

Supplementary Materials for “A Unified Analysis of Structured Sonar-terrain Data using Bayesian Functional Mixed Models”

Hongxiao Zhu

Department of Statistics, Virginia Tech, Blacksburg, VA 24061

Philip Caspers

Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA 24061

Jeffrey S. Morris

The University of Texas M.D. Anderson Cancer Center, Houston, TX 77230

Xiaowei Wu

Department of Statistics, Virginia Tech, Blacksburg, VA 24061

and

Rolf Müller

Department of Mechanical Engineering, Virginia Tech, Blacksburg, VA 24061

December 14, 2016

1 More Details of the Sonar-terrain Experiment

We collected echoes using a dual frequency bistatic sonar head. A schematic illustration of the device is provided in Figure 1. The sonar head consists of two transducers: one

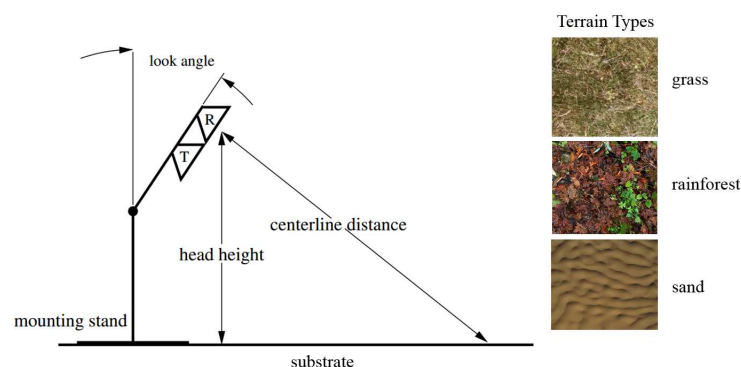


Figure 1: The setup of the sonar head in the sonar-terrain experiment.

transmitter (denoted as T) used to ensonify the environment and one receiver (denoted as R) used to receive echoes. The receiver consists of a small, omnidirectional microphone mounted at the throat of a conical horn (Fletcher, 1992). The conical horn was used to narrow the directionality and provide amplification along the line of sight. The receiver and transmitter were collocated and mounted on a portable stand so that the height of the sonar head was approximately 0.5 meter above the ground, and the look angle was set so that the centerline (the line that is perpendicular to the receiver cone mouth and connects the receiver with the substrate) was one meter long. The sonar head generates an acoustic output using two commercially available piezo electric emission elements at resonant frequencies of 25 and 40 khz respectively.

Three terrain substrates were considered: grass, rainforest, and sand. Example patches are demonstrated in Figure 1. The rainforest substrate is a simulated tropical rainforest floor composed of randomly scattered fallen branches, damp loose dirt, and leaves. Echo footprints of grass and rainforest were taken outdoors, while the sand footprints were measured at an indoor area with a few inches of sand spread over a concrete floor.

2 MCMC Algorithms for the FMM models

For notational convenience, based on the model (4) in the main text, we pool the two design matrices for the fixed effects to form $\tilde{\mathbf{X}} = [\mathbf{V}, \mathbf{X}]$, and pool corresponding coefficients to form $\tilde{\mathbf{B}}_{lk}^* = [(\mathbf{G}_{lk}^*)^T, B_{lk}^*]^T$. Model (4) then becomes:

$$\mathbf{d}_{lk} = \tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^* + \mathbf{Z}\mathbf{U}_{lk}^* + \mathbf{E}_{lk}^*, \quad (1)$$

where $\tilde{\mathbf{X}}$ is a design matrix of size N by $A + 1$, and $\tilde{\mathbf{B}}_{lk}^*$ is a vector of length $A + 1$.

2.1 Algorithm for Gfmm

We design the following Markov Chain Monte Carlo algorithm for posterior sampling in Gfmm:

[**Step 0.**] Initialize $\tilde{\mathbf{B}}^*$, $\{q_{lk}\}$, $\{s_{lk}\}$, $\{\tau_{alk}\}$ and set prior parameters. In particular, we use Henderson’s mixed model equations to obtain estimates for $\tilde{\mathbf{B}}_{lk}^*$ and the variance compo-

nents (Searle et al., 1992, page 275-286) for each (j, k) , and use these estimates as initial values.

[**Step 1.**] Update $(\gamma_{alk} | \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \{q_{lk}\}, \{s_{lk}\})$ and $(\tilde{B}_{alk}^* | \gamma_{alk}, \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \{q_{lk}\}, \{s_{lk}\})$ for $a = 1, \dots, A + 1$, where $\tilde{\mathbf{B}}_{(-a)lk}^*$ is the vector of $\tilde{\mathbf{B}}_{lk}^*$ with the a th component removed.

From model (1), we see that $\mathbf{d}_{lk} | \tilde{\mathbf{B}}_{lk}^*, \{q_{lk}\}, \{s_{lk}\} \sim N(\tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^*, \Sigma_{lk})$, where $\Sigma_{lk} = q_{lk}\mathbf{Z}\mathbf{Z}^T + s_{lk}\mathbf{I}_N$. We first update γ_{alk} by calculating the conditional odds:

$$\begin{aligned} \text{Conditional Odds} &= \frac{f(\gamma_{alk} = 1 | \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk})}{f(\gamma_{alk} = 0 | \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk})} \\ &= \frac{f(\mathbf{d}_{lk} | \gamma_{alk} = 1, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk})}{f(\mathbf{d}_{lk} | \gamma_{alk} = 0, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk})} \cdot \frac{f(\gamma_{alk} = 1)}{f(\gamma_{alk} = 0)} \\ &= \text{Conditional Bayes Factor} \cdot \text{Prior Odds}. \end{aligned}$$

Further derivations show that

$$\begin{aligned} \text{Conditional Bayes Factor} &= \frac{f(\mathbf{d}_{lk} | \gamma_{alk} = 1, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk})}{f(\mathbf{d}_{lk} | \gamma_{alk} = 0, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk})} \\ &= (1 + \frac{\tau_{alk}}{V_{alk}})^{-1/2} \exp \left\{ \frac{1}{2} \zeta_{alk}^2 (1 + V_{alk}/\tau_{alk})^{-1} \right\}, \end{aligned}$$

where $V_{alk} = [(\tilde{\mathbf{X}}_{(a)lk})^T (\Sigma_{lk})^{-1} \tilde{\mathbf{X}}_{(a)lk}]^{-1}$, $\tilde{\mathbf{X}}_{(a)lk}$ is the a th column of matrix $\tilde{\mathbf{X}}_{lk}$, $\zeta_{alk} = \hat{B}_{alk}^* / \sqrt{V_{alk}}$, and $\hat{B}_{alk}^* = V_{alk} \tilde{\mathbf{X}}_{(a)lk}^T \Sigma_{lk}^{-1} \tilde{\mathbf{d}}_{lk}$ for $\tilde{\mathbf{d}}_{lk} = \mathbf{d}_{lk} - \sum_{l=1, l \neq a}^{A+1} \tilde{\mathbf{X}}_{(l)lk} \tilde{B}_{llk}^*$.

Given $\{\gamma_{alk}\}$, we then update $\tilde{\mathbf{B}}_{lk}^*$ by $(\tilde{B}_{alk}^* | \gamma_{alk} = 0, \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk}) = \delta_0$, and $(\tilde{B}_{alk}^* | \gamma_{alk} = 1, \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \Sigma_{lk}) \sim N(\mu_{alk}^0, V_{alk}^0)$, where $\mu_{alk}^0 = \hat{B}_{alk}^* / (1 + V_{alk}/\tau_{alk})$ and $V_{alk}^0 = V_{alk} / (1 + V_{alk}/\tau_{alk})$.

[**Step 2.**] Update $\{q_{lk}\}$ and $\{s_{lk}\}$ using Metropolis-Hastings. First propose new values of q_{lk}, s_{lk} using log transform, e.g., $\log(\tilde{q}_{lk}) = \log(q_{lk}) + \delta\epsilon$, for $\epsilon \sim N(0, 1)$ and δ is the step size. Second calculate the proposal ratio of the random walk proposal:

$$\frac{f(q_{lk} | \tilde{q}_{lk})}{f(\tilde{q}_{lk} | q_{lk})} = \frac{\tilde{q}_{lk}}{q_{lk}}.$$

The new parameters $\{\tilde{q}_{lk}, \tilde{s}_{lk}\}$ are accepted with probability $\min\{o_{lk}, 1\}$, where

$$o_{lk} = \frac{f(\mathbf{d}_{lk} | \tilde{q}_{lk}, \tilde{s}_{lk}, \cdot) \pi(\tilde{q}_{lk}) \pi(\tilde{s}_{lk}) f(q_{lk} | \tilde{q}_{lk}) f(s_{lk} | \tilde{s}_{lk})}{f(\mathbf{d}_{lk} | q_{lk}, s_{lk}, \cdot) \pi(q_{lk}) \pi(s_{lk}) f(\tilde{q}_{lk} | q_{lk}) f(\tilde{s}_{lk} | s_{lk})}.$$

Here we have assumed $q_{lk} \sim \text{Inv-Gamma}(a_{lk}, b_{lk})$, therefore

$$\log \frac{\pi(\tilde{q}_{lk})}{\pi(q_{lk})} = (a_{lk} + 1)(\log(q_{lk}) - \log(\tilde{q}_{lk})) + b_{lk}(1/q_{lk} - 1/\tilde{q}_{lk}).$$

Similar formula holds for the ratio $\pi(\tilde{s}_{lk})/\pi(s_{lk})$. The log likelihood ratio takes the form

$$\log \frac{f(\mathbf{d}_{lk}|\tilde{q}_{lk}, \tilde{s}_{lk}, \cdot)}{f(\mathbf{d}_{lk}|q_{lk}, s_{lk}, \cdot)} = \frac{|\tilde{\Sigma}_{lk}|^{-1/2} \exp\{-1/2(\mathbf{d}_{lk} - \tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^*)^T \tilde{\Sigma}_{lk}^{-1}(\mathbf{d}_{lk} - \tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^*)\}}{|\Sigma_{lk}|^{-1/2} \exp\{-1/2(\mathbf{d}_{lk} - \tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^*)^T \Sigma_{lk}^{-1}(\mathbf{d}_{lk} - \tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^*)\}}$$

where $\Sigma_{lk} = q_{lk}\mathbf{Z}\mathbf{Z}^T + s_{lk}\mathbf{I}_N$, and $\tilde{\Sigma}_{lk} = \tilde{q}_{lk}\mathbf{Z}\mathbf{Z}^T + \tilde{s}_{lk}\mathbf{I}_N$

[Step 3.] Update \mathbf{U}_{lk}^* from $(\mathbf{U}_{lk}^*|\cdot) \sim N(\boldsymbol{\mu}_{\mathbf{U}_{lk}}, \mathbf{V}_{\mathbf{U}_{lk}})$, where $\mathbf{V}_{\mathbf{U}_{lk}} = ((s_{lk})^{-1}\mathbf{Z}^T\mathbf{Z} + (\mathbf{I}_N(q_{lk})^{-1})^{-1})^{-1}$, $\boldsymbol{\mu}_{\mathbf{U}_{lk}} = \mathbf{V}_{\mathbf{U}_{lk}}\mathbf{Z}^T(s_{lk})^{-1}(\mathbf{d}_{lk} - \tilde{\mathbf{X}}\tilde{\mathbf{B}}_{lk}^*)$. This step is optional.

Repeat Steps 1-3 until reaching a pre-specified maximum number of iterations.

2.2 Algorithm for Rfmm

We denote the scaling parameters $\boldsymbol{\theta}_{lk} = \{(\lambda_{ilk})_i, (\phi_{mlk})_m, (\psi_{alk})_a\}$, and denote $\boldsymbol{\nu}^2 = \{(\nu_{lk}^E)^2\}_{lk}, \{(\nu_{lk}^U)^2\}_{lk}, \{(\nu_{aj}^G)^2\}_{aj}$. We design the following Markov Chain Monte Carlo algorithm for posterior sampling in Rfmm:

[Step 0.] Initialize $\{\tilde{\mathbf{B}}_{lk}^*\}$, $\{\boldsymbol{\theta}_{lk}\}$, $\boldsymbol{\nu}^2$, and set the prior parameters (a^E, b^E) , (a^U, b^U) , (a^G, b^G) , and (a^π, b^π) . In particular, we use Henderson's mixed model equations to obtain MLE estimates for $\tilde{\mathbf{B}}_{lk}^*$, \mathbf{U}_{lk}^* and the variance components (Searle et al., 1992, page 275-286) for each (j, k) , and use these estimates to generate initial values for $\{\boldsymbol{\theta}_{lk}\}$, $\boldsymbol{\nu}^2$. We set prior parameters (a^E, b^E) , (a^U, b^U) , (a^G, b^G) , and (a^π, b^π) by controlling the mean of the prior distributions to be the initial values and the variances to be reasonably large.

[Step 1.] Update $(\gamma_{alk}|\mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\theta}_{lk}, \boldsymbol{\nu}^2)$ and $(\tilde{B}_{alk}^*|\gamma_{alk}, \mathbf{d}_{lk}, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\theta}_{lk}, \boldsymbol{\nu}^2)$ for $a = 1, \dots, A+1$, where $\tilde{\mathbf{B}}_{(-a)lk}^*$ is the vector of $\tilde{\mathbf{B}}_{lk}^*$ with the a th component removed. For computation convenience, we firstly pre-multiply model (1) by $\boldsymbol{\Lambda}_{lk}^{-1/2}$, where $\boldsymbol{\Lambda}_{lk} = \text{diag}\{\lambda_{1lk}, \dots, \lambda_{Nlk}\}$.

This transforms model (1) to:

$$\mathbf{d}_{lk}^+ = \tilde{\mathbf{X}}_{lk}^+ \tilde{\mathbf{B}}_{lk}^* + \mathbf{Z}_{lk}^+ \mathbf{U}_{lk}^* + \mathbf{E}_{lk}^+$$

where $\mathbf{d}_{lk}^+ = \boldsymbol{\Lambda}_{lk}^{-1/2}\mathbf{d}_{lk}$, $\tilde{\mathbf{X}}_{lk}^+ = \boldsymbol{\Lambda}_{lk}^{-1/2}\tilde{\mathbf{X}}$, $\mathbf{Z}_{lk}^+ = \boldsymbol{\Lambda}_{lk}^{-1/2}\mathbf{Z}$ and $\mathbf{E}_{lk}^+ = \boldsymbol{\Lambda}_{lk}^{-1/2}\mathbf{E}_{lk}^*$. We see that $\mathbf{d}_{lk}^+|\tilde{\mathbf{B}}_{lk}^*, \boldsymbol{\theta}_{lk}, \boldsymbol{\nu}^2 \sim N(\tilde{\mathbf{X}}_{lk}^+ \tilde{\mathbf{B}}_{lk}^*, \boldsymbol{\Sigma}_{lk}^+)$, where $\boldsymbol{\Sigma}_{lk}^+ = \mathbf{Z}_{lk}^+ \boldsymbol{\Phi}_{lk} (\mathbf{Z}_{lk}^+)^T + \mathbf{I}_N$, and $\boldsymbol{\Phi}_{lk} = \text{diag}\{\phi_{mlk}\}_m$.

We first update γ_{alk} by calculating the conditional odds:

$$\begin{aligned}
\text{Conditional Odds} &= \frac{f(\gamma_{alk} = 1 | \mathbf{d}_{lk}^+, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2)}{f(\gamma_{alk} = 0 | \mathbf{d}_{lk}^+, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2)} \\
&= \frac{f(\mathbf{d}_{lk}^+ | \gamma_{alk} = 1, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2)}{f(\mathbf{d}_{lk}^+ | \gamma_{alk} = 0, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2)} \cdot \frac{f(\gamma_{alk} = 1)}{f(\gamma_{alk} = 0)} \\
&= \text{Conditional Bayes Factor} \cdot \text{Prior Odds}.
\end{aligned}$$

Further derivations show that

$$\begin{aligned}
\text{Conditional Bayes Factor} &= \frac{f(\mathbf{d}_{lk}^+ | \gamma_{alk} = 1, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2)}{f(\mathbf{d}_{lk}^+ | \gamma_{alk} = 0, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2)} \\
&= (1 + \frac{\psi_{alk}}{V_{alk}})^{-1/2} \exp \left\{ \frac{1}{2} \zeta_{alk}^2 (1 + V_{alk}/\psi_{alk})^{-1} \right\},
\end{aligned}$$

where $V_{alk} = [(\tilde{\mathbf{X}}_{(a)lk}^+)^T (\boldsymbol{\Sigma}_{lk}^+)^{-1} \tilde{\mathbf{X}}_{(a)lk}^+]^{-1}$, $\tilde{\mathbf{X}}_{(a)lk}^+$ is the a th column of matrix $\tilde{\mathbf{X}}_{lk}^+$, $\zeta_{alk} = \hat{B}_{alk}^*/\sqrt{V_{alk}}$, and $\hat{B}_{alk}^* = V_{alk}(\tilde{\mathbf{X}}_{(a)lk}^+)^T (\boldsymbol{\Sigma}_{lk}^+)^{-1} \tilde{\mathbf{d}}_{lk}^+$ for $\tilde{\mathbf{d}}_{lk}^+ = \mathbf{d}_{lk}^+ - \sum_{l=1, l \neq a}^{A+1} \tilde{\mathbf{X}}_{(l)lk}^+ \tilde{B}_{ilk}^*$.

Given $\{\gamma_{alk}\}$, we then update $\tilde{\mathbf{B}}_{lk}^*$ by $(\tilde{B}_{alk}^* | \gamma_{alk} = 0, \mathbf{d}_{lk}^+, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2) = \delta_0$, and $(\tilde{B}_{alk}^* | \gamma_{alk} = 1, \mathbf{d}_{lk}^+, \tilde{\mathbf{B}}_{(-a)lk}^*, \boldsymbol{\Sigma}_{lk}^+(\boldsymbol{\theta}_{lk}), \boldsymbol{\nu}^2) \sim N(\mu_{alk}^0, V_{alk}^0)$, where $\mu_{alk}^0 = \hat{B}_{alk}^*/(1 + V_{alk}/\psi_{alk})$ and $V_{alk}^0 = V_{alk}/(1 + V_{alk}/\psi_{alk})$.

[Step 2.] Update $(\mathbf{U}_{lk}^* | \tilde{\mathbf{B}}_{lk}^*, \mathbf{d}_{lk}^+, \mathbf{Z}_{lk}^+, \tilde{\mathbf{X}}_{lk}^+, \boldsymbol{\nu}^2) \sim N(\boldsymbol{\mu}_u, \mathbf{V}_u)$, where

$$\begin{aligned}
\boldsymbol{\mu}_u &= \{(\mathbf{Z}_{lk}^+)^T \mathbf{Z}_{lk}^+ + \boldsymbol{\Phi}_{lk}^{-1}\}^{-1} (\mathbf{Z}_{lk}^+)^T (\mathbf{d}_{lk}^+ - \tilde{\mathbf{X}}_{lk}^+ \tilde{\mathbf{B}}_{lk}^*) \\
\mathbf{V}_u &= \{(\mathbf{Z}_{lk}^+)^T \mathbf{Z}_{lk}^+ + \boldsymbol{\Phi}_{lk}^{-1}\}^{-1}.
\end{aligned}$$

[Step 3.] Update $(\boldsymbol{\theta}_{lk} | \tilde{\mathbf{B}}_{lk}^*, \mathbf{U}_{lk}^*, \mathbf{d}_{lk}, \boldsymbol{\nu}^2)$. In particular,

$$(\lambda_{ilk}^{-1} | \mathbf{d}_{ilk}, \tilde{\mathbf{B}}_{lk}^*, \mathbf{U}_{lk}^*, (\nu_{lk}^E)^2) \sim \text{Inv-Gauss} \left(\sqrt{\frac{(\nu_{lk}^E)^2}{(d_{ilk} - \tilde{\mathbf{X}}_i^T \tilde{\mathbf{B}}_{lk}^* - \mathbf{Z}_i^T \mathbf{U}_{lk}^*)^2}}, (\nu_{lk}^E)^2 \right),$$

where $\tilde{\mathbf{X}}_i$ is the i th row of $\tilde{\mathbf{X}}$ and \mathbf{Z}_i is the i th row of \mathbf{Z} . Similarly,

$$(\phi_{mlk}^{-1} | U_{mlk}^*, (\nu_{lk}^U)^2) \sim \text{Inv-Gauss} \left(\sqrt{\frac{(\nu_{lk}^U)^2}{(U_{mlk}^*)^2}}, (\nu_{lk}^U)^2 \right),$$

where U_{mlk}^* is the m th component of \mathbf{U}_{lk}^* . To update $\{\psi_{alk}\}$, we have two cases. Conditional on $\gamma_{alk} = 1$, $(\psi_{alk}^{-1} | \tilde{B}_{alk}^*, \gamma_{alk} = 1, (\nu_{aj}^G)^2) \sim \text{Inv-Gauss}(\sqrt{(\nu_{aj}^G)^2 / (\tilde{B}_{alk}^*)^2}, (\nu_{aj}^G)^2)$. Conditional

on $\gamma_{alk} = 0$, we update ψ_{alk} from the prior: $(\psi_{alk}|\tilde{B}_{alk}^*, (\nu_{aj}^G)^2, \gamma_{alk} = 0) \sim f(\psi_{alk}|(\nu_{aj}^G)^2)$, where $f(\psi_{alk}|(\nu_{aj}^G)^2)$ is the density of $\text{Exponential}((\nu_{aj}^G)^2/2)$.

[**Step 4.**] Update parameters $(\nu^2|\{\theta_{lk}\}, \{\gamma_{alk}\})$. In particular, $(\nu_{lk}^E)^2|(\lambda_{ilk})_i \sim \text{Gamma}(N + a^E, 1/2 \sum_i \lambda_{ilk} + b^E)$, $(\nu_{lk}^U)^2|(\phi_{mlk})_m \sim \text{Gamma}(M + a^U, 1/2 \sum_m \phi_{mlk} + b^U)$, where M is the total number of random effects. Finally, we have $(\nu_{aj}^G)^2|(\psi_{alk})_a \sim \text{Gamma}(K_j + a^G, 1/2 \sum_k \psi_{alk} + b^G)$, where K_j is the number of wavelet coefficients in scale j .

[**Step 5.**] Update $(\pi_{aj}|\{\gamma_{alk}\}_k) \sim \text{Beta}(\sum_k \gamma_{alk} + a^\pi, K_j - \sum_k \gamma_{alk} + b^\pi)$.

Repeat Steps 1-5 until reaching a pre-specified maximum number of iterations.

3 Integrating Out Random Effects in the Predictive Likelihood

The model selection and discriminant analysis both require calculating the posterior predictive likelihood for data from a new footprint, i.e., computing $f(\mathbf{D}^s|\mathbf{V}^s, \mathbf{X}^s, \Theta^{(g)})$ by plugging in posterior samples of parameters (denoted by $\Theta^{(g)}$, g is the index for posterior samples). Since the random effect for a new footprint cannot be estimated during the training procedure, we need to integrate out the random effect \mathbf{U}_{lk}^* from model (1) when computing the posterior predictive likelihood. We now discuss details of this integration.

We can write model (1) for w echo envelopes from a new footprint as

$$\mathbf{d}_{lk}^s = \tilde{\mathbf{X}}^s \tilde{\mathbf{B}}_{lk}^* + u_{lk}^s \mathbf{1}_w + \mathbf{E}_{lk}^s, \quad (2)$$

where \mathbf{d}_{lk}^s , $\mathbf{1}_w$ and \mathbf{E}_{lk}^s are vectors of length w , and u_{lk}^s is the unknown random effect for the new footprint.

In the Gfmm case, both u_{lk}^s and \mathbf{E}_{lk}^s are Gaussian, i.e., $u_{lk}^s \sim N(0, q_{lk})$, $\mathbf{E}_{lk}^s \sim N(0, s_{lk} \mathbf{I}_w)$. Therefore the likelihood for \mathbf{d}_{lk}^s is again multivariate normal, i.e., $\mathbf{d}_{lk}^s \sim N(\tilde{\mathbf{X}}^s \tilde{\mathbf{B}}_{lk}^*, \Sigma_{lk}^s)$, where $\Sigma_{lk}^s = q_{lk} \mathbf{J} + s_{lk} \mathbf{I}_w$. Here \mathbf{J} is a w by w matrix of ones, and \mathbf{I}_w is a w dimensional identity matrix.

In the Rfmm case, since both u_{lk}^s and components in \mathbf{E}_{lk}^s follow scaled mixture of normal distributions (in particular, DE distributions), the analytical solution for the integrated predictive likelihood is not easy to obtain. We therefore approximate the predictive

likelihood of \mathbf{d}_{lk}^s using the trapezoidal rule. We found that the trapezoidal rule is more accurate and numerically efficient than Monte Carlo approximation; the latter requires a lot of (more than 1000) Monte Carlo samples. In particular, since all components of \mathbf{E}_{lk}^s are independent $DE(0, 1/\nu_{lk}^E)$, we have

$$f(\mathbf{d}_{lk}^s | \tilde{\mathbf{X}}^s, \tilde{\mathbf{B}}_{lk}^*, \boldsymbol{\Theta}_{lk}) = \int f(\mathbf{d}_{lk}^s | \tilde{\mathbf{X}}^s, \tilde{\mathbf{B}}_{lk}^*, u_{lk}^s, \nu_{lk}^E) f(u_{lk}^s | \nu_{lk}^U) du_{lk}^s. \quad (3)$$

Here $f(\mathbf{d}_{lk}^s | \tilde{\mathbf{X}}^s, \tilde{\mathbf{B}}_{lk}^*, u_{lk}^s, \nu_{lk}^E) = \prod_{i=1}^w DE(d_{ilk}^s - \tilde{\mathbf{X}}_i^s \tilde{\mathbf{B}}_{lk}^* - u_{lk}^s, 1/\nu_{lk}^E)$, where $\tilde{\mathbf{X}}_i^s$ is the i th row of the design matrix $\tilde{\mathbf{X}}$, $\mathbf{d}_{lk}^s = (d_{1lk}^s, \dots, d_{wlk}^s)^T$, and $\boldsymbol{\Theta}_{lk}$ contains all parameters except u_{lk}^s . The integration in (3) is then approximated by the trapezoidal rule. In particular, we first partition the range of u_{lk}^s into a dense grid, and then evaluate $f(\mathbf{d}_{lk}^s | \tilde{\mathbf{X}}^s, \tilde{\mathbf{B}}_{lk}^*, u_{lk}^s, \nu_{lk}^E)$ at each grid point. The integration in (3) can finally be approximated numerically using the trapezoidal rule. In this calculation, the range of u_{lk}^s is determined using the posterior samples of u_{lk}^s in the training procedure.

4 More Results in the Sensitivity Analysis

In this section, we display additional results in the sensitivity analysis. In Figure 2, we plot the estimates of the three terrain mean effects under the five different setups. In Figures 3 – 5, we plot the estimates of the three pairwise contrast effects between terrains under the five different setups, where the yellow color indicates significantly nonzero regions detected by SimBaS. In these figures, $|T|$ denotes the number of grid points, *db4* and *db10* denotes the wavelets basis Daubechies 4 and Daubechies 10 respectively.

From Figure 2 - 5, we observe similar patterns to the channel contrast effect shown in the main text. In particular, the estimates are similar across cases (1) - (4). The estimates for case (5) show evidently more fluctuations on both the mean and the 95% simultaneous credible band. The differences between the flagged regions mostly appear in places where the SimBaS is close to 0.05, which are weakly significant regions.

In Figure 6 - 8, we display the wavelet basis functions corresponding to the original analysis and the sensitivity analyses in cases (4) - (5). In these figures, $l = 0$ indicates the scaling function (father wavelets), and $l > 1$ indicates different resolution levels of the (mother) wavelet functions. From these plots, we see that higher order of the Daubechies

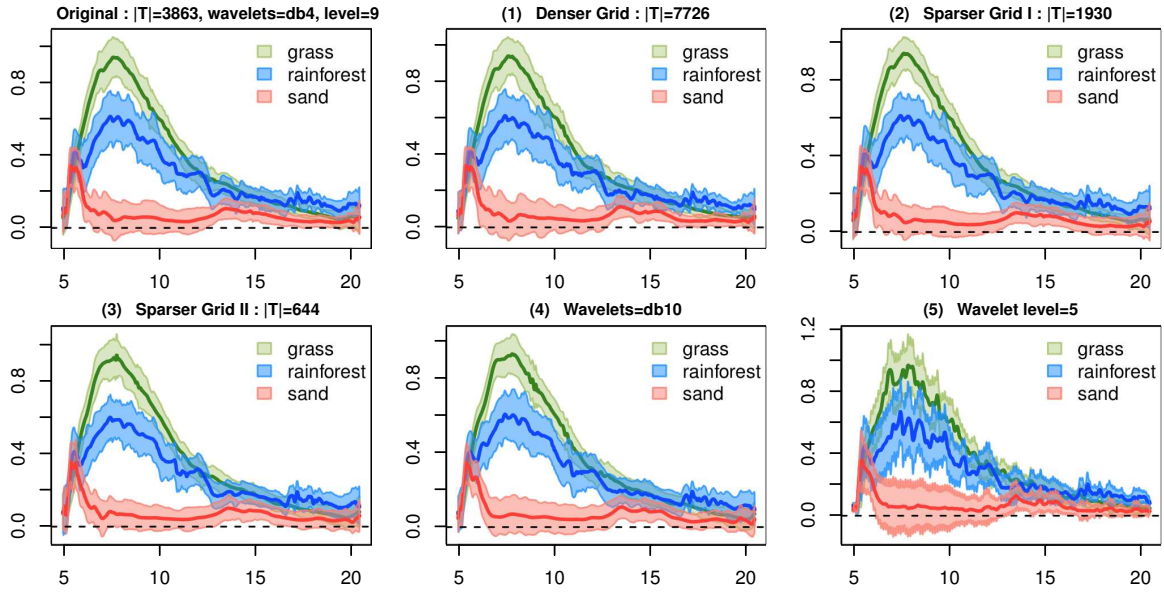


Figure 2: The estimated terrain mean effects in the five sensitivity analyses.

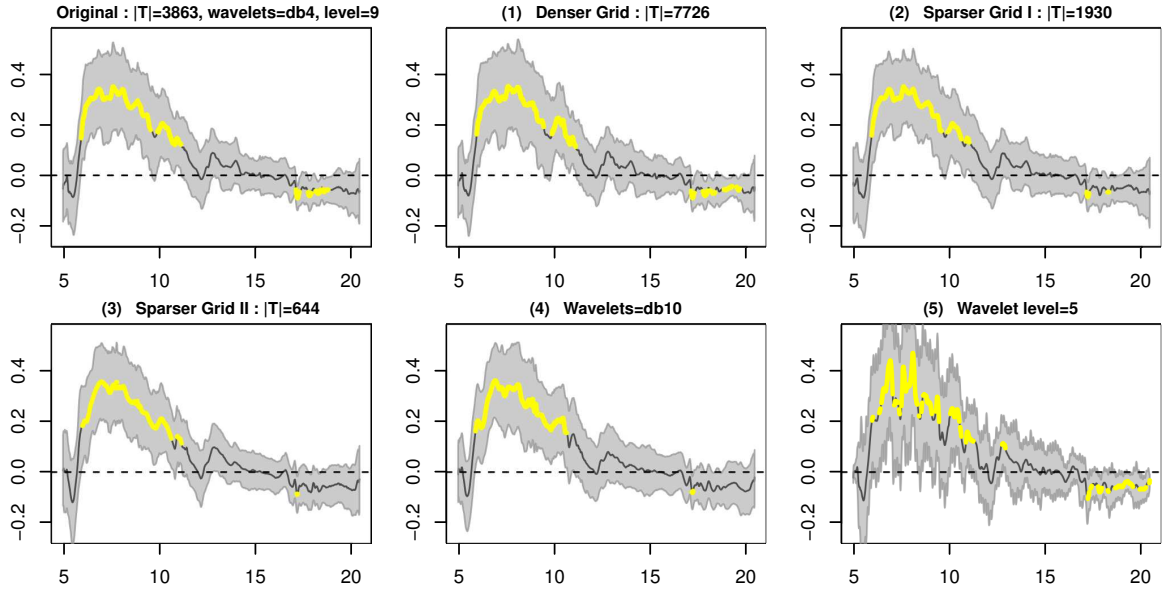


Figure 3: The estimated contrast effect grass - rainforest and the flagged nonzero regions (in yellow color) in the five sensitivity analyses.

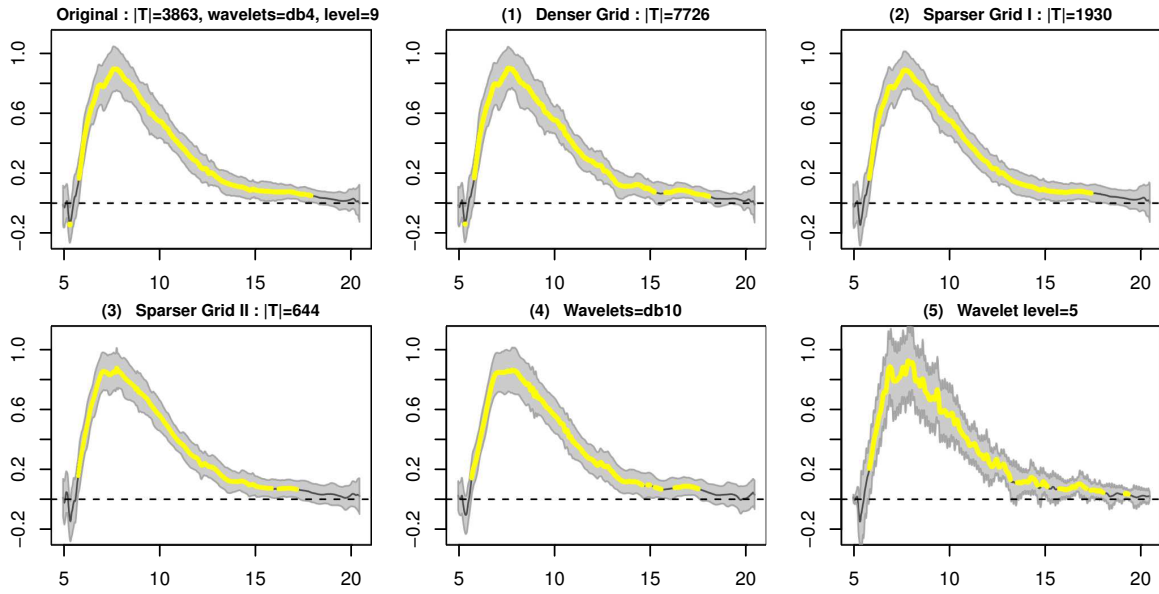


Figure 4: The estimated contrast effect grass - sand and the flagged nonzero regions (in yellow color) in the five sensitivity analyses.

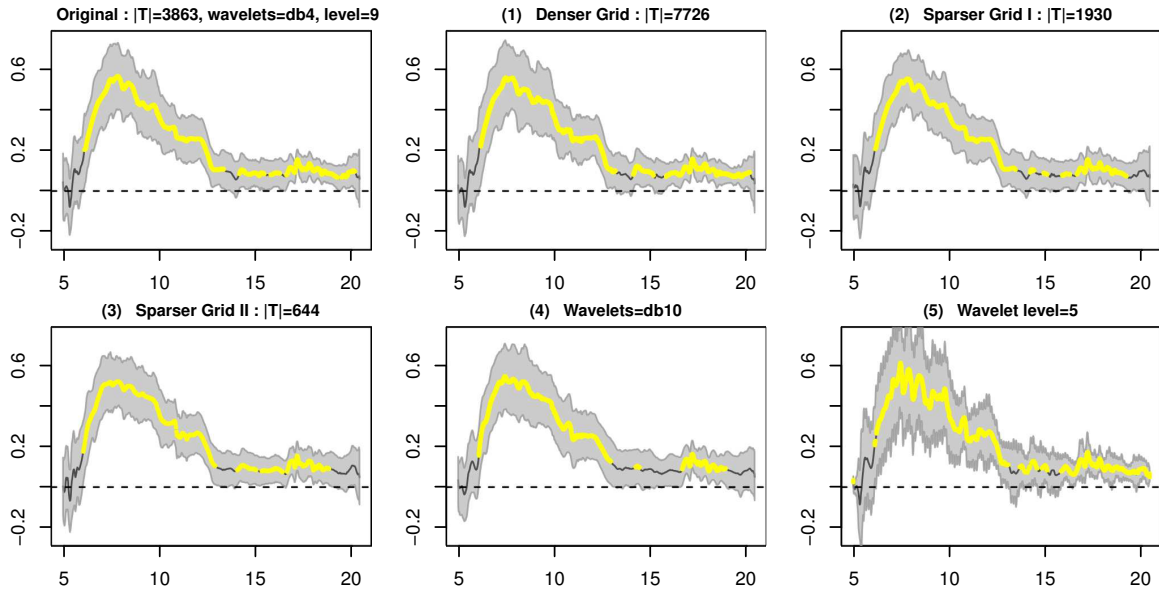


Figure 5: The estimated contrast effect rainforest - sand and the flagged nonzero regions (in yellow color) in the five sensitivity analyses.

wavelets (e.g., *db10*) demonstrates wider supports in lower resolution levels (e.g., levels $l = 0$ to $l = 3$) as well as higher degree of smoothness and fluctuations. Furthermore, in the low wavelet level setup (e.g., number of levels = 5 in case 5), almost all basis functions have narrow supports. This provides a sensible interpretation to the more wiggling estimation results seen in case (5).

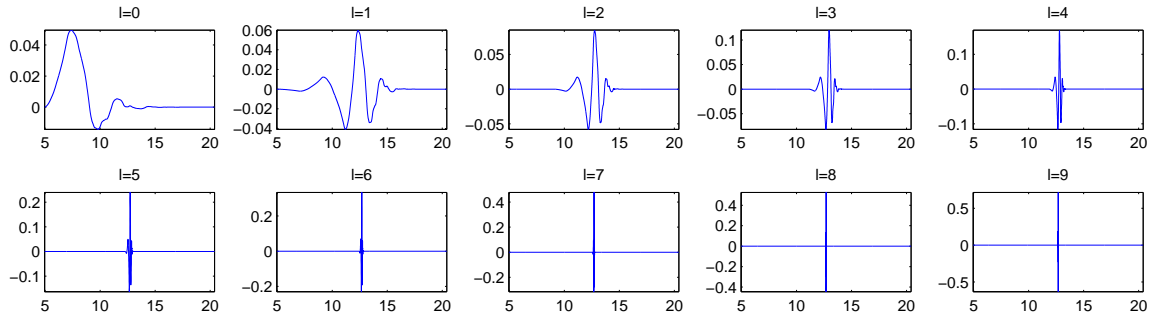


Figure 6: Basis functions in the original analysis, where we set wavelet='db4', number of levels=9.

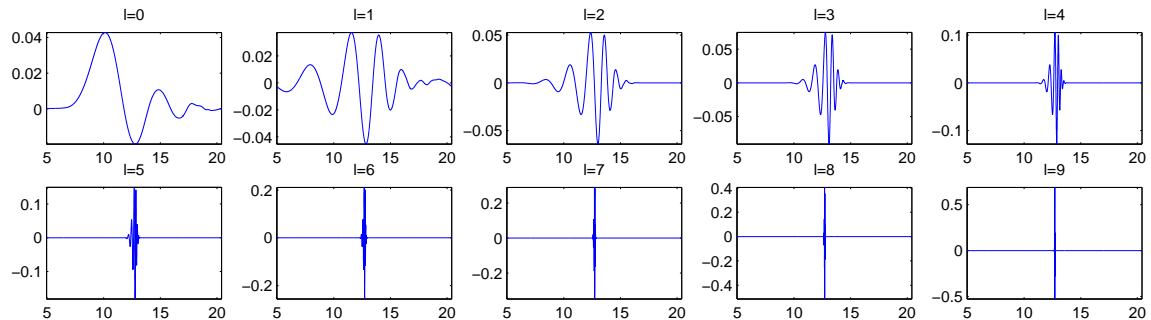


Figure 7: Basis functions in sensitivity analysis case (4), where wavelet='db10', number of levels=9.

References

Fletcher, N. H. (1992), *Acoustic Systems in Biology* Oxford University Press.

Searle, S. R., Cassela, G., and McCulloch, C. E. (1992), *Variance Components*, New York, USA: John Wiley & Sons, Inc.

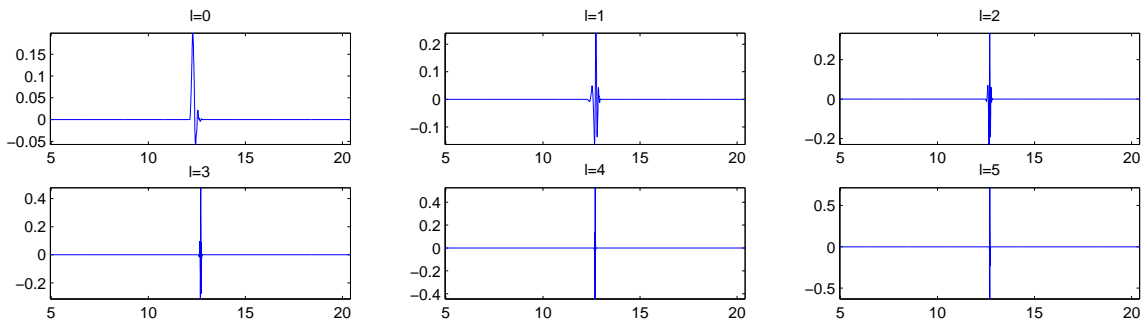


Figure 8: Basis functions in sensitivity analysis case (5), where wavelet='db4', number of levels=5.