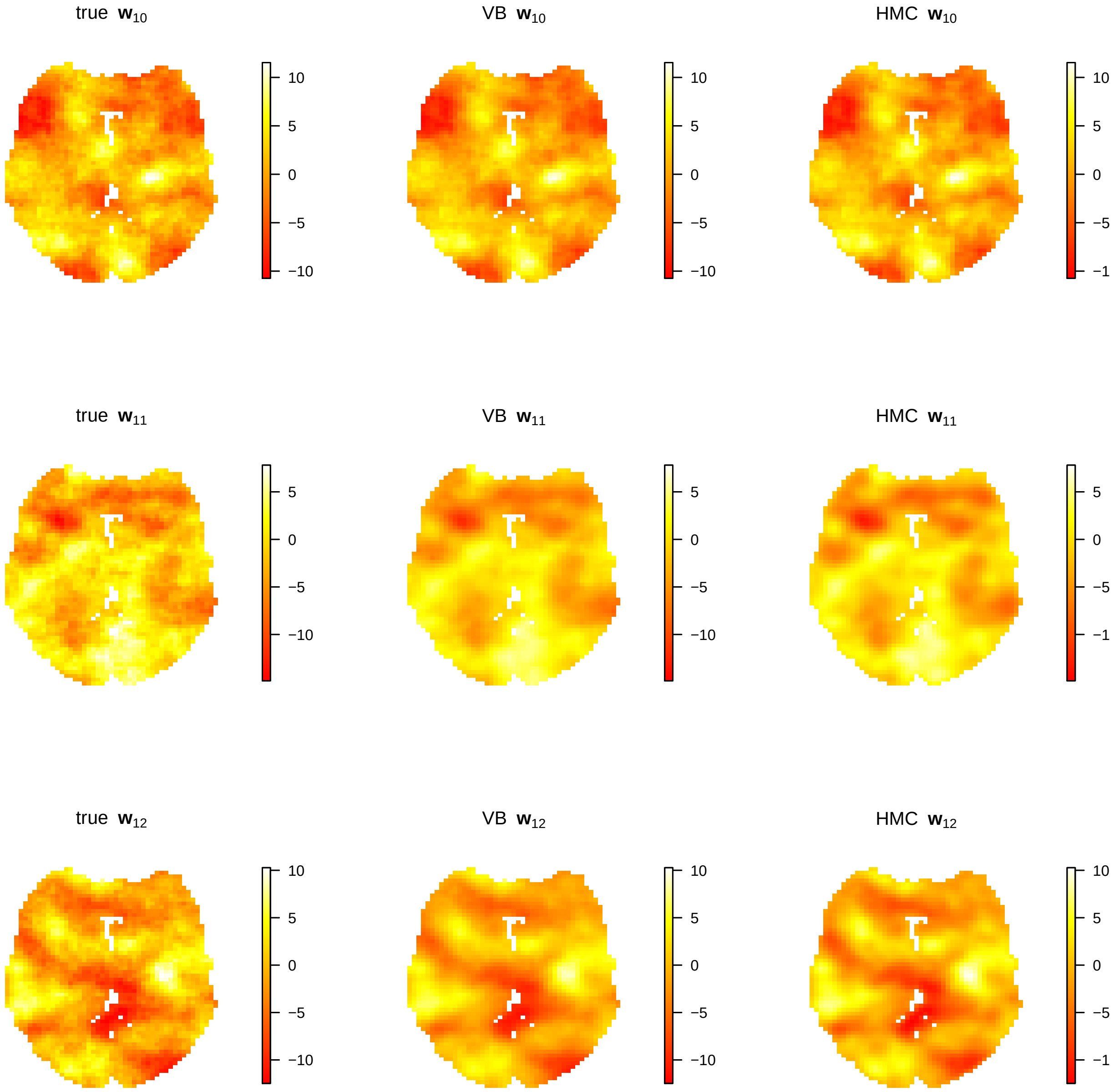

Figure 6. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_{10}, \mathbf{w}_{11}, \mathbf{w}_{12}$ from HMC and VB for Simulation Two. The estimates are compared with true image in each row.

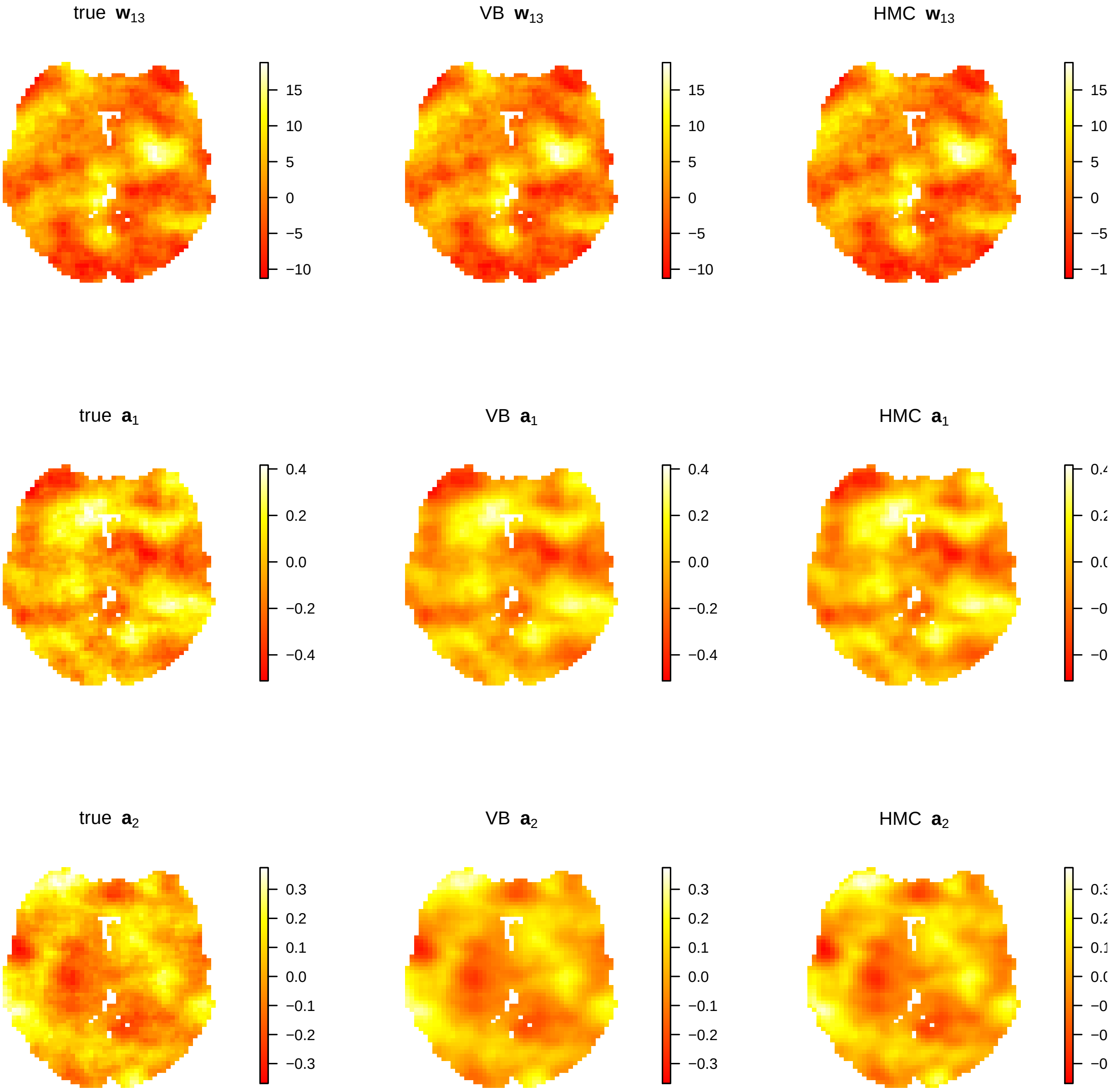

Figure 7. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{w}_{13}$, $\mathbf{a}_1$, $\mathbf{a}_2$ from HMC and VB for Simulation Two. The estimates are compared with true image in each row.

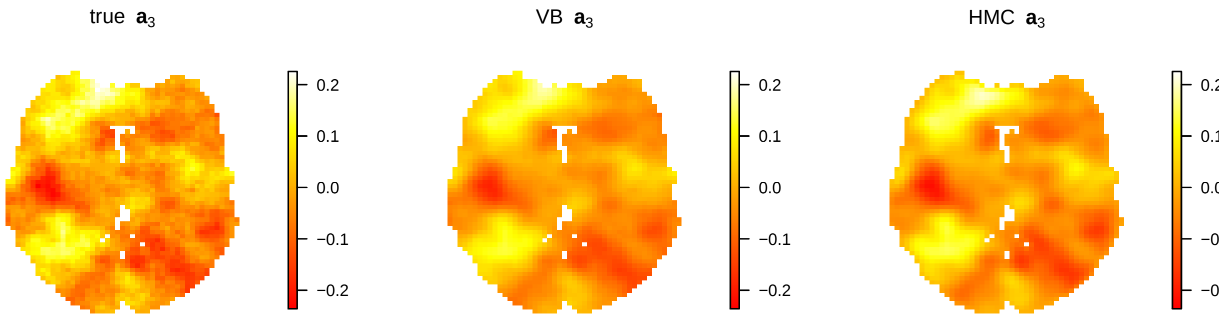

Figure 8. Image of average (over simulation replicates) posterior mean estimate of $\mathbf{a}_3$ from HMC and VB for Simulation Two. The estimates are compared with true image.

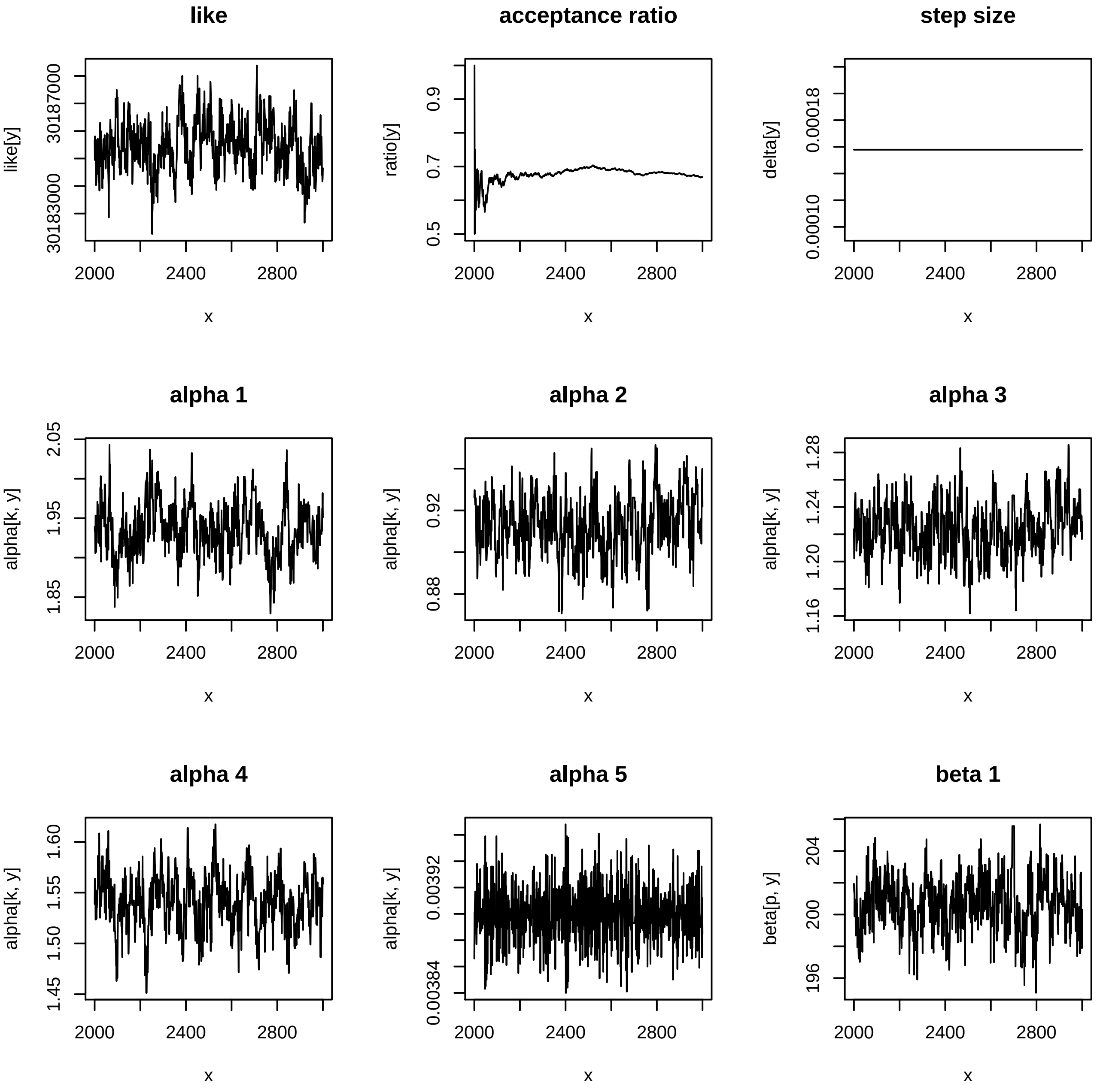

Figure 9. Traceplot for the parameters from HMC. The chain runs for 3000 iterations, with first 2000 as burn-in and thrown away. The three figures on top row (from left to right) are likelihood, acceptance ratio of Metropolis-Hastings step, and leapfrog step size δ respectively. The rest shows the trace plots from $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

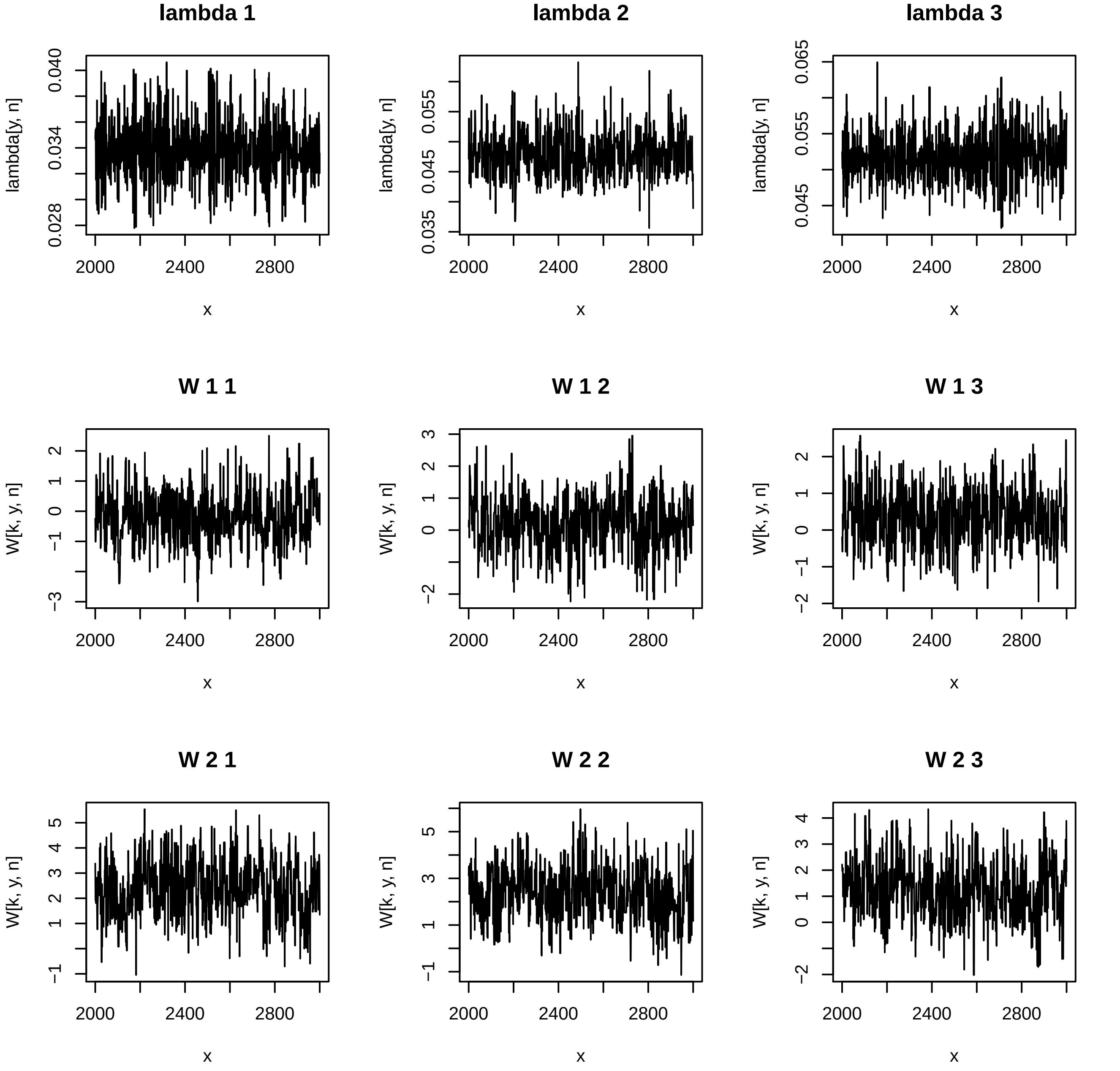

Figure 10. Traceplot for the parameters from HMC. The chain runs for 3000 iterations, with first 2000 as burn-in and thrown away. The top row represents the trace plots for $\lambda_1, \lambda_2, \lambda_3$. The second and third row shows trace plots from w_{11}, w_{12}, w_{13} and w_{21}, w_{22}, w_{23}. We just show the trace plots from first three voxels out of 56527 voxels due to a limited space.

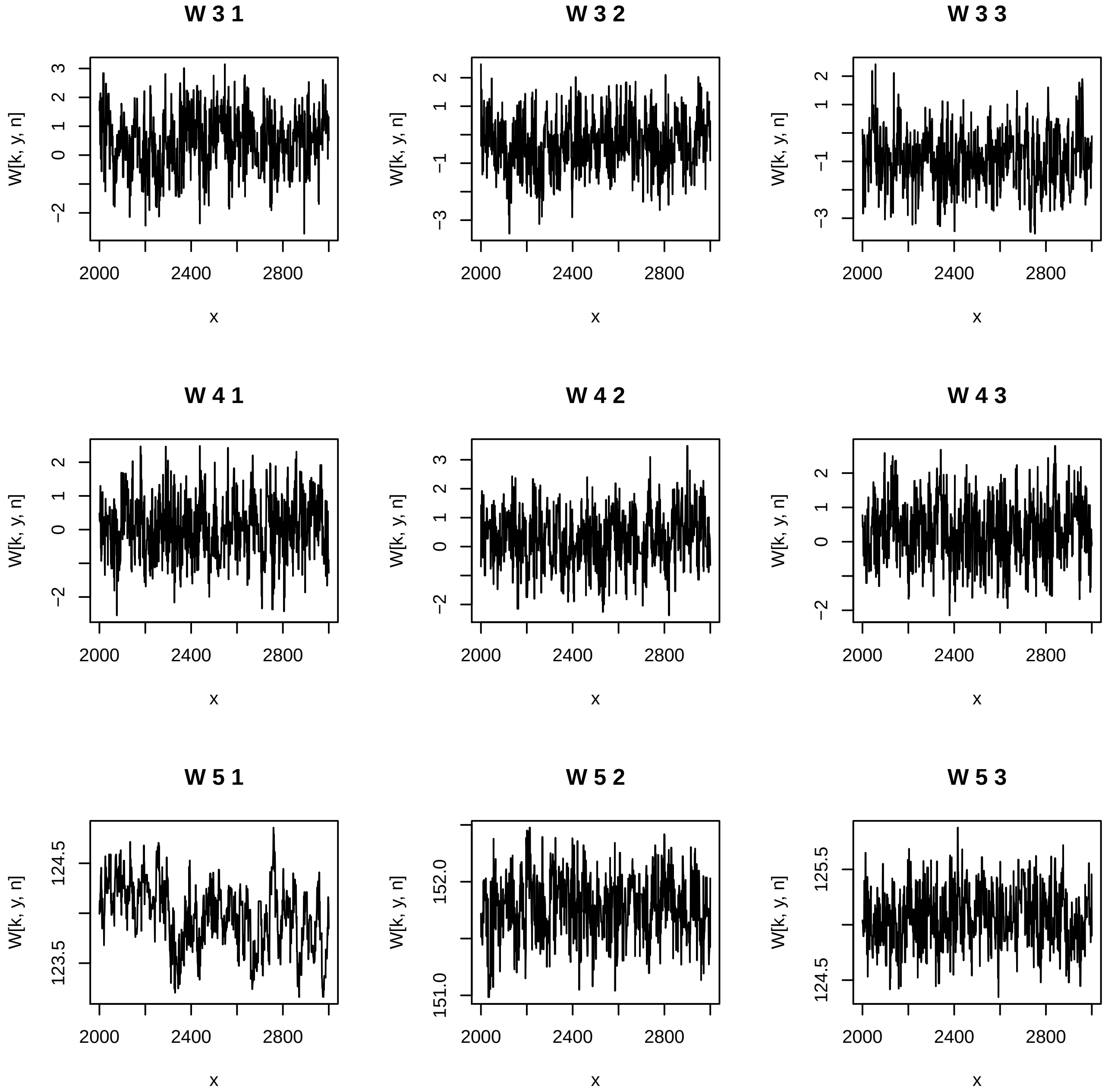

Figure 11. Traceplot for the parameters ($\mathbf{w}_3$ to $\mathbf{w}_5$) from HMC. The chain runs for 3000 iterations, with first 2000 as burn-in and thrown away.

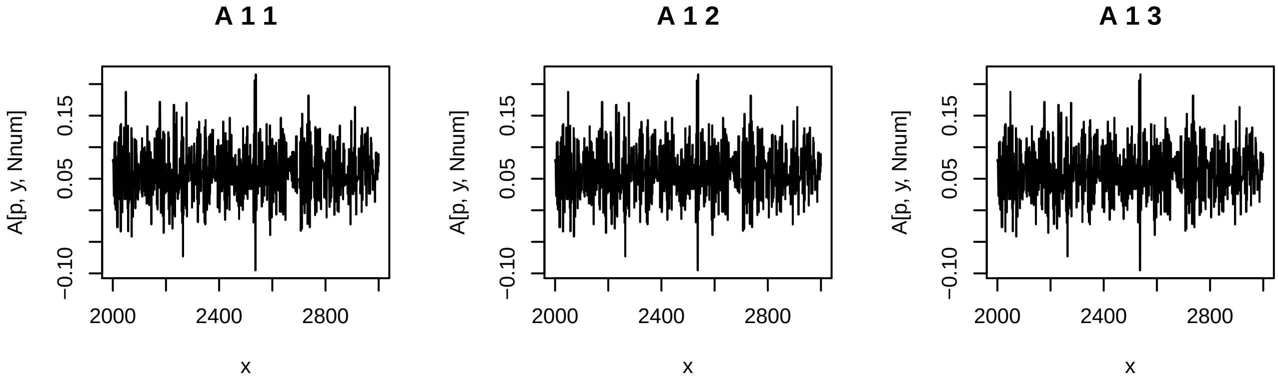

Figure 12. Traceplot for the auto-regressive coefficient $\mathbf{a}_1$ from HMC. The chain runs for 3000 iterations, with first 2000 as burn-in and thrown away.

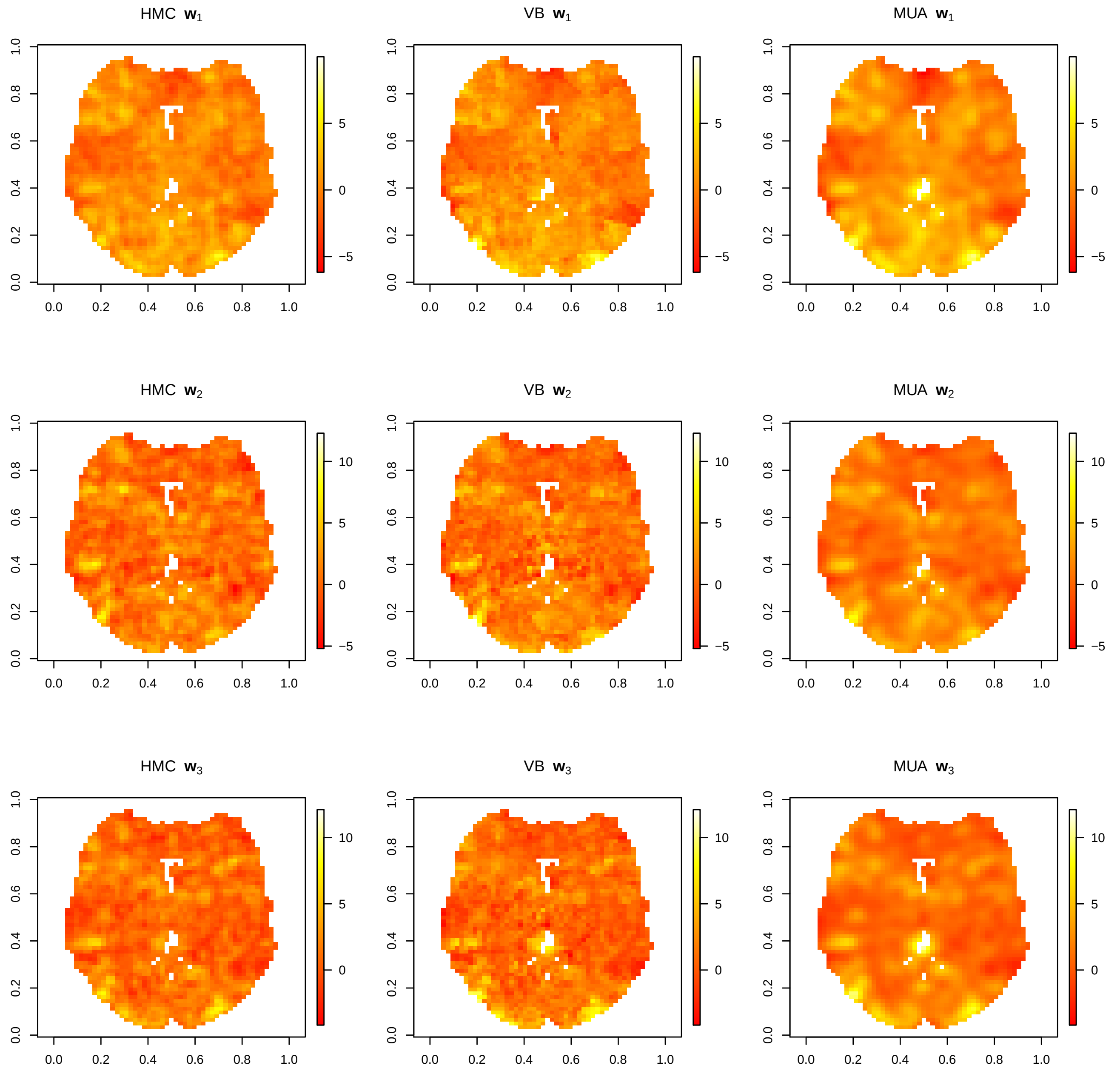

Figure 13. Image of posterior mean estimate of $\mathbf{w}_1 - \mathbf{w}_3$ from HMC, VB and MUA. These are the estimates from 26^{th} slices on the z-axis. We only provide this slice due to a limited space. The result is similar in other slices.

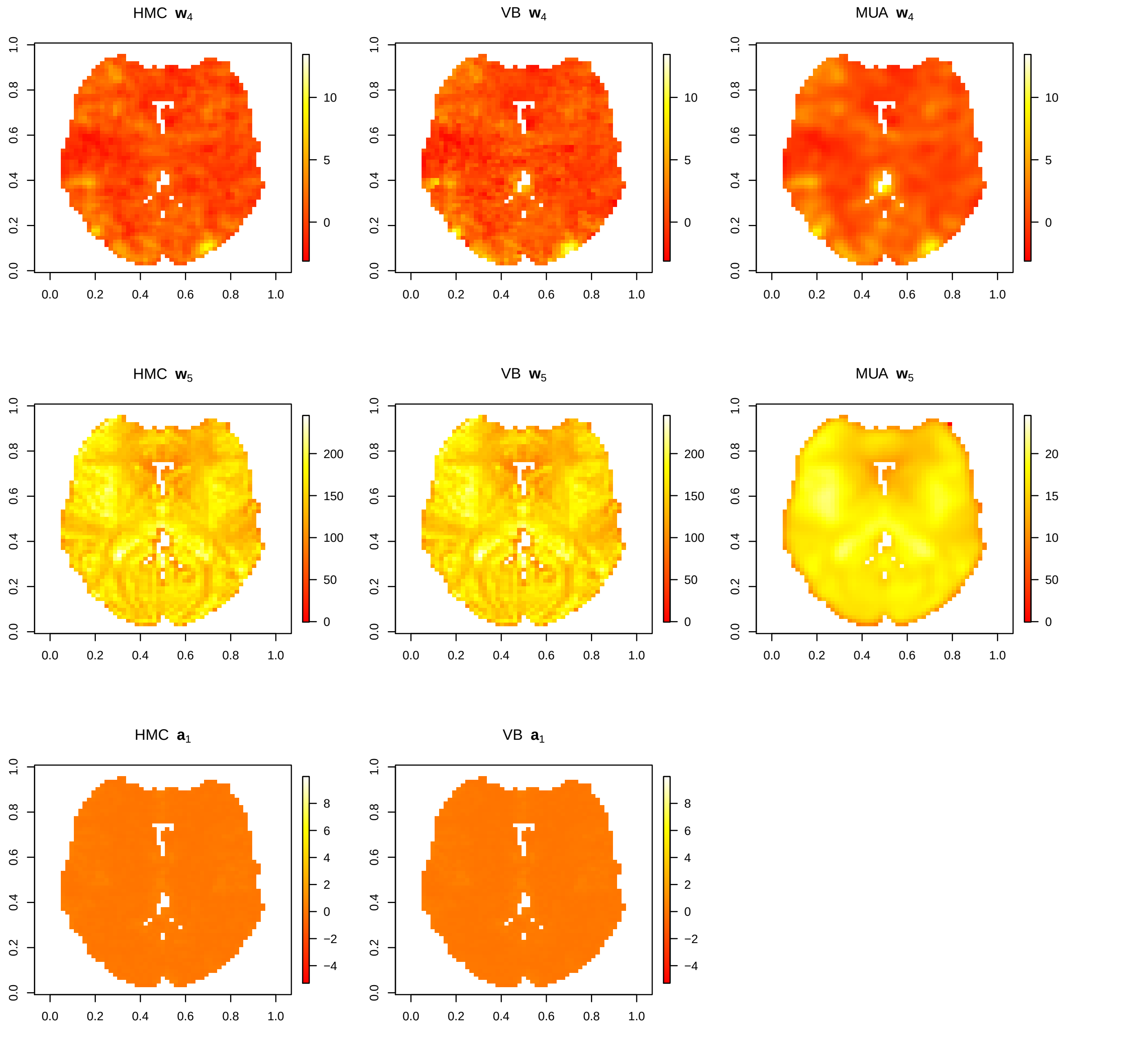

Figure 14. Image of posterior mean estimate of $\mathbf{w}_4, \mathbf{w}_5, \mathbf{a}_1$ from HMC, VB and MUA. These are the estimates from 26^{th} slices on the z-axis. Because MUA do not provide estimates of auto-regressive coefficients, we omit it here.

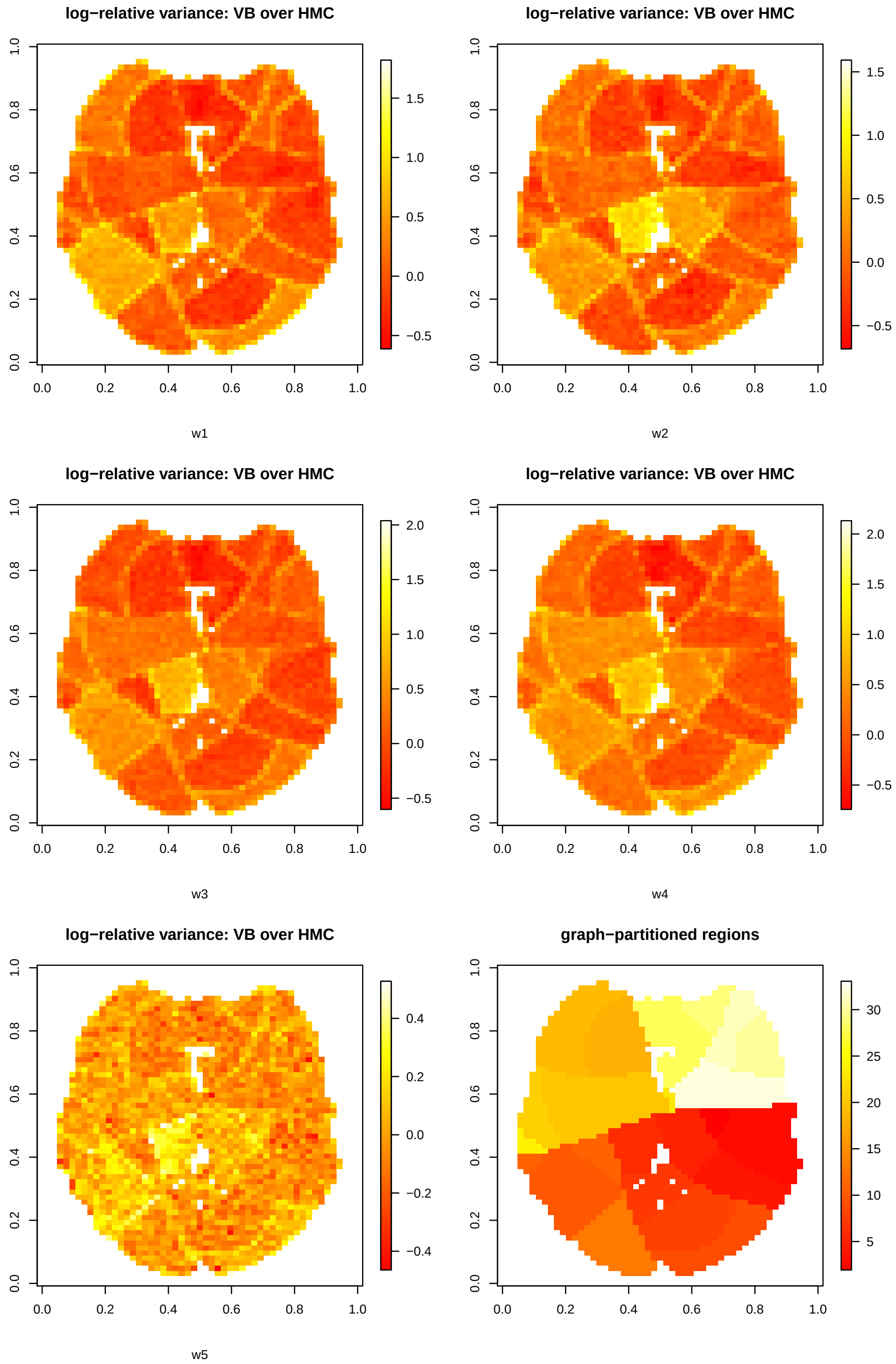

Figure 15. Log-relative ratio of marginal posterior variance from VB over HMC. The first five image corresponds to $\mathbf{w}_1$ to $\mathbf{w}_5$, the last one is the graph-partitioned regions by SPM VB. This is also the 26^{th} slice.

Prepared using simauth.cls