1. Motivation: A multilevel CAR model for forest restoration

The Sabah biodiversity experiment

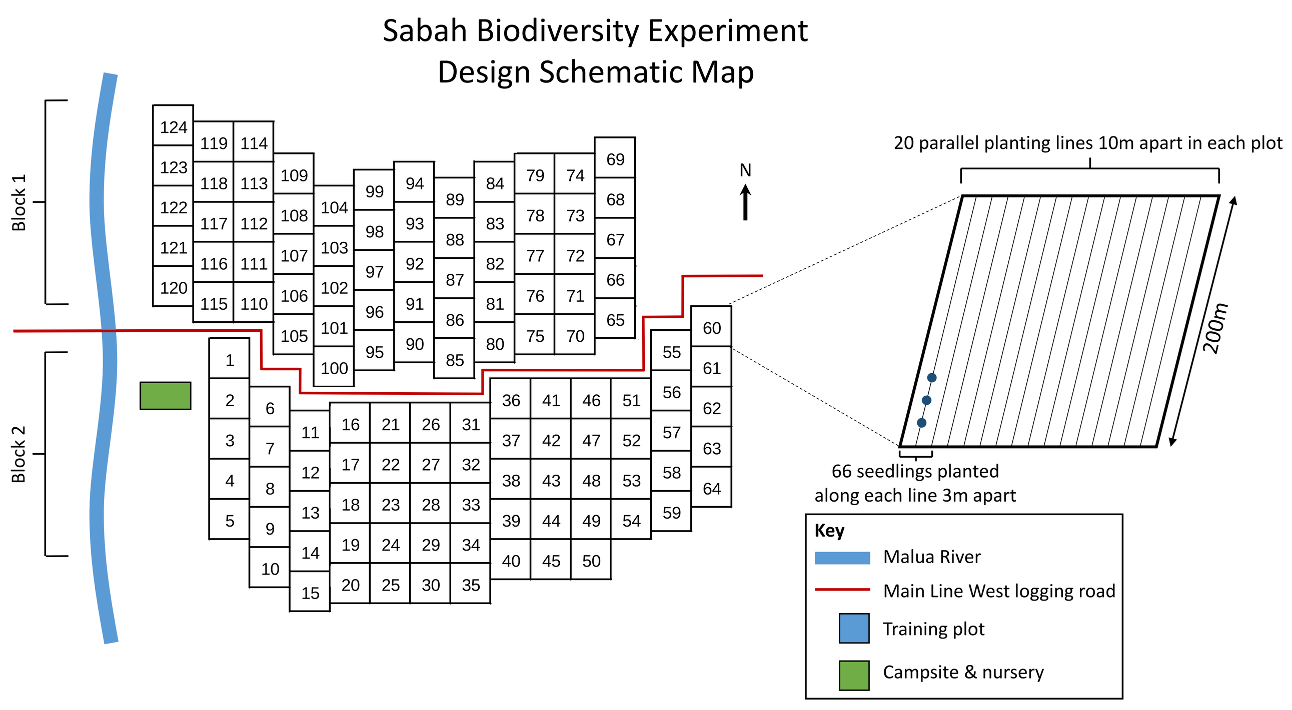
Measurements of more than 7000 trees in 124 study plots are collected over 10 years and researchers want to investigate:

- tree growth rate $\sim$ biodiversity + species?
- any spatial correlation?
  - tree level: $\varepsilon_{ij}|\delta(\varepsilon_{ij}) \sim \mathcal{N}(\rho_e \sum \delta(\varepsilon_{ij}), \sigma_e^2)$
  - plot level: $b_j|\delta(b_j) \sim \mathcal{N}(\rho_p \sum \delta(b_j), \sigma_p^2)$

A multilevel CAR model

$$\boldsymbol{y} = \boldsymbol{X\beta} + Z\boldsymbol{b} + \boldsymbol{\varepsilon};\ \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \Sigma_b),\ \boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \Sigma_\varepsilon)$$

$\Sigma_b = (I - \rho_p W_p)^{-1} \otimes diag(\Sigma_{b_j}),\ \Sigma_{b_j} = \sigma_b^2 I_{n_j}; \Sigma_\varepsilon = \sigma^2(I - \rho_e W_e)^{-1}$



(?)

2. The likelihood evaluation

The log-likelihood function $\ell(\theta; y)$ is

$$\propto \frac{\log(|\Sigma_\varepsilon||\Sigma_b|)}{-2} + \log\int \exp\left\{\frac{\tilde{\boldsymbol{e}}^T\Sigma_\varepsilon^{-1}\tilde{\boldsymbol{e}}^T + \boldsymbol{b}^T\Sigma_b^{-1}\boldsymbol{b}}{-2}\right\} \mathrm{d}\boldsymbol{b} \propto -\frac{1}{2}\left(\log|\Sigma_y| + \boldsymbol{e}^T\Sigma_y^{-1}\boldsymbol{e}^T\right)$$

$\Sigma_y = Z^T\Sigma_b Z + \Sigma_\varepsilon,\ \tilde{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{X\beta} - Z\boldsymbol{b},\ \boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X\beta}$

$|\Sigma_y|$ and Σ_y^{-1} takes $\mathcal{O}(n^3)$ flops but Σ_ε^{-1} and Σ_b^{-1} are *sparse*!

3. The Monte Carlo likelihood

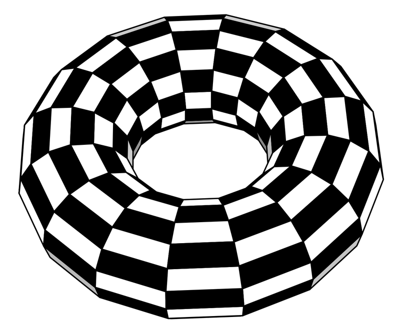
For the likelihood $L(\theta; y) = \int f_\theta(y, b)\,\mathrm{d}b$, generate s samples of $b^{*(i)} \sim f_\psi(b|Y = y)$. (?)

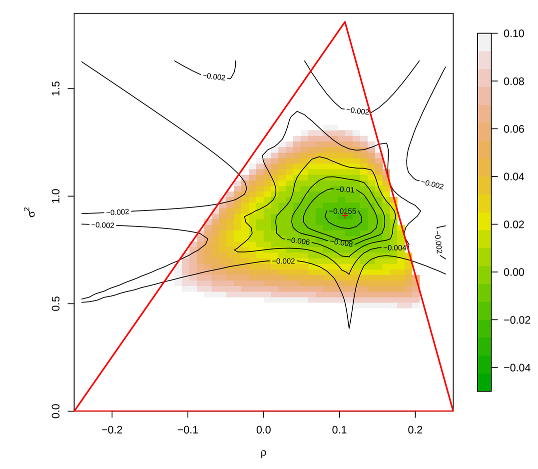
MC likeihood:

$\ell_\psi^s(\theta) = \log \frac{1}{s}\sum_i^s \frac{f_\theta(y, b^{*(i)})}{f_\psi(y, b^{*(i)})}$

MC MLE: $\hat{\theta}_\psi^s = \arg\max_{\theta\in\Theta} \hat{\ell}_\psi^s(\theta)$

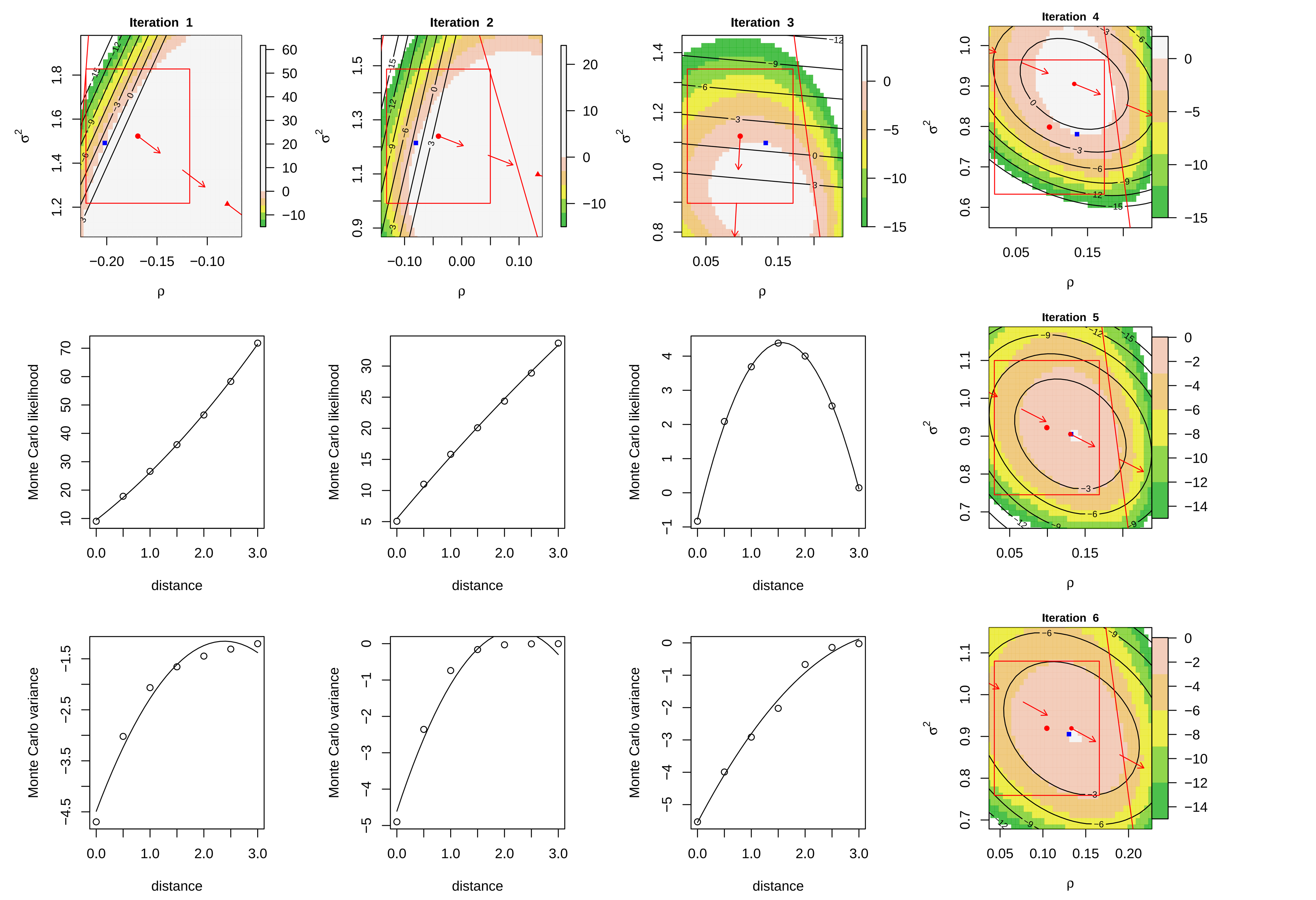
MC error: depends on $d(\theta, \psi)$ and scales as $1/\sqrt{s}$.





4. MC-MLE through RSM

Use evaluations at *design points* to fit a low-order polynomials approximation to the target function within the *design region*. Search for the maximum by exploring along the *steepest ascent path*. (?)



5. Conclusion

- simulation-based likelihood inference procedure with stable implementation and acceptable computation time
  - regardless of initial value
  - 1 CPU hour for the SBE data
- fully automated and easy adaptation to more complicated spatial models.
- available in the package **mclcar**.

References