Supplementary Material for "Bayesian Spatial Binary Regression for Label Fusion in Structural Neuroimaging"

D. Andrew Brown* Christopher S. McMahan* Russell T. Shinohara†

Kristin A. Linn†

for the Alzheimer's Disease Neuroimaging Initiative‡

*School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, USA

†Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics, and Center for Biomedical Image Computing and Analytics, Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

‡Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`

Here we have included additional details concerning MCMC implementation. We also have additional figures referred to in the manuscript that are helpful, but not essential, to the reader.

1. MCMC DETAILS FOR IMPLEMENTATION

For posterior inference, we use a Metropolis-Hastings-within-Gibbs sampling scheme (Müller, 1991, and references therein). The necessary full conditional distributions are given in Section 1 of the Supplementary Material. Sampling $\boldsymbol{\delta}$ can be done following Gamerman (1997). Depending on $\boldsymbol{T}$, every voxel contributes information to either $\boldsymbol{\phi}_r$ or $\boldsymbol{\eta}_r$, but not both, due to the definitions of $\xi(\cdot, r)$ and $\psi(\cdot, r)$. Following Albert and Chib (1993), we introduce latent variables $\zeta^1_{vr} \mid T_v, \boldsymbol{\beta}_r, \phi_{vr} \overset{ind}{\sim} N(T_v(\boldsymbol{x}_{vr}^\top \boldsymbol{\beta}_r + \phi_{vr}), 1)$, $\zeta^0_{vr} \mid T_v, \boldsymbol{\gamma}_r, \eta_{vr} \overset{ind}{\sim} N((1 - T_v)(\boldsymbol{z}_{vr}^\top \boldsymbol{\gamma}_r + \eta_{vr}), 1)$, $v = 1, \ldots, V$. This induces conditional conjugacy, thus facilitating straightforward Gibbs sampling.

A convenient feature of CAR models is the availability of easily interpretable conditional distributions that follow from the Markov property. The CAR prior on $\boldsymbol{\phi}_r$ implies that $\phi_{vr} \mid \boldsymbol{\phi}_{(-v),r}, \tau_{\phi_r} \sim N(\rho\overline{\phi}_v, (w_{v\cdot}\tau_{\phi_r})^{-1})$, where $\boldsymbol{\phi}_{(-v),r} = (\phi_{1r}, \ldots, \phi_{v-1,r}, \phi_{v+1,r}, \ldots, \phi_{Vr})^\top \in \mathbb{R}^{V-1}$, $w_{v\cdot} = \sum_{k=1}^V w_{vk}$, and $\overline{\phi}_v = w_{v\cdot}^{-1} \sum_{k=1}^V w_{vk}\phi_{kr}$. Combined with the augmented data ζ^1_{vr}, this yields $\phi_{vr} \mid \zeta^1_{vr}, \boldsymbol{\phi}_{(-v),r}, \tau_\phi, \boldsymbol{\beta}_r, T_v \sim N(\mu^*_{v,r}, \sigma^{2,*}_v)$, where $\sigma^{2,*}_v = (T_v + \tau_\phi w_{v\cdot})^{-1}$ and $\mu^*_{v,r} = \sigma^{2,*}(T_v(\zeta^1_{vr} - \boldsymbol{x}_{vr}^\top \boldsymbol{\beta}_r) + \rho_\phi w_{v\cdot}\tau_\phi \overline{\phi}_v)$. The Markov property still holds since, conditional on its neighbors, ϕ_{vr} is independent of the rest of the field. Thus, we can update independent elements of $\boldsymbol{\phi}_r$ simultaneously (e.g., in parallel or through 'vectorized' calculations in `R`) while conditioning on all of their common neighbors. Such simultaneous updating is accomplished through *chromatic sampling* (Brown et al., 2021, and references therein), so named due to its association with coloring the GMRF graph. With simultaneous updates, chromatic Gibbs sampling dramatically reduces the cost of sequential single-site updating in `R` while avoiding the computational burden of drawing from high-dimensional Gaussian distributions. In fact, Brown et al. (2021) show that its computational cost increases at a much lower rate in `R` than block sampling via multivariate Gaussian draws. When coded in `C++` without parallelization,

chromatic sampling and single-site sampling are computationally similar, differing only in the order of site updates. For our application in Section 4, we found the convergence behavior of both schemes to be case-dependent. An in-depth investigation of this issue would be outside the scope of the current paper.

The following full conditional distributions follow readily from the full posterior density in Section 2,

$$\begin{aligned}
\boldsymbol{\beta}_r|\boldsymbol{\phi}_r, \tau_{\boldsymbol{\phi}_r} &\overset{ind.}{\sim} N\left(\boldsymbol{\mu}^*_{\boldsymbol{\beta}_r}, \boldsymbol{\Sigma}^*_{\boldsymbol{\beta}_r}\right), \quad r = 1, ..., R \\
\boldsymbol{\gamma}_r|\boldsymbol{\eta}_r, \tau_{\boldsymbol{\eta}_r} &\overset{ind.}{\sim} N\left(\boldsymbol{\mu}^*_{\boldsymbol{\gamma}_r}, \boldsymbol{\Sigma}^*_{\boldsymbol{\gamma}_r}\right), \quad r = 1, ..., R \\
\tau_{\boldsymbol{\phi}_r}|\boldsymbol{\alpha}_{\boldsymbol{\phi}_r}, \boldsymbol{\phi}_r &\overset{ind.}{\sim} Ga\left(a^*_{\boldsymbol{\phi}_r}, b^*_{\boldsymbol{\phi}_r}\right), \quad r = 1, ..., R \\
\tau_{\boldsymbol{\eta}_r}|\boldsymbol{\alpha}_{\boldsymbol{\eta}_r}, \boldsymbol{\eta}_r &\overset{ind.}{\sim} Ga\left(a^*_{\boldsymbol{\eta}_r}, b^*_{\boldsymbol{\eta}_r}\right), \quad r = 1, ..., R \\
T_v|\boldsymbol{\delta}, \boldsymbol{\phi}_v, \boldsymbol{\eta}_v, \boldsymbol{Y}_v &\overset{ind.}{\sim} Bernoulli\{p^*_{v1}/(p^*_{v1} + p^*_{v0})\}
\end{aligned} \tag{1}$$

where $\boldsymbol{\phi}_v = (\phi_{v1}, ..., \phi_{vR})^\top$, $\boldsymbol{\eta}_v = (\eta_{v1}, ..., \eta_{vR})^\top$, $\boldsymbol{Y}_v = (Y_{v1}, ..., Y_{vR})^\top$, and

$$
\begin{aligned}
\boldsymbol{\mu}^*_{\boldsymbol{\beta}} &= \boldsymbol{\Sigma}^*_{\boldsymbol{\beta}} \boldsymbol{X}_r^\top (\boldsymbol{D} - \rho_{\boldsymbol{\phi}_r} \boldsymbol{W}) \boldsymbol{\phi}_r \\
\boldsymbol{\mu}^*_{\boldsymbol{\gamma}_r} &= \boldsymbol{\Sigma}^*_{\boldsymbol{\gamma}} \boldsymbol{Z}_r^\top (\boldsymbol{D} - \rho_{\boldsymbol{\eta}_r} \boldsymbol{W}) \boldsymbol{\eta}_r \\
\boldsymbol{\Sigma}^*_{\boldsymbol{\beta}_r} &= \left\{ \boldsymbol{X}_r^\top (\boldsymbol{D} - \rho_{\boldsymbol{\phi}_r} \boldsymbol{W}) \boldsymbol{X}_r + \tau_{\boldsymbol{\phi}_r}^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\phi}_r}^{-1} \right\}^{-1} \\
\boldsymbol{\Sigma}^*_{\boldsymbol{\gamma}_r} &= \left\{ \boldsymbol{Z}_r^\top (\boldsymbol{D} - \rho_{\boldsymbol{\eta}_r} \boldsymbol{W}) \boldsymbol{Z}_r + \tau_{\boldsymbol{\eta}_r}^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\eta}_r}^{-1} \right\}^{-1} \\
a^*_{\boldsymbol{\phi}_r} &= a_{\boldsymbol{\phi}_r} + V/2 \\
a^*_{\boldsymbol{\eta}_r} &= a_{\boldsymbol{\eta}_r} + V/2 \\
b^*_{\boldsymbol{\phi}_r} &= b_{\boldsymbol{\phi}_r} + (\boldsymbol{\phi}_r - \boldsymbol{X}_r \boldsymbol{\beta}_r)^\top (\boldsymbol{D} - \rho_{\boldsymbol{\phi}_r} \boldsymbol{W}) (\boldsymbol{\phi}_r - \boldsymbol{X}_r \boldsymbol{\beta}_r)/2 \\
b^*_{\boldsymbol{\eta}_r} &= b_{\boldsymbol{\eta}_r} + (\boldsymbol{\eta}_r - \boldsymbol{Z}_r \boldsymbol{\gamma}_r)^\top (\boldsymbol{D} - \rho_{\boldsymbol{\eta}_r} \boldsymbol{W}) (\boldsymbol{\eta}_r - \boldsymbol{Z}_r \boldsymbol{\gamma}_r)/2 \\
p^*_{v1} &= \{ g^{-1}(\boldsymbol{c}_v^\top \boldsymbol{\delta}) \} \prod_{r=1}^{R} \Phi(\phi_{vr})^{Y_{vr}} \{1 - \Phi(\phi_{vr})\}^{(1-Y_{vr})} \\
p^*_{v0} &= \{ 1 - g^{-1}(\boldsymbol{c}_v^\top \boldsymbol{\delta}) \} \prod_{r=1}^{R} \Phi(\eta_{vr})^{(1-Y_{vr})} \{1 - \Phi(\eta_{vr})\}^{Y_{vr}}.
\end{aligned}
$$

The remaining steps are to sample $\boldsymbol{\delta}$ and the spatial random effects $\boldsymbol{\phi}_r$ and $\boldsymbol{\eta}_r$.

Introduce ζ^1_{vr} and ζ^0_{vr} and let $\mathcal{V}_\ell = \{v : T_v = \ell\},\ \ell = 0, 1$. Define $Y_{vr} = I(\zeta^1_{vr} \geq 0)$ for $v \in \mathcal{V}_1$, and similarly $Y_{vr} = I(\zeta^0_{vr} < 0)$ for $v \in \mathcal{V}_0$. Letting $\boldsymbol{\mathcal{T}} = \text{diag}(T_1, \ldots, T_V) \in \mathbb{R}^{V \times V}$, $\boldsymbol{\zeta}^1_r = (\zeta^1_{1r}, \ldots, \zeta^1_{Vr})^\top \in \mathbb{R}^V$, and $\boldsymbol{\zeta}^0_r = (\zeta^0_{1r}, \ldots, \zeta^0_{Vr})^\top \in \mathbb{R}^V$, this can be expressed as $\boldsymbol{\zeta}^1_r \mid \boldsymbol{\beta}_r, \phi_{vr}, \boldsymbol{\mathcal{T}} \sim N_V\{\boldsymbol{\mathcal{T}}(\boldsymbol{\phi}_r + \boldsymbol{X}_r \boldsymbol{\beta}_r), \boldsymbol{I}\}$ and $\boldsymbol{\zeta}^0_r \mid \boldsymbol{\beta}_r, \phi_{vr}, \boldsymbol{\mathcal{T}} \sim N_V\{(\boldsymbol{I} - \boldsymbol{\mathcal{T}})(\boldsymbol{\eta}_r + \boldsymbol{Z}_r \boldsymbol{\gamma}_r), \boldsymbol{I}\}$. Using the augmented data likelihood, the full conditional distributions of the spatial effects are given by $\boldsymbol{\phi}_r | \boldsymbol{\zeta}^1_r, \tau_{\boldsymbol{\phi}_r}, \boldsymbol{\beta}_r \sim N_V(\boldsymbol{\mu}^*_{\boldsymbol{\phi}_r}, \boldsymbol{\Sigma}^*_{\boldsymbol{\phi}_r})$ and $\boldsymbol{\eta}_r | \boldsymbol{\zeta}^0_r, \tau_{\boldsymbol{\eta}_r}, \boldsymbol{\gamma}_r \sim N_V(\boldsymbol{\mu}^*_{\boldsymbol{\eta}_r}, \boldsymbol{\Sigma}^*_{\boldsymbol{\eta}_r})$, for $r = 1, ..., R$, where $\boldsymbol{\Sigma}^*_{\boldsymbol{\phi}_r} = \left(\boldsymbol{\mathcal{T}} + \tau_{\boldsymbol{\phi}_r}(\boldsymbol{D} - \rho_\phi \boldsymbol{W})\right)^{-1}$,
$\boldsymbol{\Sigma}^*_{\boldsymbol{\eta}_r} = \left(\boldsymbol{I} - \boldsymbol{\mathcal{T}} + \tau_{\boldsymbol{\eta}_r}(\boldsymbol{D} - \rho_\eta \boldsymbol{W})\right)^{-1}$, $\boldsymbol{\mu}^*_{\boldsymbol{\phi}_r} = \boldsymbol{\Sigma}^*_{\boldsymbol{\phi}_r} \boldsymbol{\mathcal{T}}^\top (\boldsymbol{\zeta}^1_r - \boldsymbol{\mathcal{T}} \boldsymbol{X}_r \boldsymbol{\beta}_r)$, and $\boldsymbol{\mu}^*_{\boldsymbol{\eta}_r} = \boldsymbol{\Sigma}^*_{\boldsymbol{\eta}_r} (\boldsymbol{I} - \boldsymbol{\mathcal{T}})^\top (\boldsymbol{\zeta}^0_r - (\boldsymbol{I} - \boldsymbol{\mathcal{T}}) \boldsymbol{Z}_r \boldsymbol{\gamma}_r)$. For the conditional distribution of $\boldsymbol{\zeta}$, it can be shown that, for all v, r, $\zeta^1_{vr} \mid Y_{vr}, \phi_{vr}, \beta_r, T_v \overset{ind}{\sim} I(Y_{vr} = 0) TN\{T_v(\phi_{vr} + \boldsymbol{x}^\top_{vr} \boldsymbol{\beta}_r), 1, (-\infty, 0)\} + I(Y_{vr} = 1) TN\{T_v(\phi_{vr} + \boldsymbol{x}^\top_{vr} \boldsymbol{\beta}_r), 1, (0, \infty)\}$, and $\zeta^0_{vr} \mid Y_{vr}, \eta_{vr}, \beta_r, T_v \overset{ind}{\sim} I(Y_{vr} = 1) TN\{(1 - T_v)(\eta_{vr} + \boldsymbol{z}^\top_{vr} \boldsymbol{\gamma}_r), 1, (-\infty, 0)\} + I(Y_{vr} = 0) TN\{(1 - T_v)(\eta_{vr} + \boldsymbol{z}^\top_{vr} \boldsymbol{\gamma}_r), 1, (0, \infty)\}$, where $TN(\mu, \sigma^2, (a, b))$

denotes a Gaussian distribution with mean μ and variance σ^2 truncated to the interval (a, b).

Under our proposed model (1) - (5) in the manuscript, the full conditional density of $\boldsymbol{\delta}$ is $\pi(\boldsymbol{\delta}|\boldsymbol{T}, \boldsymbol{C}) \propto \exp\{-\boldsymbol{\delta}^\top \boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1}\boldsymbol{\delta}/2 + \sum_{v=1}^{V} T_v\theta_v - b(\theta_v)\}$, where $\boldsymbol{C} \in \mathbb{R}^{V\times J}$ is the design matrix constructed from $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_V$, $\theta_v = \log[g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta})/\{1 - g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta})\}]$, and $b(\cdot) = \log\{1 + \exp(\cdot)\}$. Define $\boldsymbol{T}(\boldsymbol{\delta}) = (T_1(\boldsymbol{\delta}), \ldots, T_V(\boldsymbol{\delta}))^\top \in \mathbb{R}^V$ and $\boldsymbol{Q}(\boldsymbol{\delta}) = \text{diag}(Q_1(\boldsymbol{\delta}), \ldots, Q_V(\boldsymbol{\delta}))^\top$ with $T_v(\boldsymbol{\delta}) = \boldsymbol{c}_v^\top\boldsymbol{\delta} + \{T_v - g^{-1}(\mathbf{c}_v^\top\boldsymbol{\delta})\}\dot{g}(g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta}))$ and $Q_v^{-1}(\boldsymbol{\delta}) = \ddot{b}(\theta_v)\{\dot{g}(g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta}))\}^2$ for $v = 1, \ldots, V$. Then, given the current iterate $\boldsymbol{\delta}^{(t)}$, the proposal distribution is $\boldsymbol{\delta}^{(t+1)} \mid \boldsymbol{\delta}^{(t)} \sim N_J(\boldsymbol{m}^{(t)}, \boldsymbol{V}^{(t)})$, where $\boldsymbol{m}^{(t)} = \{\boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1} + \boldsymbol{C}^\top\boldsymbol{Q}(\boldsymbol{\delta}^{(t))})\boldsymbol{C}\}^{-1}\,\boldsymbol{C}^\top\boldsymbol{Q}(\boldsymbol{\delta}^{(t)})\boldsymbol{T}(\boldsymbol{\delta}^{(t)})$ and $\boldsymbol{V}^{(t)} = \{\boldsymbol{\Sigma}_{\boldsymbol{\delta}}^{-1} + \boldsymbol{C}^\top\boldsymbol{Q}(\boldsymbol{\delta}^{(t)})\boldsymbol{C}\}^{-1}$. If the prior on $\boldsymbol{\delta}$ is elicited through the CMP approach, this proposal mechanism can still be used after suitable modification.

Under the CMP elicitation of Bedrick et al. (1996) with the logistic link, we add J covariate-binomial reponse pairs, $\{(T_{V+1}, \boldsymbol{c}_{V+1}), \ldots, (T_{V+J}, \boldsymbol{c}_{V+J})\} = \{(a_{\widetilde{p}_j}, \widetilde{\boldsymbol{c}}_j^\top)\}_{j=1}^{J}$, where $T_{V+j} = a_{\widetilde{p}_j}$ and $\widetilde{\boldsymbol{c}}_j^\top$, $j = 1, 2, ..., J$, are psuedo true statuses and linearly independent covariate vectors for the j subpopulations, respectively. Unlike binomial observations, the values T_{V+j}, are not necessarily integer-valued. The vector of transformed observations $\boldsymbol{T}(\boldsymbol{\delta})$ then has the following individual components

$$\begin{aligned} T_v(\boldsymbol{\delta}) &= \boldsymbol{c}_v^\top\boldsymbol{\delta} + \{T_v - g^{-1}(\boldsymbol{x}_v^\top\boldsymbol{\delta})\}g'(g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta})), \text{ for } v = 1, ..., V, \\ T_{V+j}(\boldsymbol{\delta}) &= \widetilde{\boldsymbol{c}}_j^\top\boldsymbol{\delta} + \{a_{\widetilde{p}_j} - \widetilde{n}_j g^{-1}(\widetilde{\boldsymbol{c}}_j^\top\boldsymbol{\delta})\}g'(\widetilde{n}_j g^{-1}(\widetilde{\boldsymbol{c}}_j^\top\boldsymbol{\delta})), \text{ for } j = 1, ..., J, \end{aligned}$$

where $\widetilde{n}_j$ are the binomial weights. Similarly, define the diagonal matrix of weights $\boldsymbol{Q}(\boldsymbol{\delta})$ to have diagonal components

$$\begin{aligned} \boldsymbol{Q}_v^{-1}(\boldsymbol{\delta}) &= b''(\theta_v)\{g'(g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta}))\}^2, \text{ for } v = 1, ..., V, \\ \boldsymbol{Q}_{V+j}^{-1}(\boldsymbol{\delta}) &= \widetilde{n}_j b''(\widetilde{\theta}_j)\{g'(\widetilde{n}_j g^{-1}(\widetilde{\boldsymbol{c}}_j^\top\boldsymbol{\delta}))\}^2, \text{ for } j = 1, ..., J, \end{aligned}$$

where $\theta_v = \log[g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\delta})/\{1 - g^{-1}(\boldsymbol{c}_v^\top\boldsymbol{\beta})\}]$ and $\widetilde{\theta}_j = \log[g^{-1}(\widetilde{\boldsymbol{c}}_j^\top\boldsymbol{\delta})/\{1 - g^{-1}(\widetilde{\boldsymbol{c}}_j^\top\boldsymbol{\delta})\}]$. The proposal density based on the previously sampled value of $\boldsymbol{\delta}$, say $\boldsymbol{\delta}^{(g-1)}$, is multivariate

normal with mean and covariance matrix given by

$$\begin{aligned} \boldsymbol{m}^{(g)} &= \{\boldsymbol{C}^\top \boldsymbol{Q}(\boldsymbol{\delta}^{(g-1)})\boldsymbol{C}\}^{-1}\boldsymbol{C}^\top \boldsymbol{Q}(\boldsymbol{\delta}^{(g-1)})\boldsymbol{T}(\boldsymbol{\delta}^{(g-1)}) \\ \boldsymbol{C}^{(g)} &= \{\boldsymbol{C}^\top \boldsymbol{Q}(\boldsymbol{\delta}^{(g-1)})\boldsymbol{C}\}^{-1}, \end{aligned}$$

where $\boldsymbol{C}$ is the usual design matrix for the covariates $\boldsymbol{c}_v$ and $\widetilde{\boldsymbol{c}}_j$. Using this mean vector and covariance matrix, the remainder of the Metropolis-Hastings algorithm follows.

In the simulation study in Subsection 3.2, the prior knowledge about pseudo-cases (i)-(iv) in the manuscript are reflected with data augmentation. With covariates (intercept, rescaled SDL, standardized intensity, interaction), we augment the model with pseudo-design matrix $\widetilde{\boldsymbol{C}} = [(1,0,5,0)^\top, (1,1,-3,-3)^\top, (1,0.5,0,0)^\top, (1,0.7,5,3.5)^\top]^\top$ and psuedo-observations $\boldsymbol{a}_{\widetilde{p}} = (392,\ 8, 50, 8)^\top$ with weights $n^* = (400, 400, 100, 400)^\top$. In the hippocampus segmentation example in Section 4 of the manuscript, the covariate vector is indexed as (intercept, rescaled SDL, gray matter inclusion indicator, interaction). The psuedo-design matrix for cases (i)-(iv) is $\widetilde{\boldsymbol{C}} = [(1,0,1,0)^\top, (1,0.5,0,0)^\top, (1,0.5,1,0.5)^\top, (1,0.1,0,0)^\top]^\top$. The pseudo-observations are created with $a_{\widetilde{p}_j} = n_j^* p_j^*,\ j = 1,\ldots,4$, with $p_1^* = 0.9999, p_2^* = 0.001, p_3^* = 0.1, p_4^* = 0.01$, and $n_1^* = n_2^* = n_3^* = n_4^* = 25000$.

2. SUPPLEMENTARY FIGURES

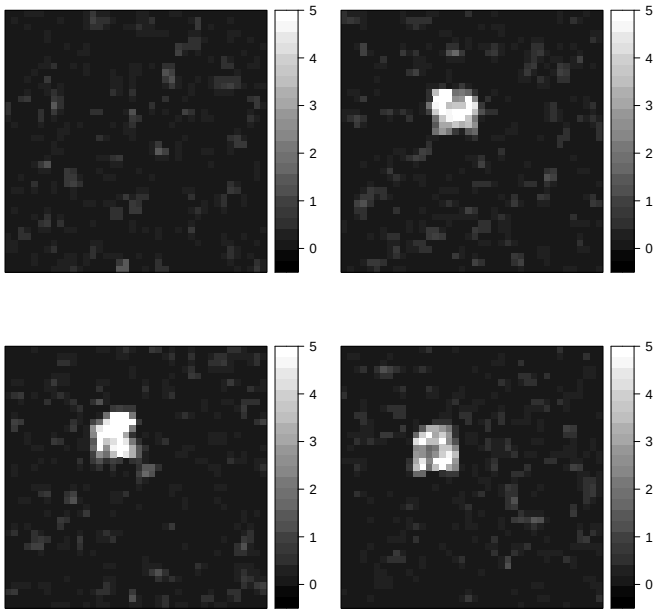

Figure 1: Simulated squared intensity differences associated to the one reliable (upper left) and three unreliable segmentations in the simulated label fusion example from Subsection 3.2.

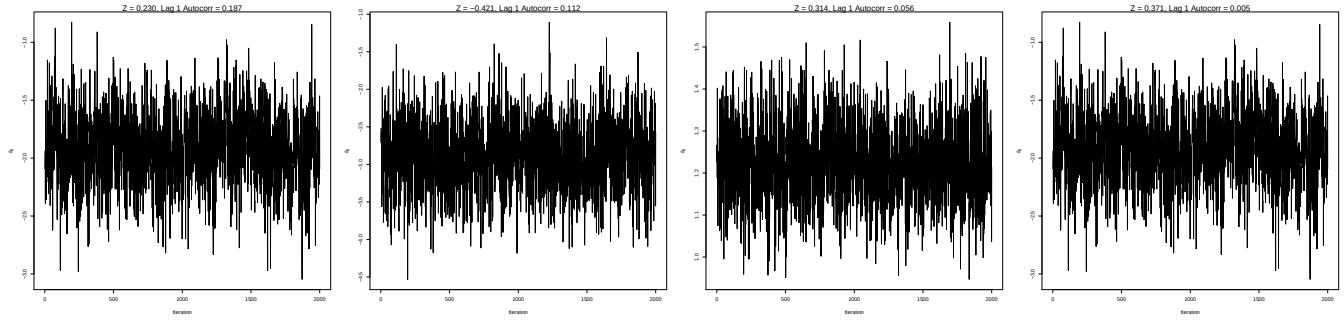

Figure 2: MCMC Trace plots of the $\boldsymbol{\delta}$ coefficients for the simulated label fusion example. The Geweke (Z) statistics and lag 1 autocorrelations are displayed above each plot.

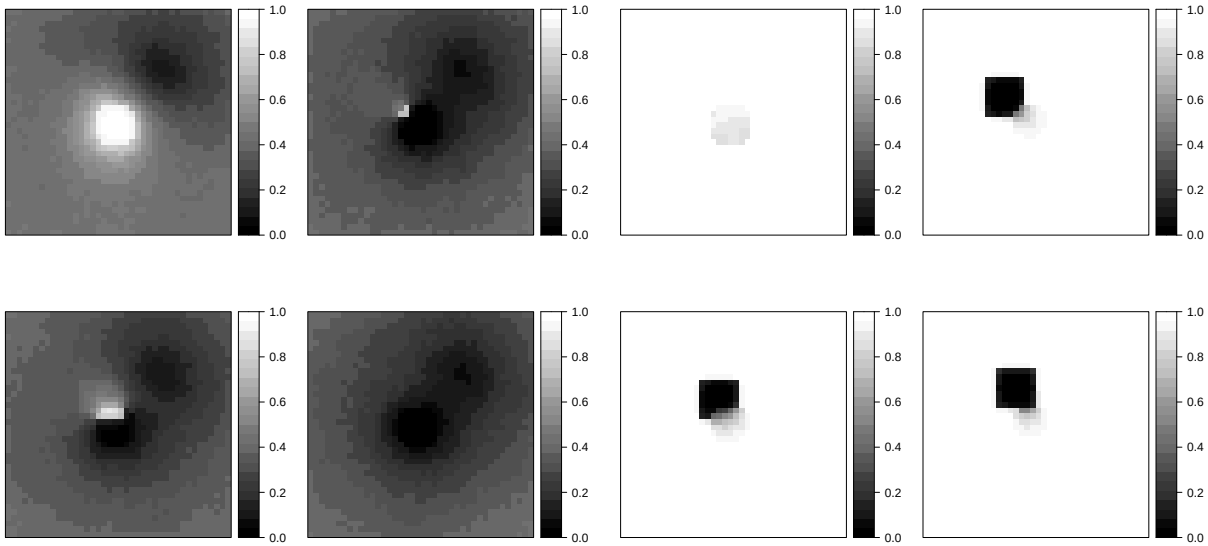

Figure 3: Posterior means of the atlas sensitivity fields (left four panels) and specificity fields (right four panels) in the simulated example.

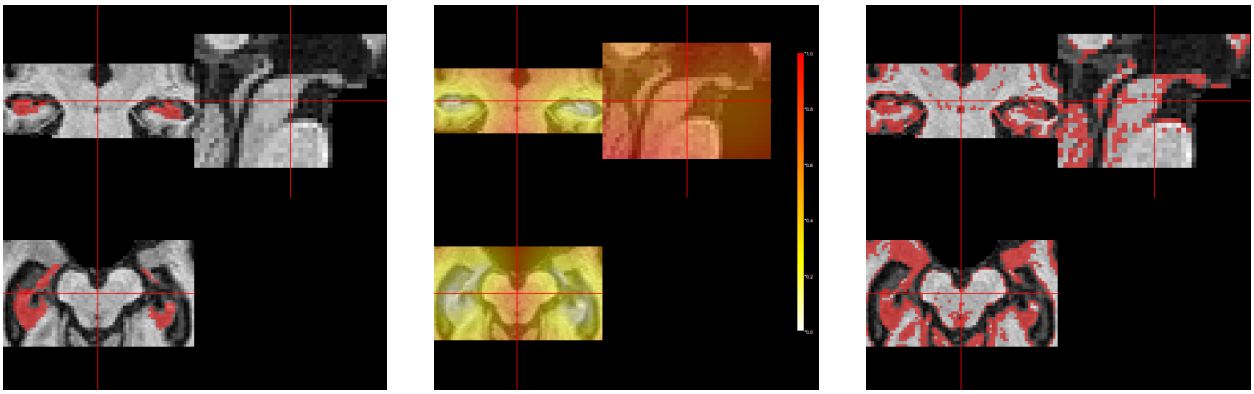

Figure 4: Manually segmented hippocampus (left), weighted and summed SDL map of the atlases (middle) and gray matter segmentation of the target (right) for the label fusion application. The gray matter segmentation is obtained by dichotomizing a tissue class segmentation obtained from the `ATROPOS` *algorithm.*

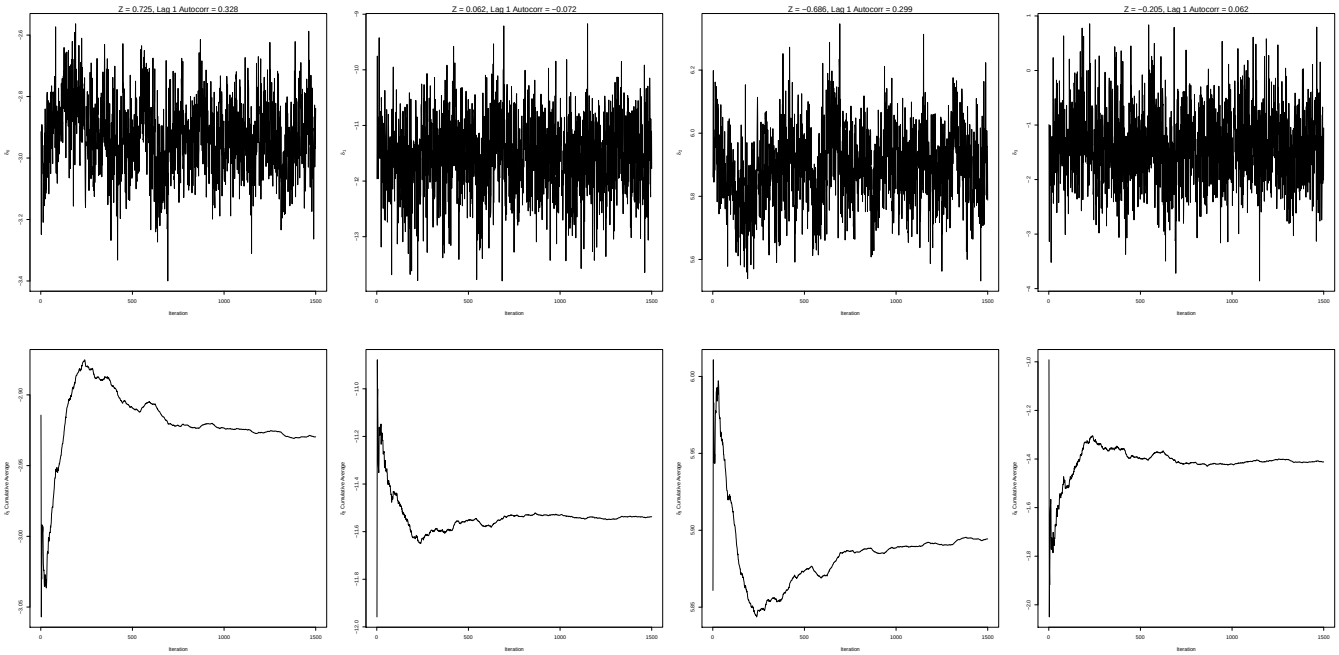

Figure 5: Trace plots and ergodic averages for the $\boldsymbol{\delta}$ coefficients for the single-brain hippocampus segmentation application. The Geweke Z statistics and lag 1 autocorrelations are displayed above the trace plots.

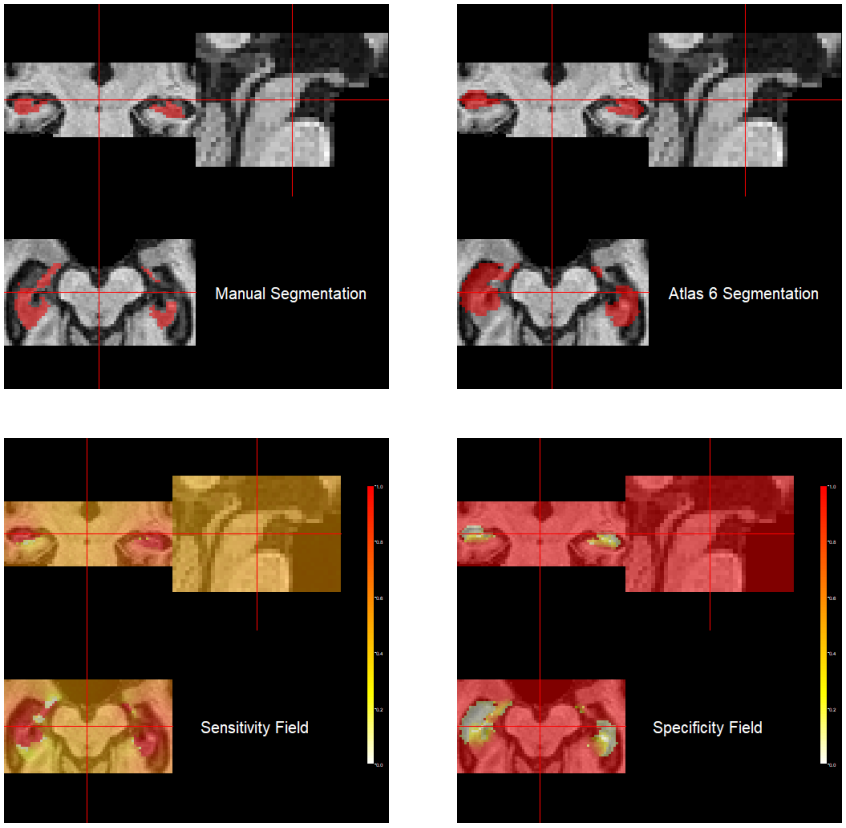

Figure 6: Manual segmentation (top left), one selected atlas segmentation (top right), and the mean sensitivity (bottom left) and specificity (bottom right) maps associated with this atlas.

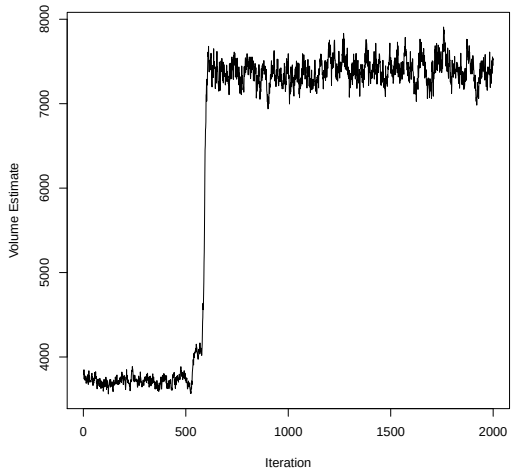

Figure 7: Trace plot of the hippocampal volume for subject 1263 mentioned in Subsection 4.2. The "jump" suggests a bi-modal marginal posterior distribution.

REFERENCES

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous data. *Journal of the American Statistical Association*, 88(422):669–679.

Brown, D. A., McMahan, C. S., and Self, S. W. (2021). Sampling strategies for fast updating of Gaussian Markov random fields. *The American Statistician*, 75:52–65.

Gamerman, D. (1997). Sampling from the posterior distribution in generalized linear mixed models. *Statistics and Computing*, 7:57–68.

Müller, P. (1991). A generic approach to posterior integration and Gibbs sampling. Technical Report, Purdue University.