# Fast and flexible gradient-based optimization of resistance surfaces

**Nate Pope, Penn State**

**IALE Workshop-4 2020**

# Outline

**1.** IBR as a nonlinear model: estimation and inference

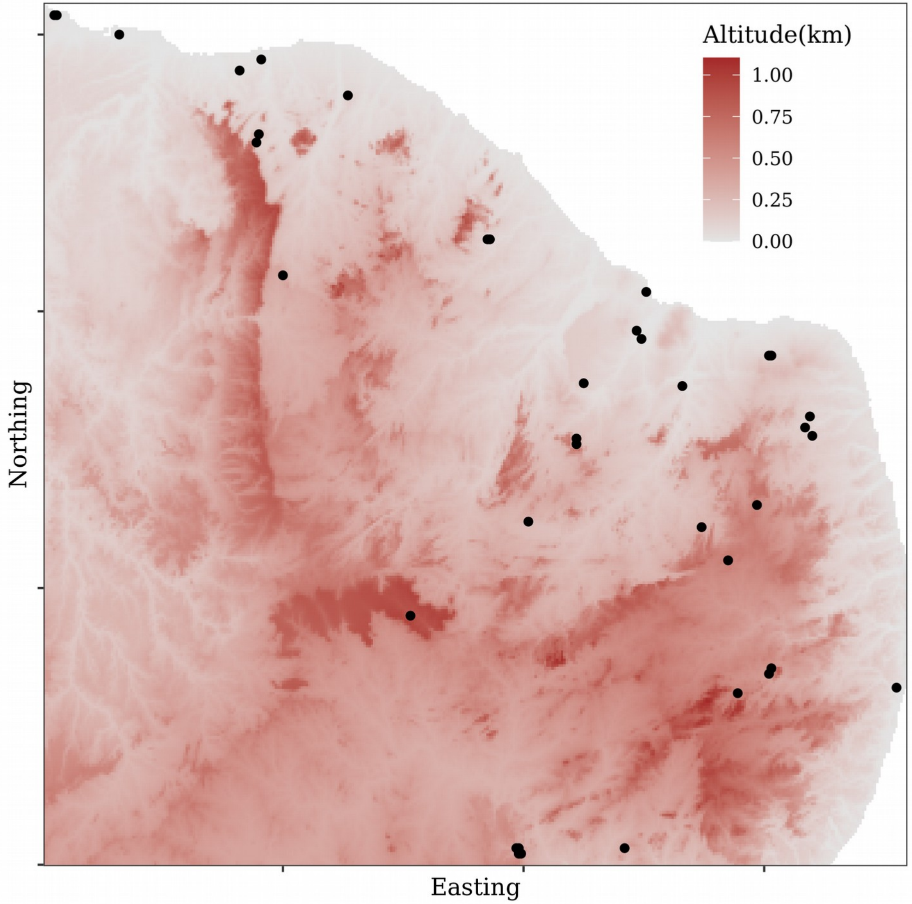
**2.** Accelerated optimization via fast calculation of gradient/Hessian

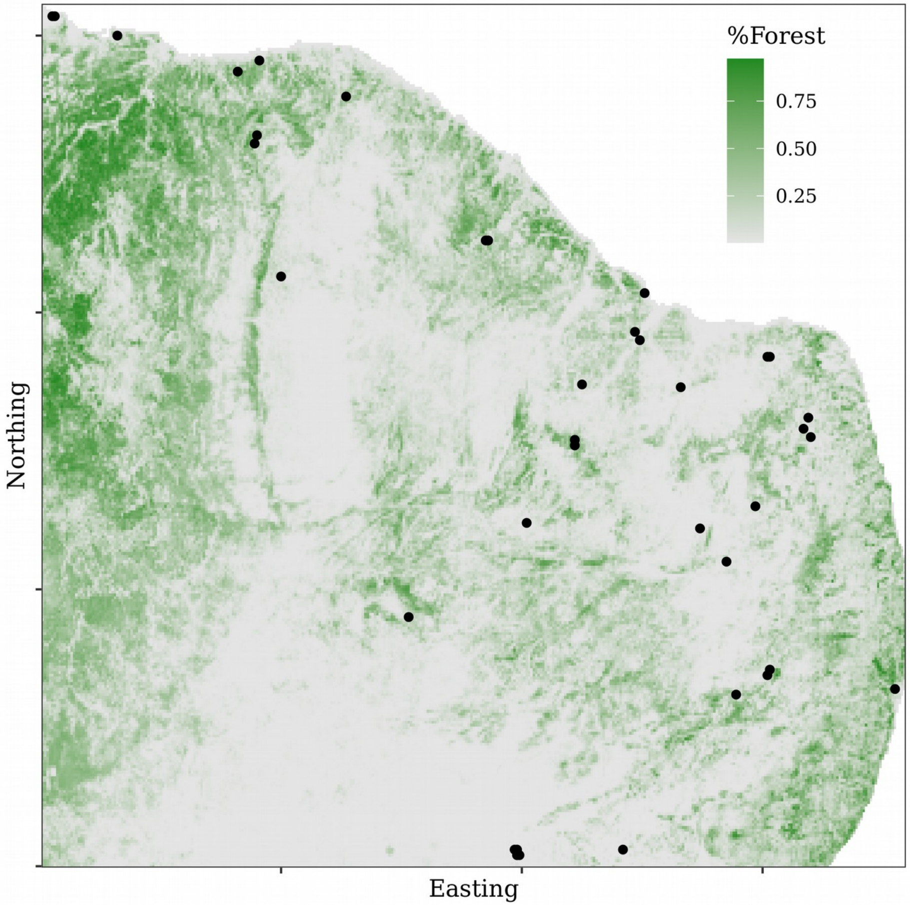
**3.** Influence/leverage diagnostics

# A motivating example

*Melipona subnitida*, “jandaira”

# A motivating example




Data from Jaffe et al 2019 Evol App; ~10K SNPs, ~40 spatial locations

# “Conductance model”

Mapping from
**spatial covariates** (multiple rasters)
→ **conductance** (single raster)

# "Conductance model"

e.g. "log-linear conductance"

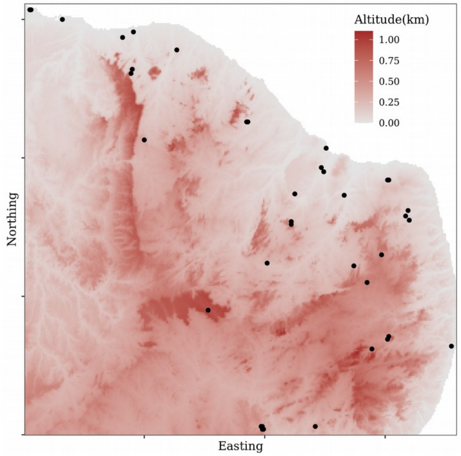
$$\boldsymbol{C_i} = \exp(\boldsymbol{x_i}'\boldsymbol{\theta})$$

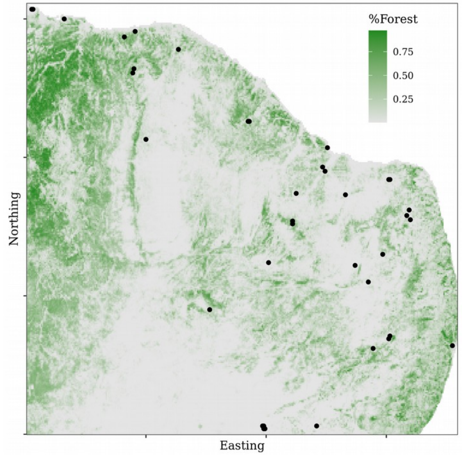
$\boldsymbol{C_i}$ is conductance in $\boldsymbol{i}$th cell

$\boldsymbol{x_i}$ is vector of spatial covariates

$\boldsymbol{\theta}$ is vector of **unknown** parameters

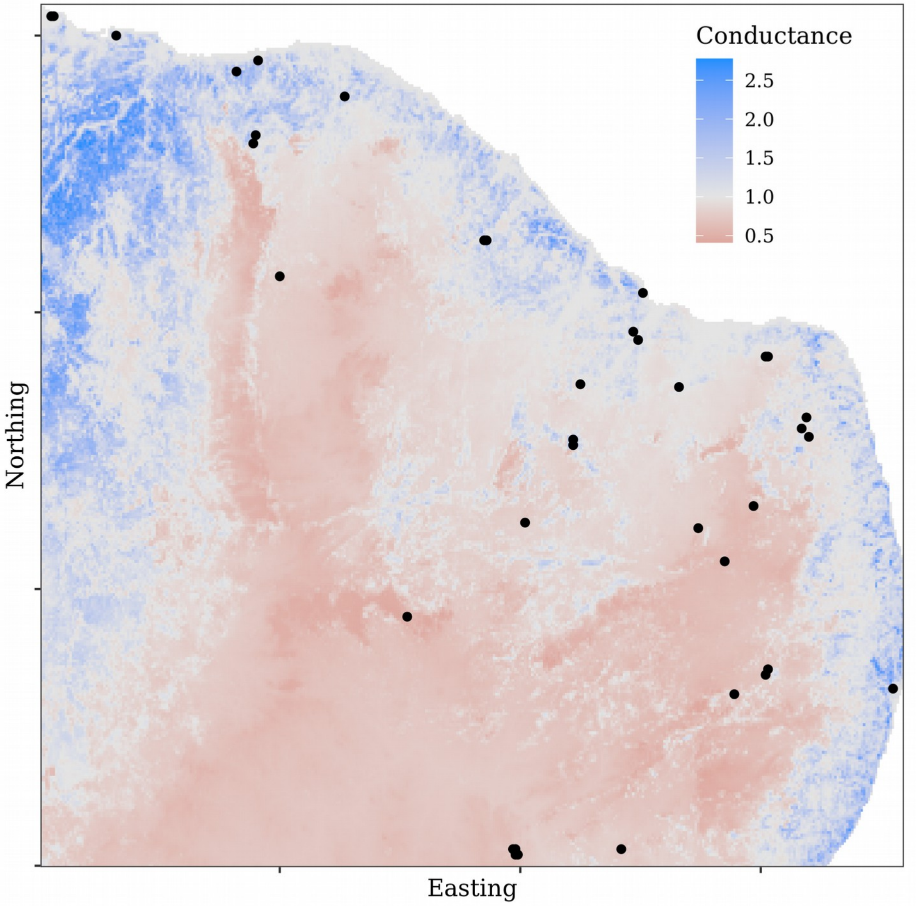
# "Conductance model"



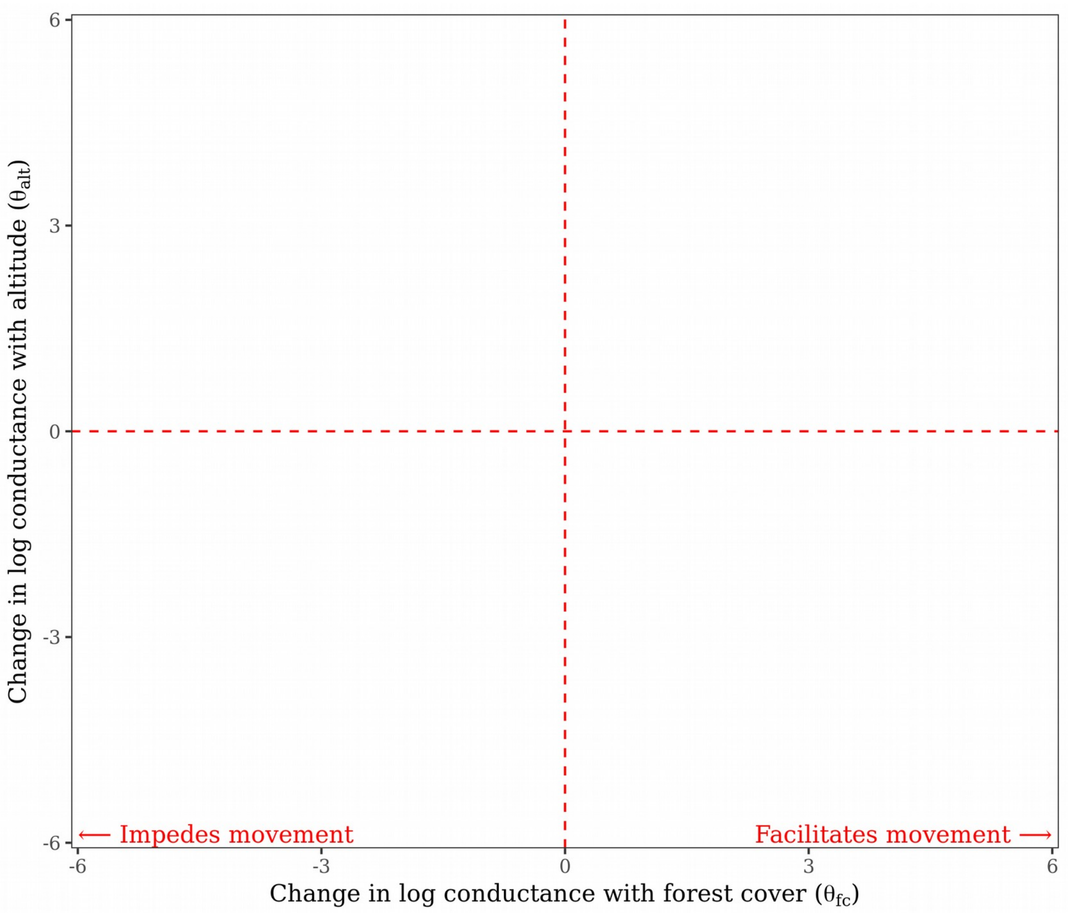




$$\boldsymbol{\theta}_{alt} = -0.89$$

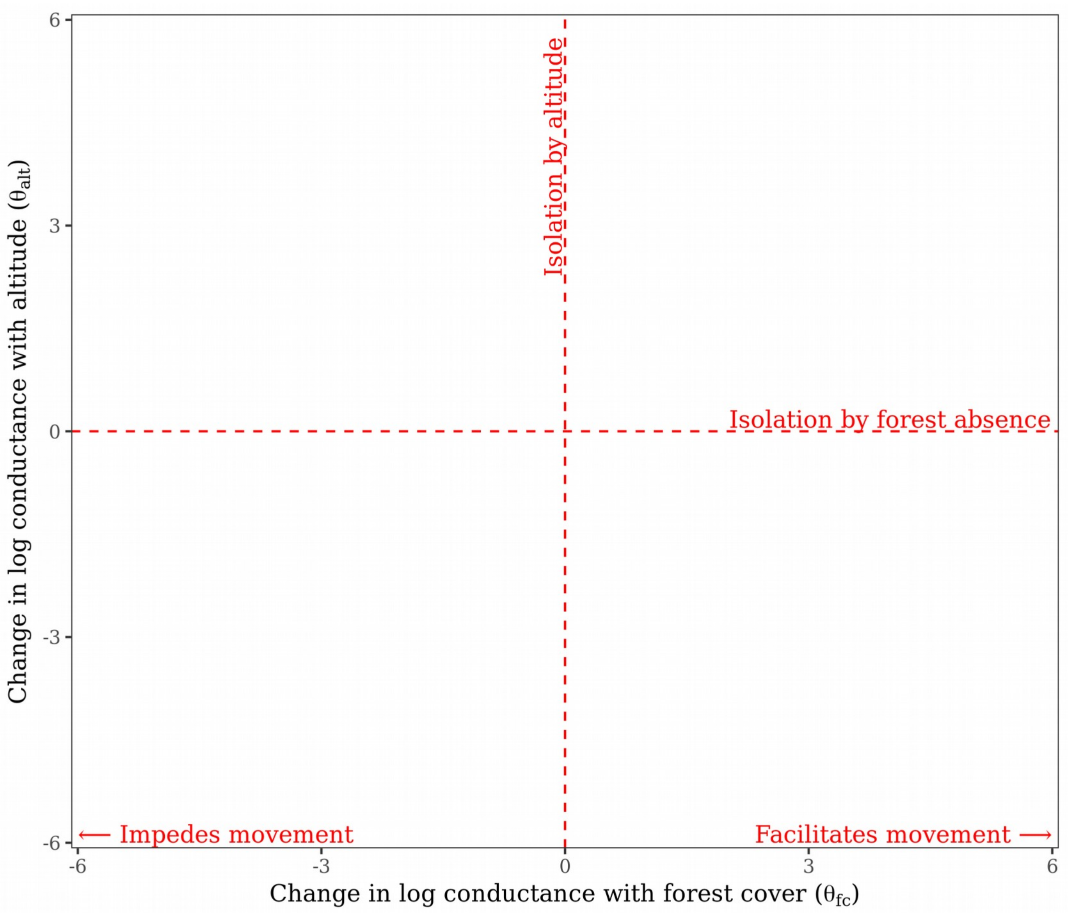
$$\boldsymbol{\theta}_{fc} = 1.13$$



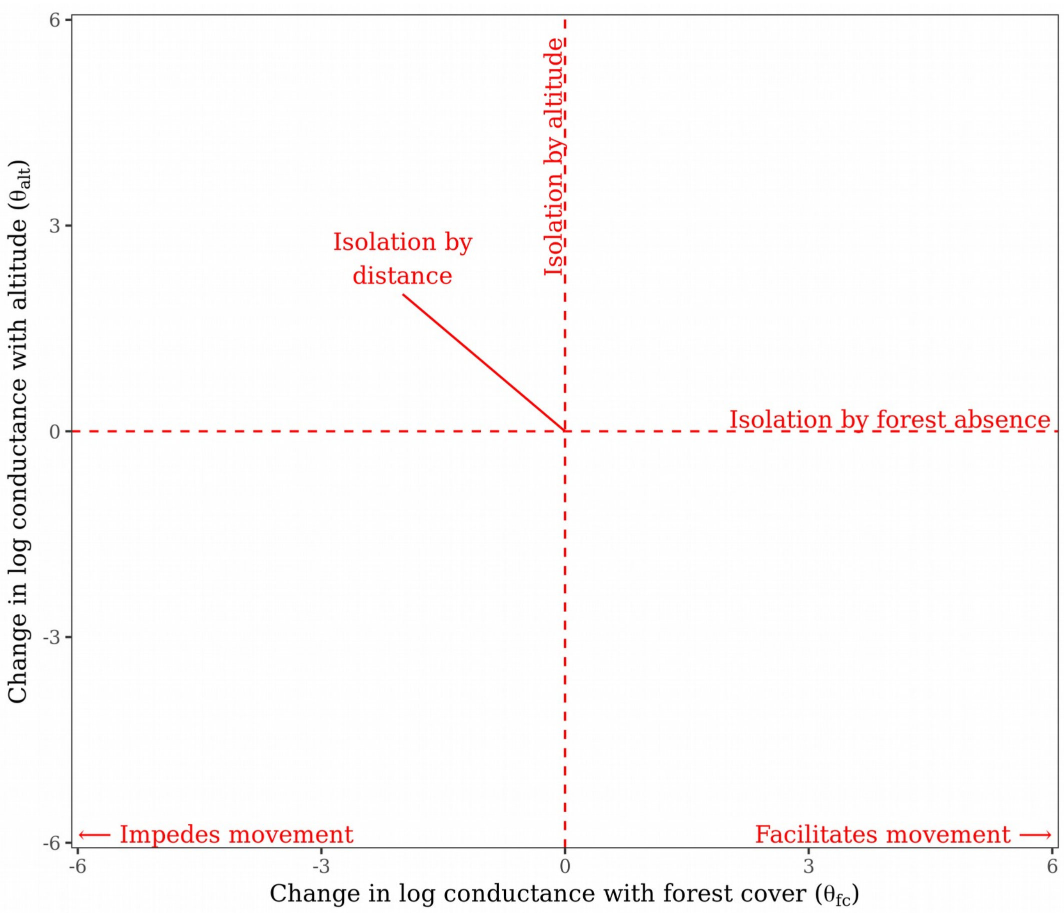
# Parameter space



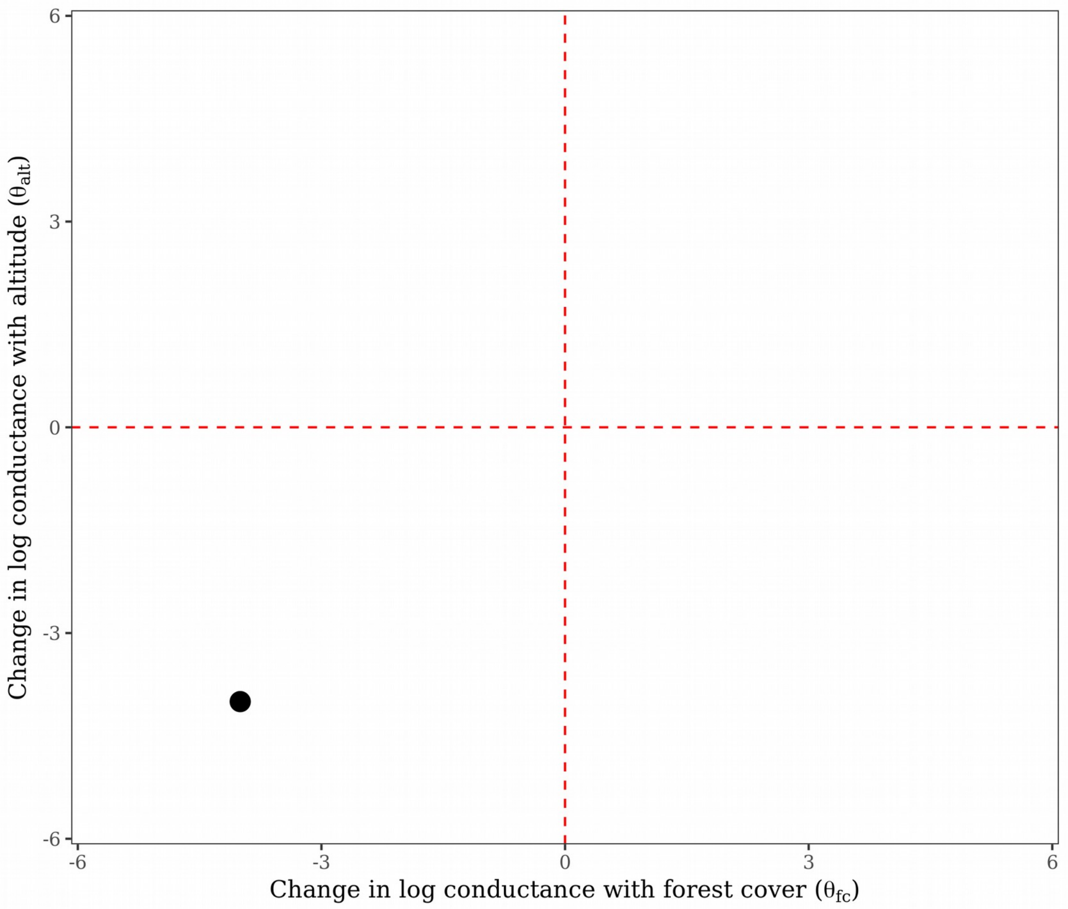
# Parameter space



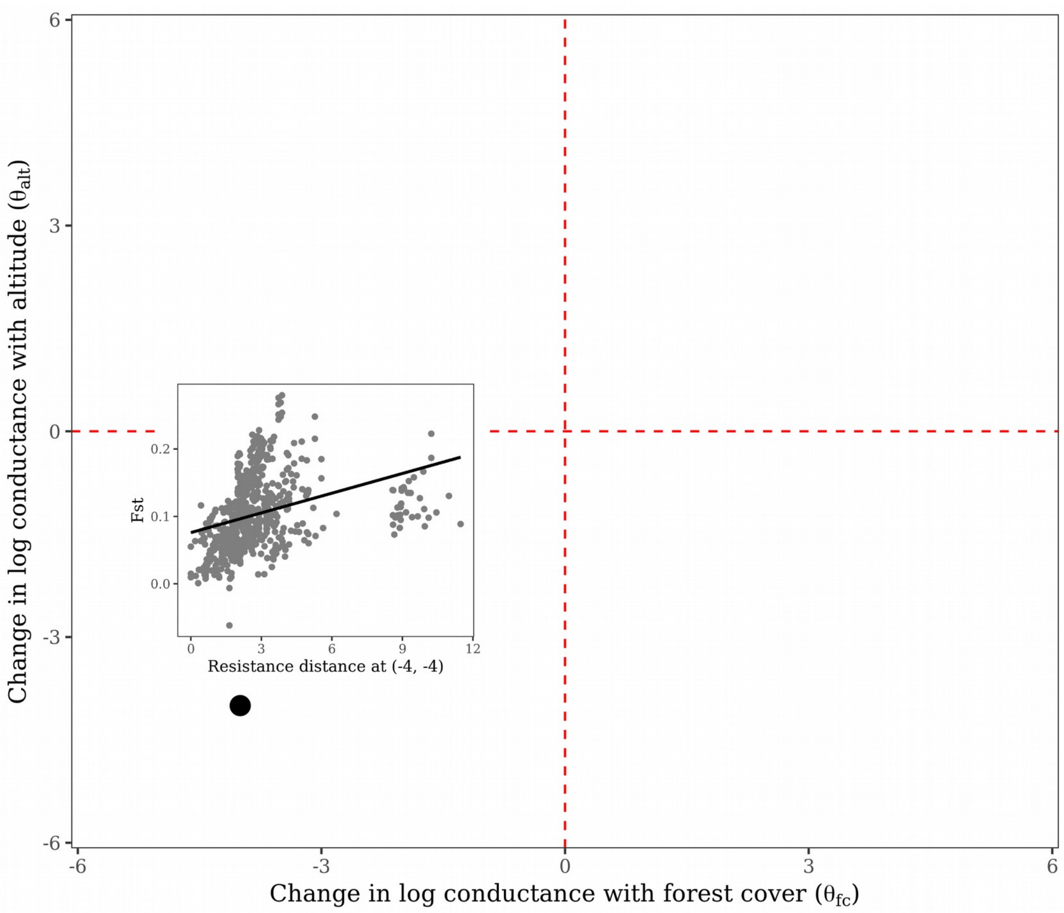
# Parameter space



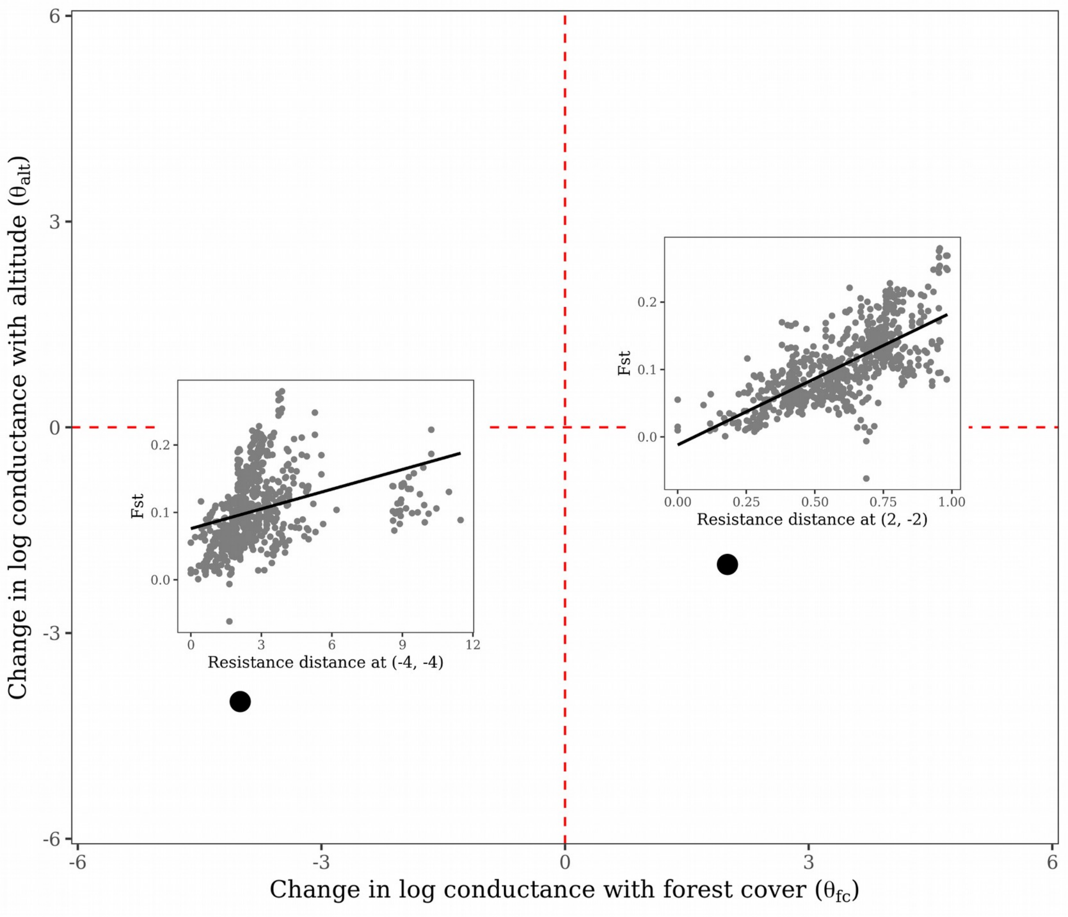
# Parameter space



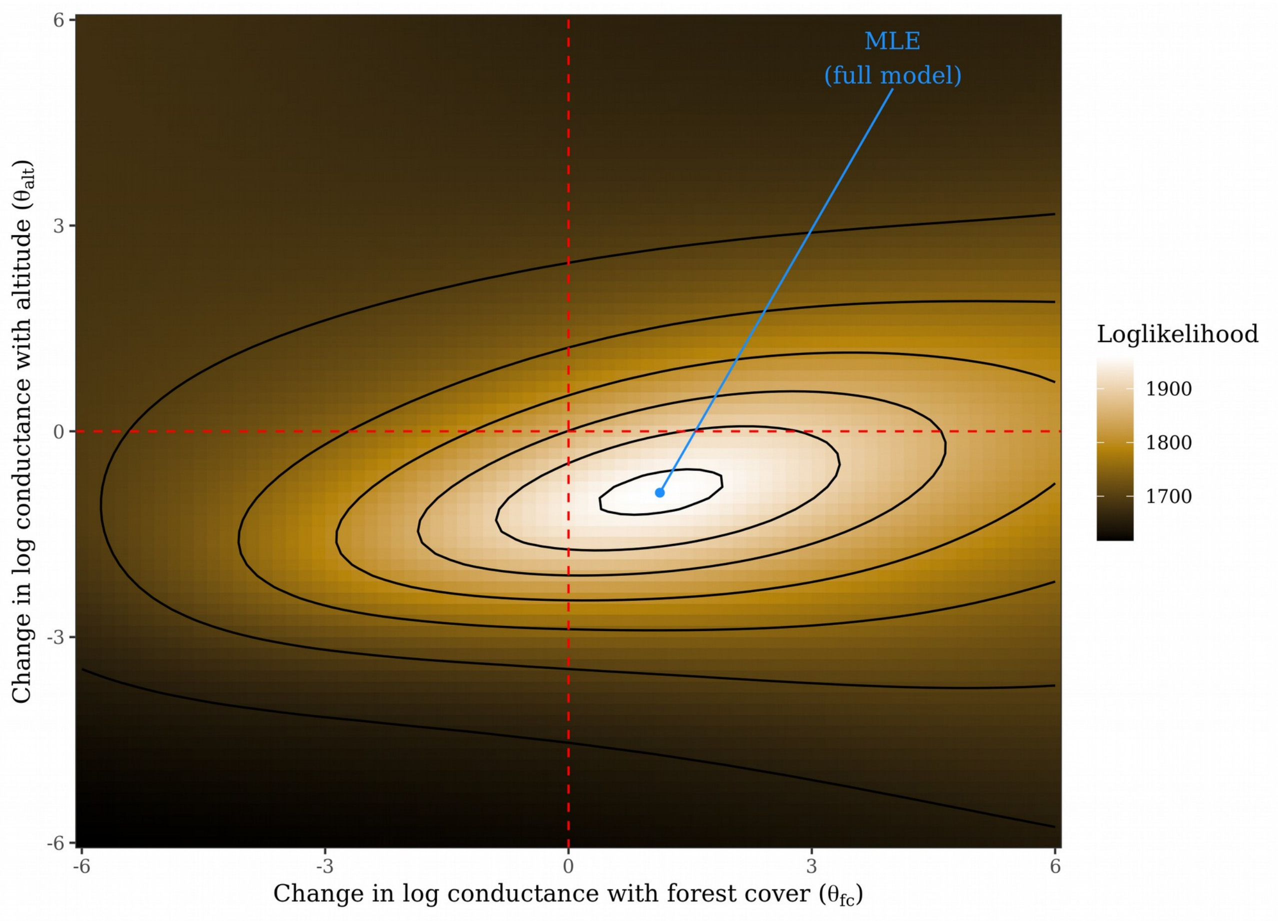
# “Measurement model”



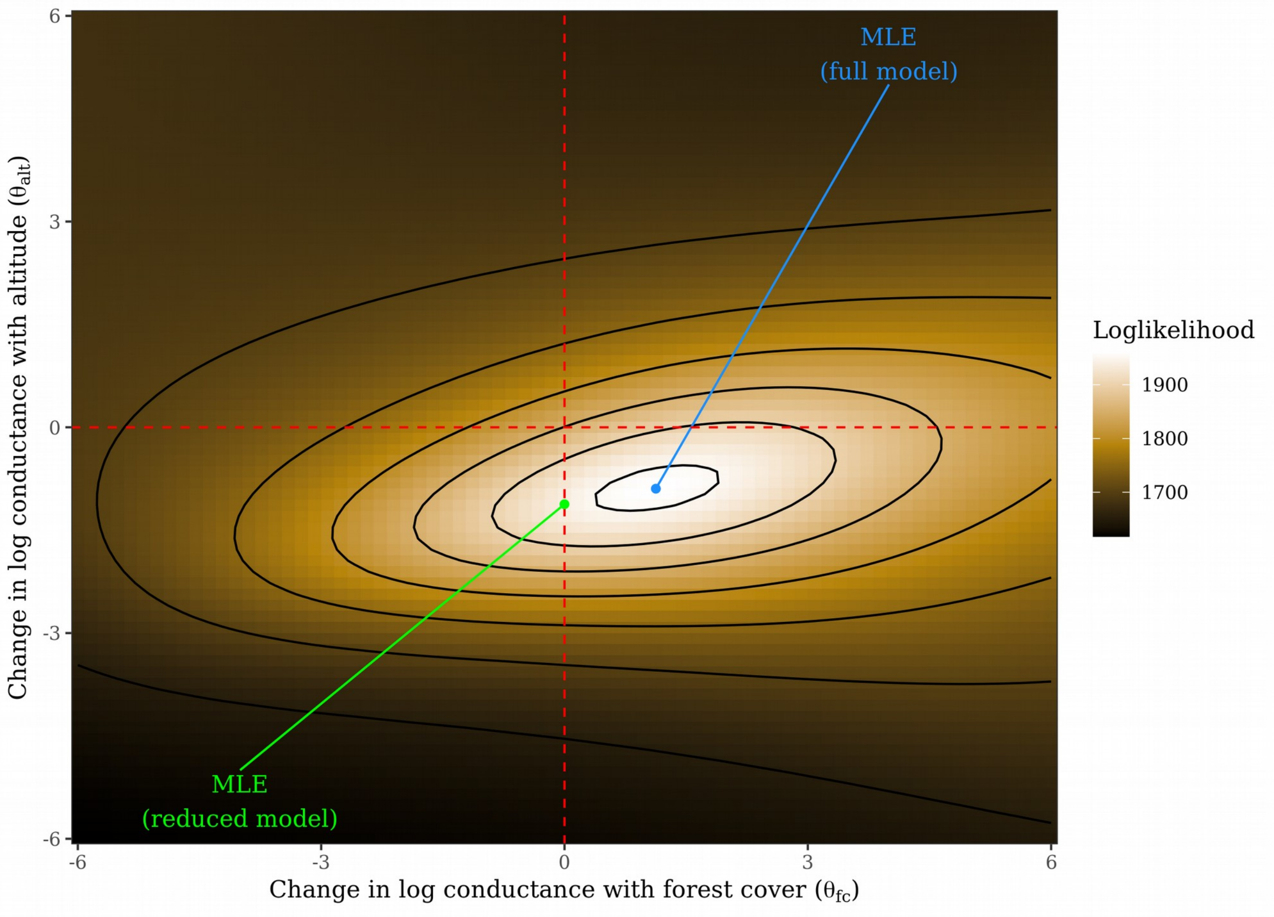
# “Measurement model”



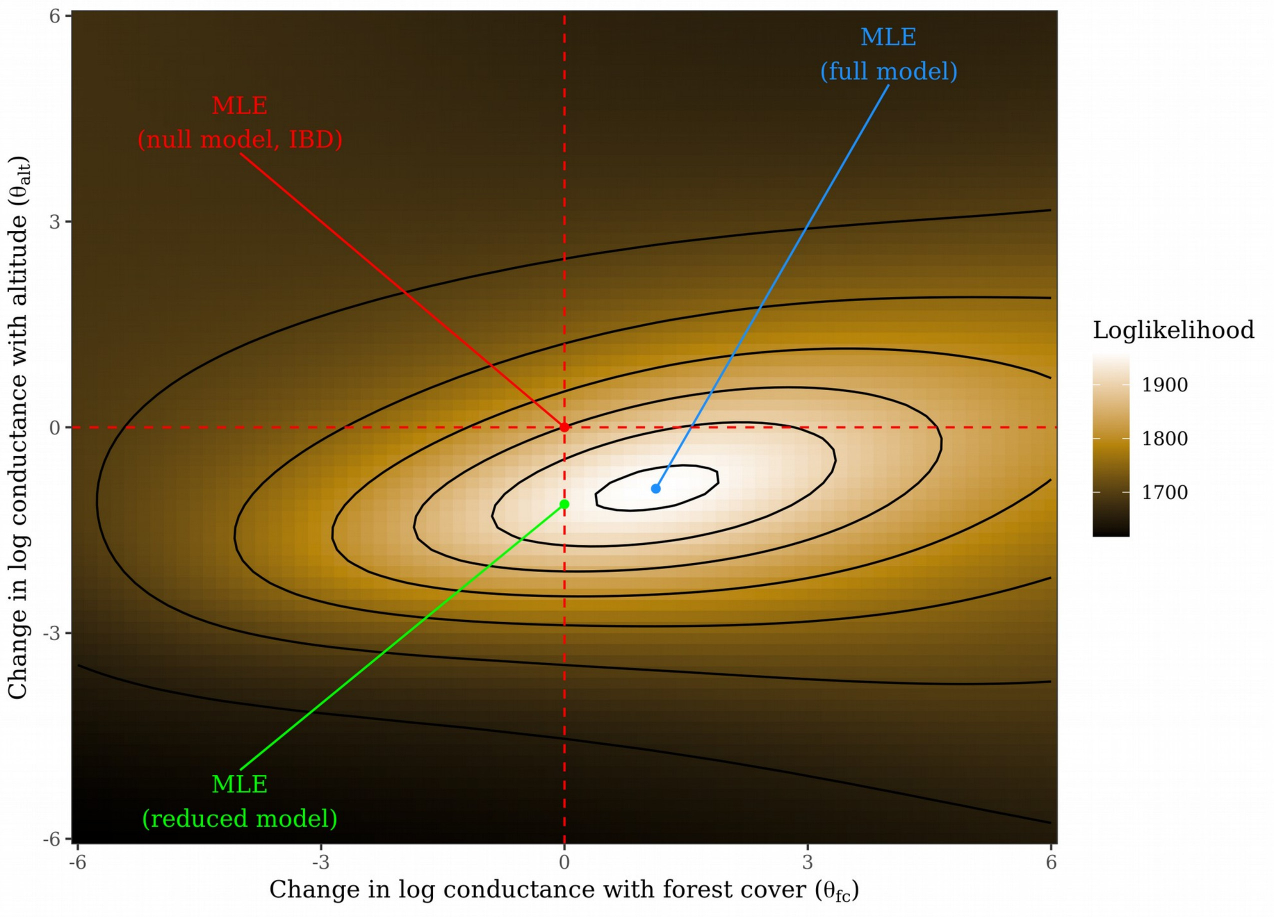
# Maximum likelihood



# Maximum likelihood



# Maximum likelihood



(this is just fitting a non-linear model by maximum likelihood)

# Inference

**Asymptotic std errors, likelihood ratios, etc.** Very fast, but sensitive to model mis-specification

**For example**, how is dependence among pairwise observations handled?

# Inference

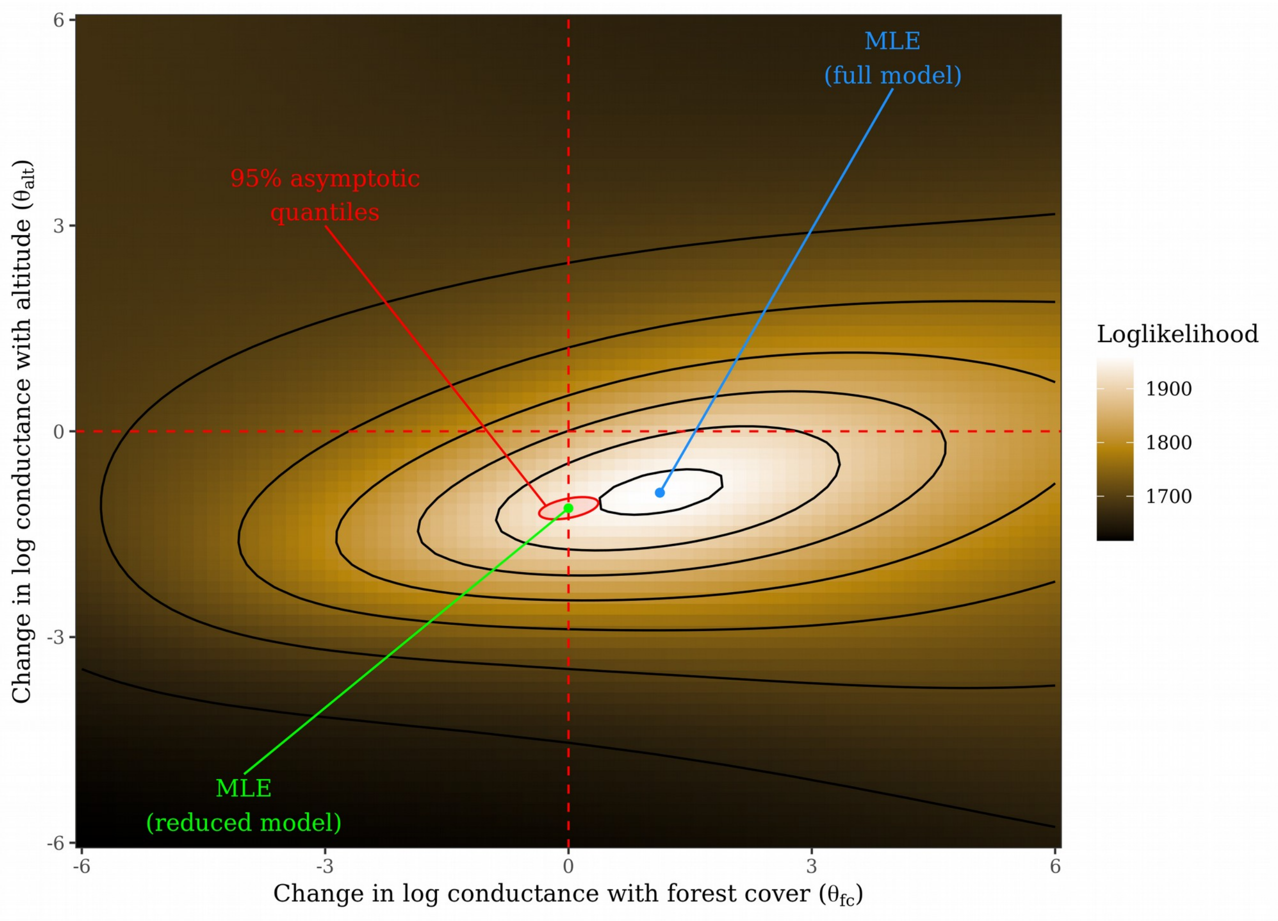
**Option 1.** Try to simulate data with similar dependence structure (e.g. via permutation)

# Inference

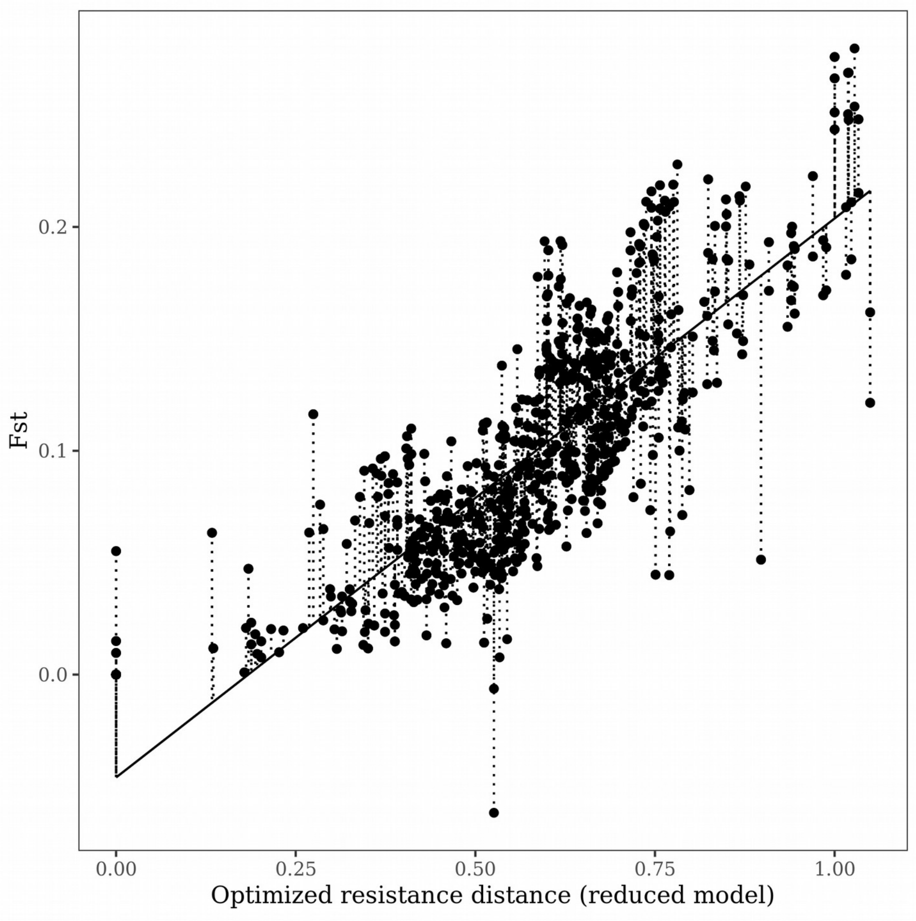
**Option 1.** Try to simulate data with similar dependence structure (e.g. via permutation)

**Option 2.** Try to model dependence structure

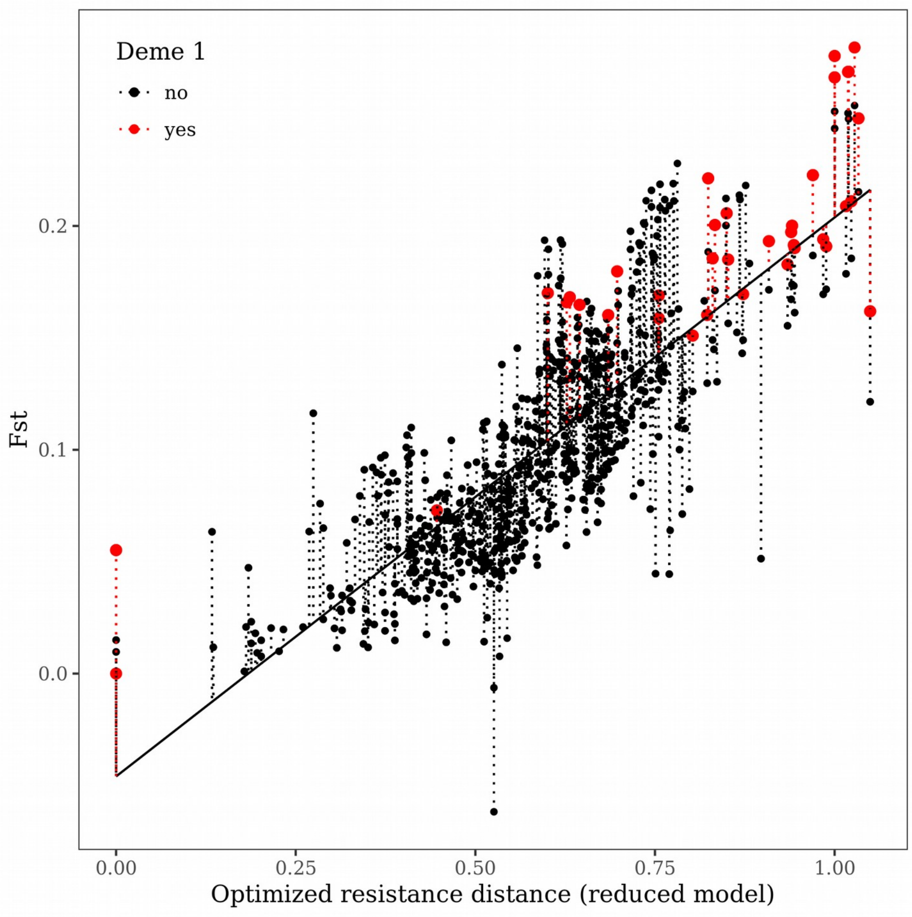
# Asymptotics (**not** modeling dependence)



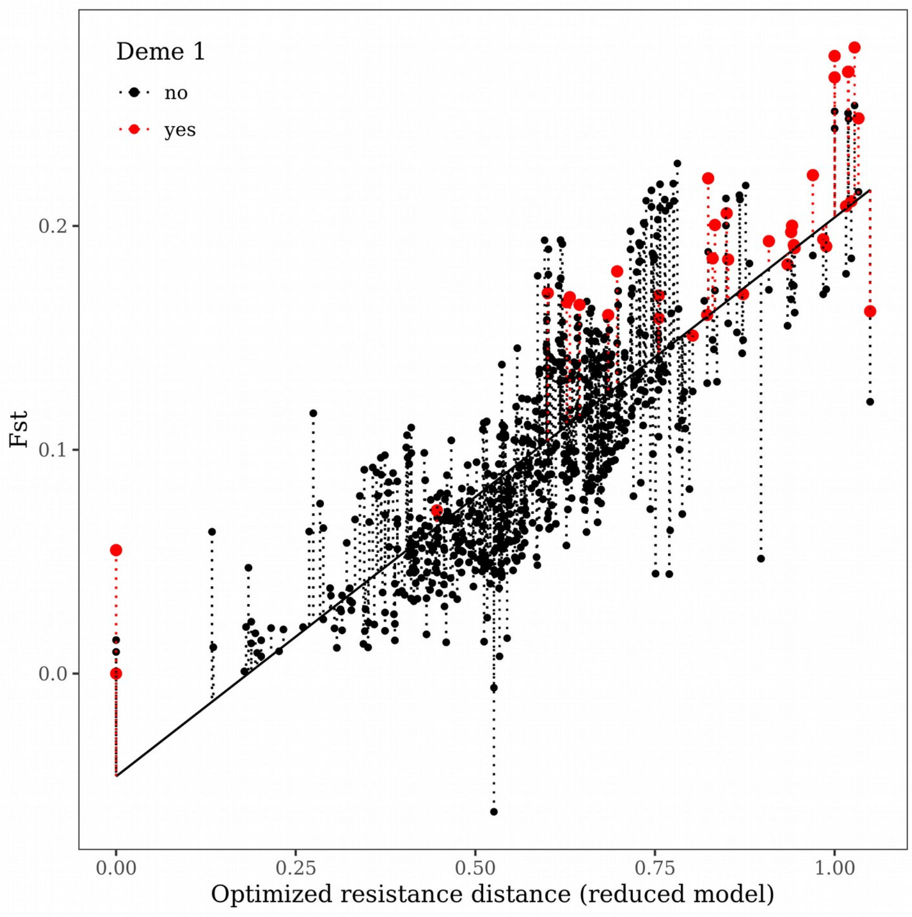
# Residual permutation tests

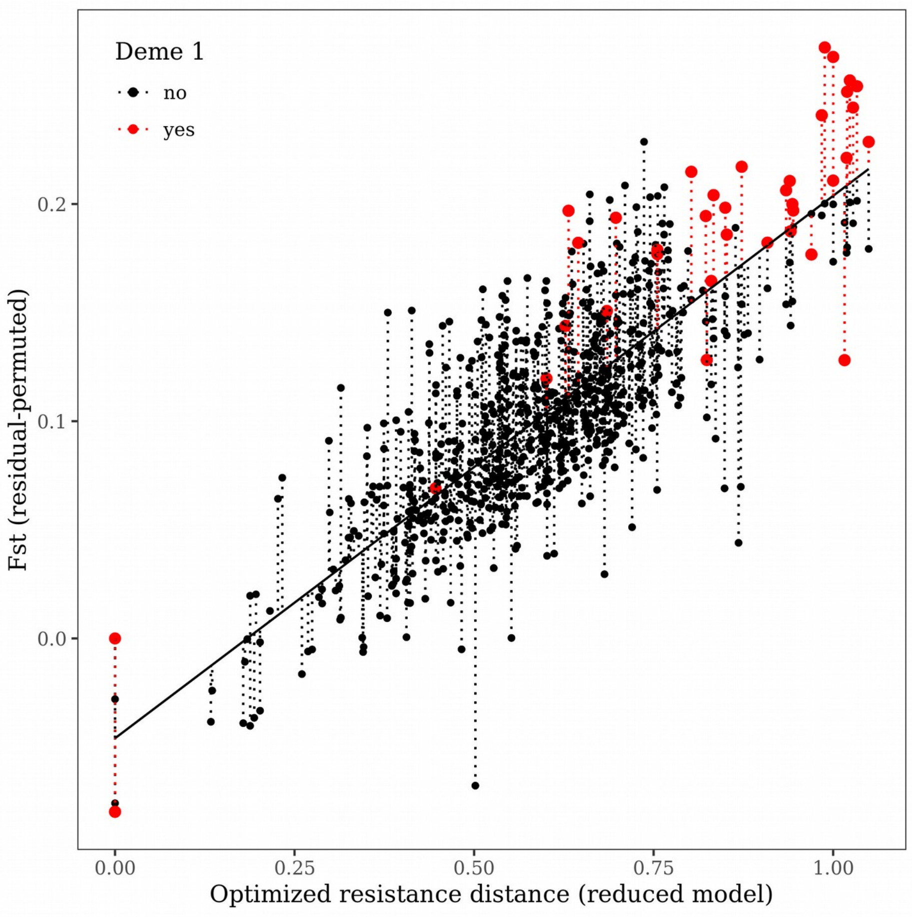


# Residual permutation tests

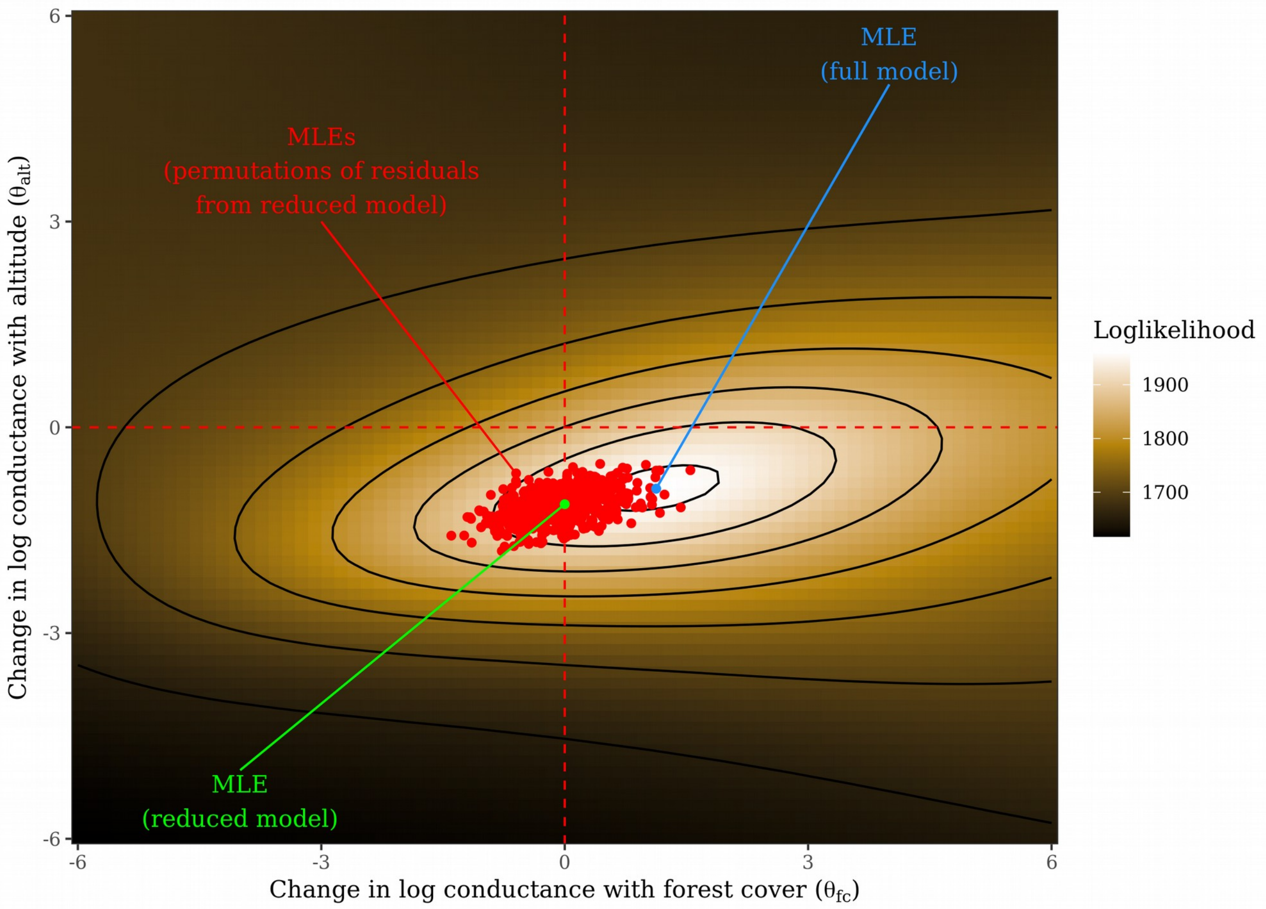


# Residual permutation tests



# Residual permutation tests



# “Measurement process”

**Nonnegative least squares.** Fast, but ignores dependence in pairwise data.

# “Measurement process”

**Nonnegative least squares.** Fast, but ignores dependence in pairwise data.

**MLPE (Clarke 2002).** Slower, less stable, but models dependence.

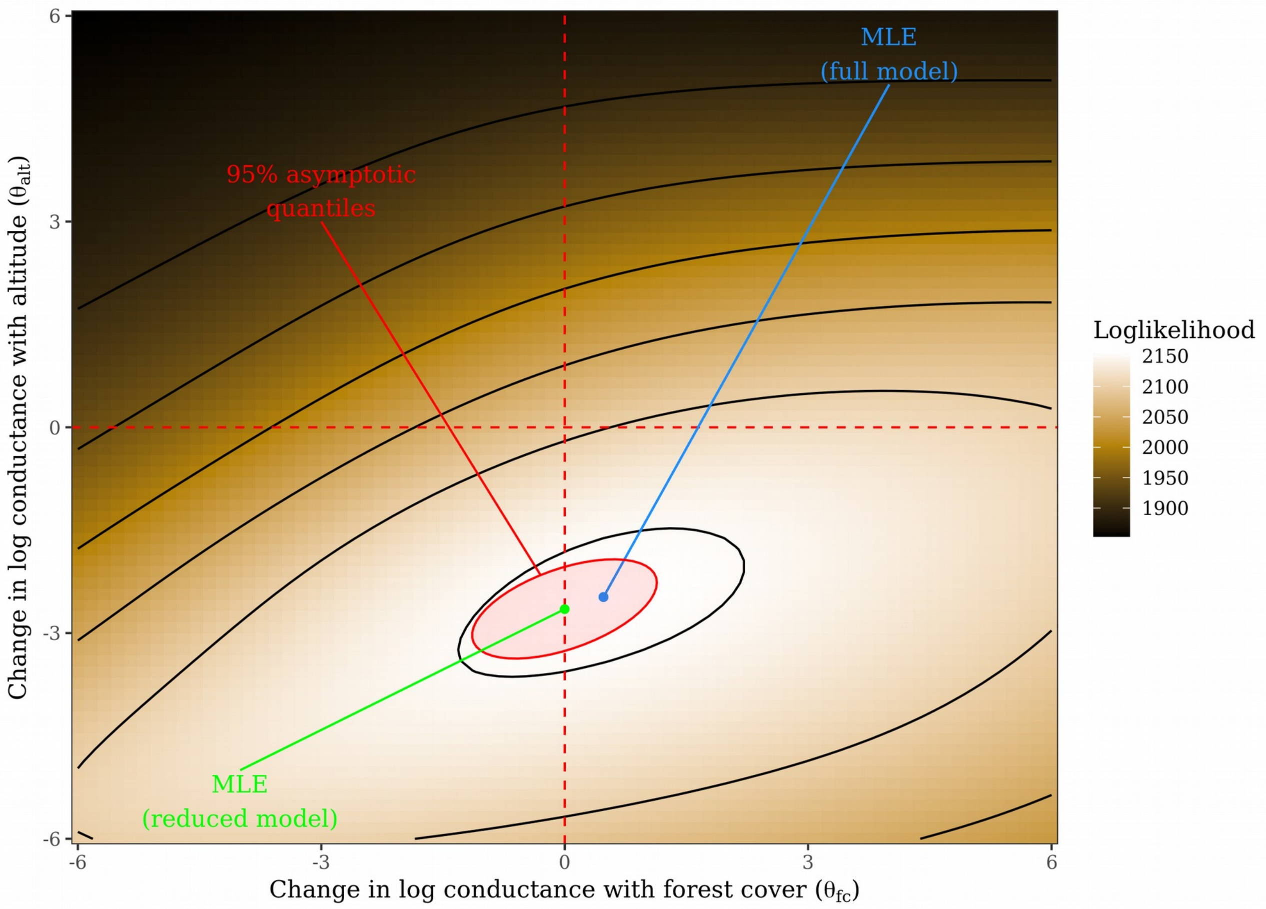
# “Measurement process”

**Nonnegative least squares.** Fast, but ignores dependence in pairwise data.

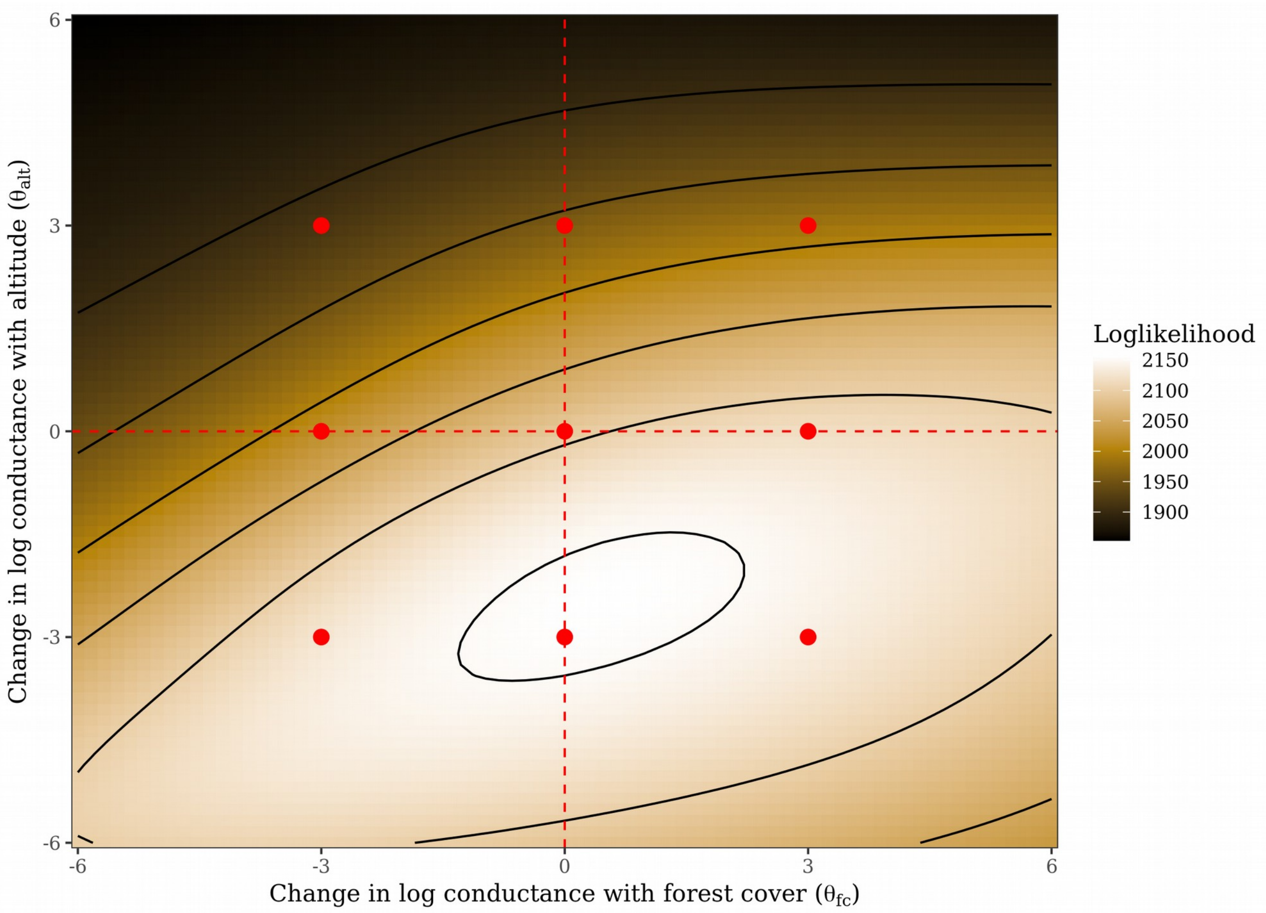
**MLPE (Clarke 2002).** Slower, less stable, but models dependence.

**Gen. Wishart (Peterson 2019).** Even slower, less stable, but principled.

# Asymptotics (modeling dependence)



# Compare to “Mantel-across-grid”



# **PROBLEM:** Likelihood is costly

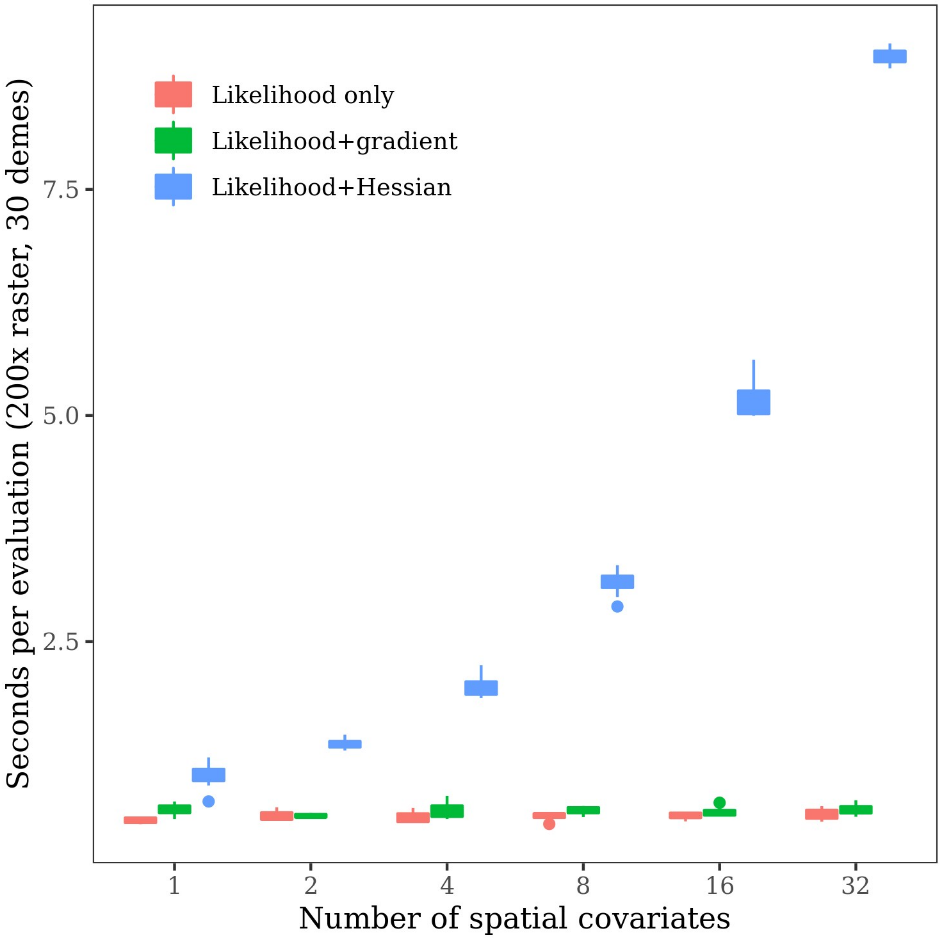
**1.** Use structure of nullspace to reduce to single Choleski decomposition per likelihood evaluation

# **PROBLEM:** Likelihood is costly

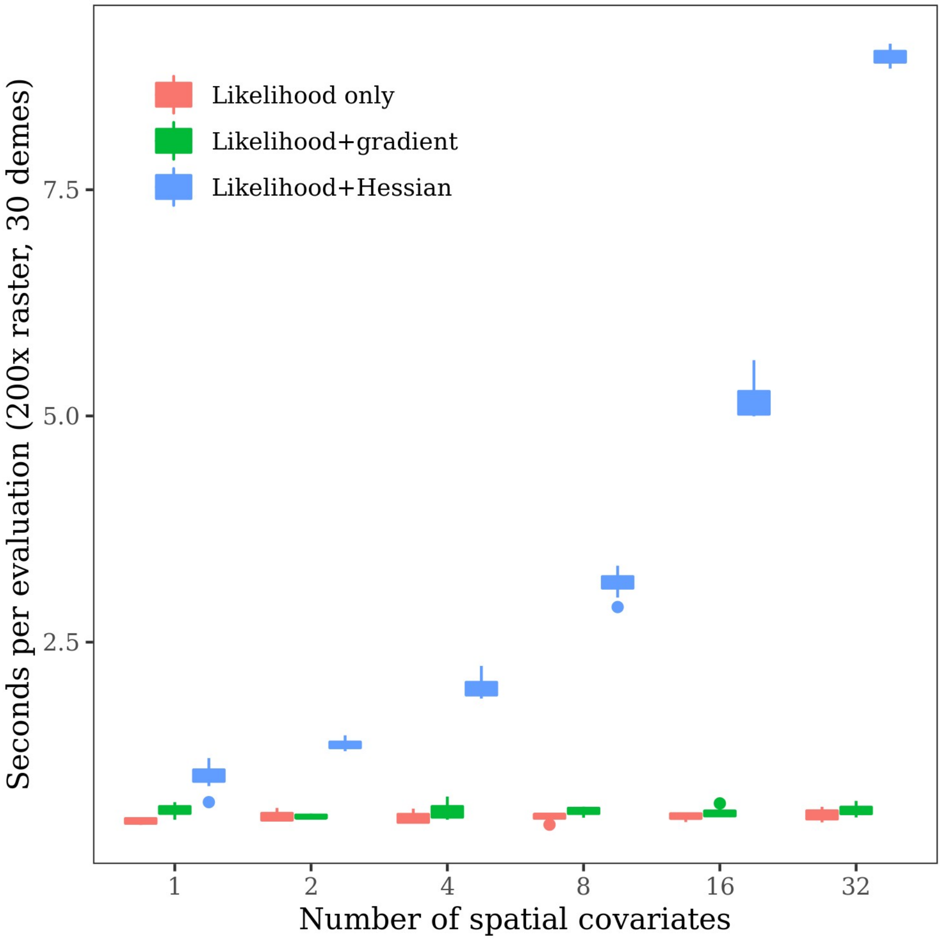
**1.** Use structure of nullspace to reduce to single Choleski decomposition per likelihood evaluation

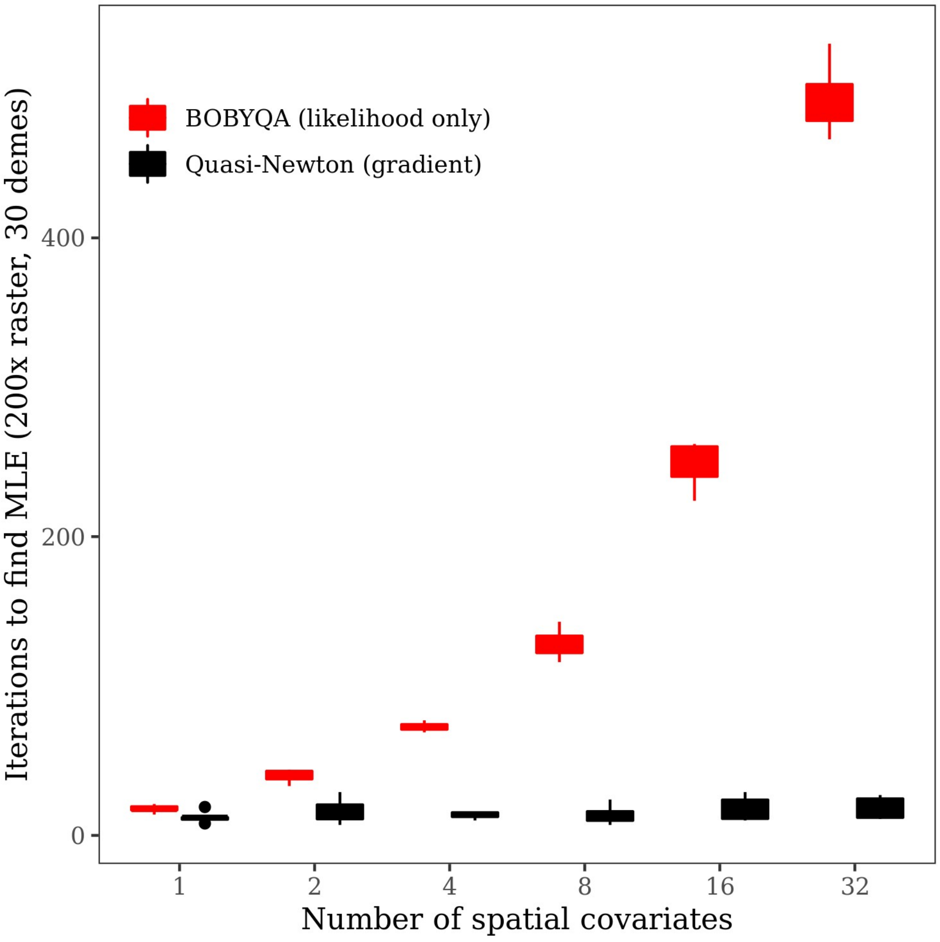
**2.** Use chain rule in reverse and dynamic programming to calculate gradient and Hessian without needing to “re-Choleski”

# Fast computation of gradient



# Gradient-accelerated optimization



# Influence and leverage

**∂θhat/∂y.** Change in parameter estimate w/ observed genetic distance.

# Influence and leverage

**∂θhat/∂y.** Change in parameter estimate w/ observed genetic distance.

**∂θhat/∂x.** Change in parameter estimate w/ spatial covariate.

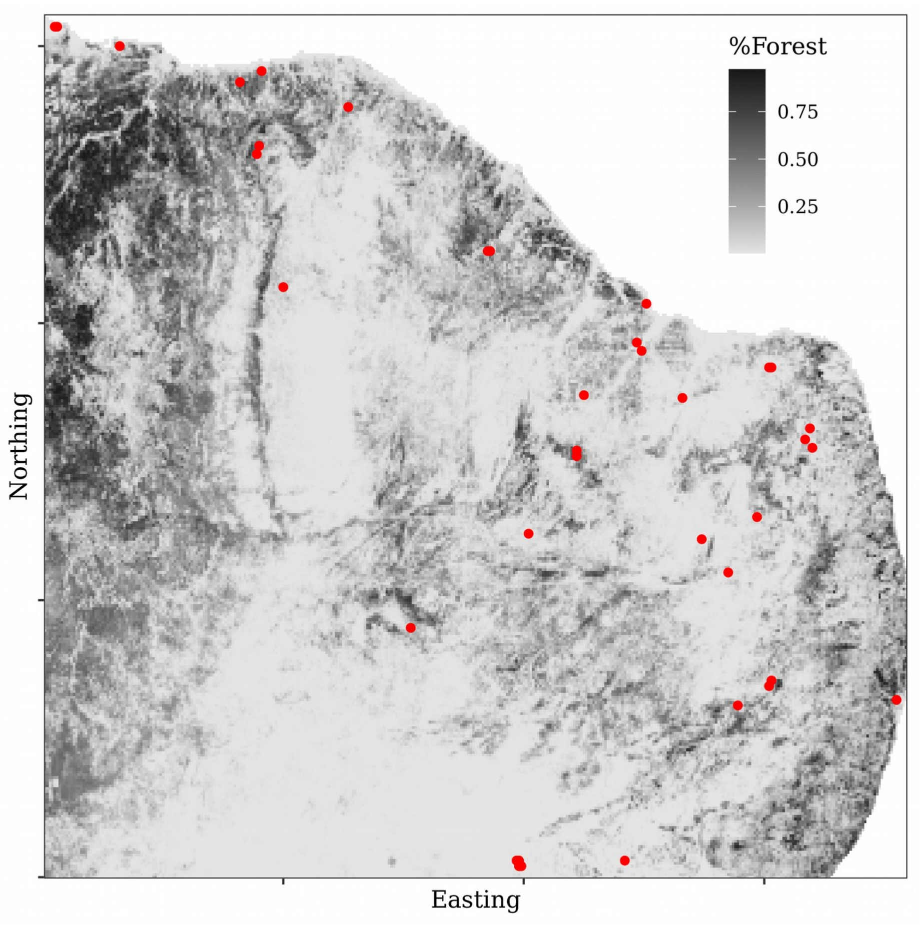
# Influence and leverage

**∂θhat/∂y.** Change in parameter estimate w/ observed genetic distance.

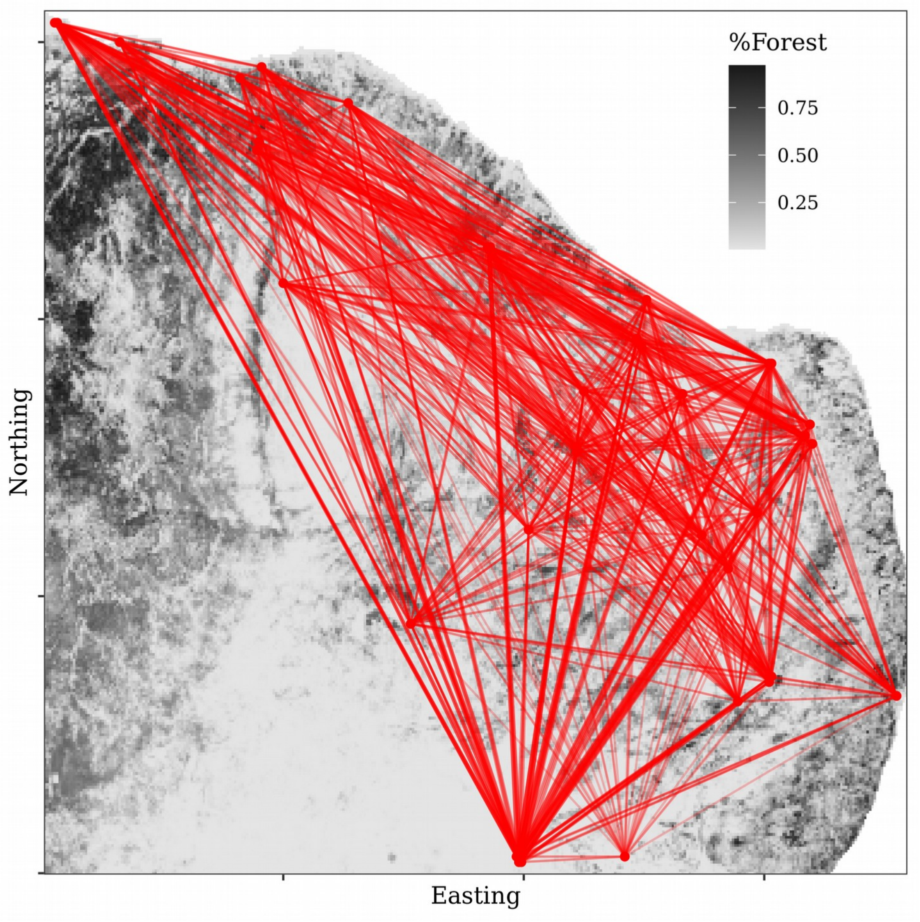
**∂θhat/∂x.** Change in parameter estimate w/ spatial covariate.

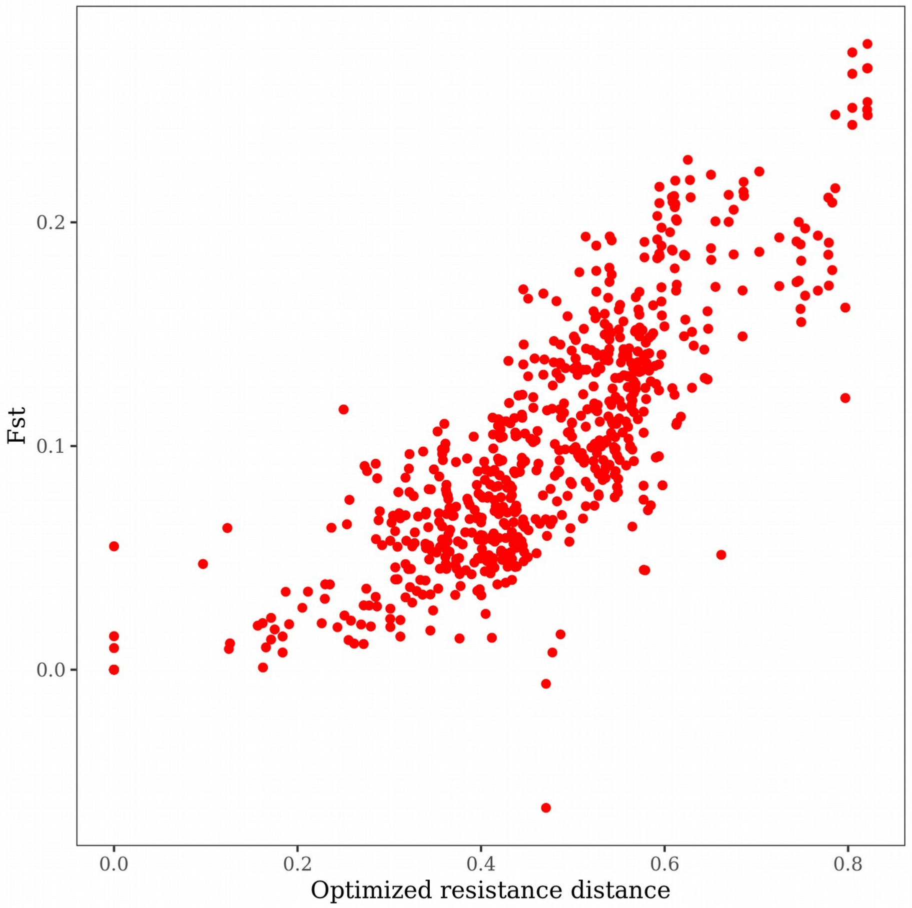
**∂yhat/∂y.** Change in fitted genetic distance w/ observed genetic distance.

# Influence and leverage

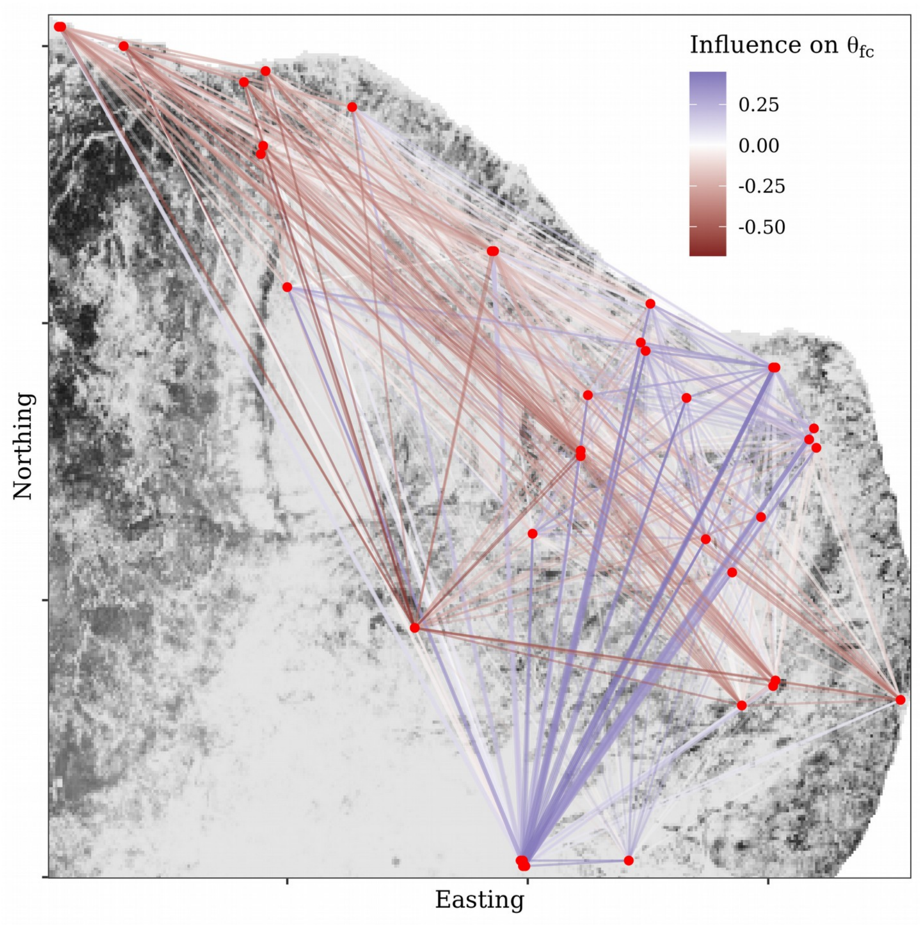


# Influence and leverage

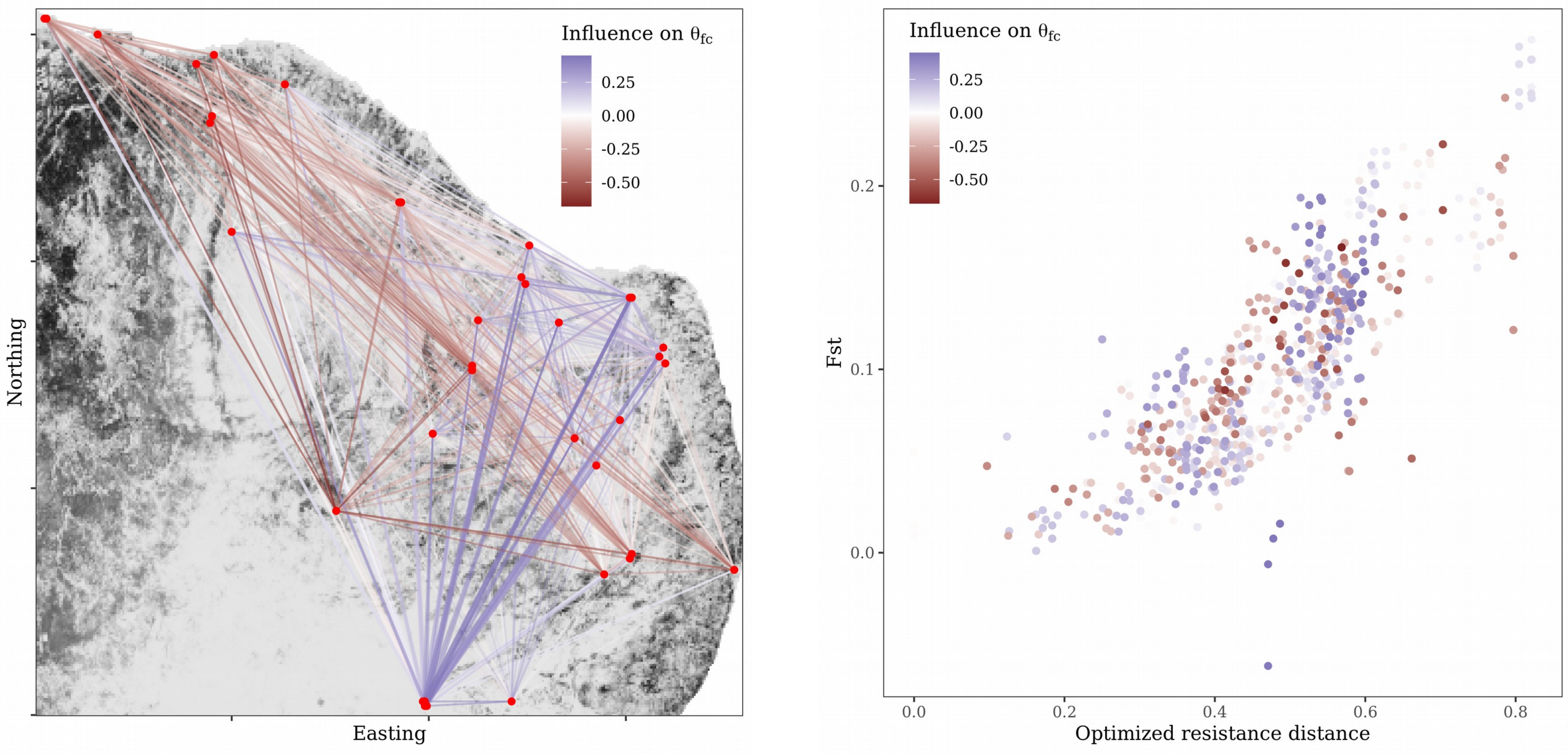


# Influence and leverage



# Influence and leverage



# github.com/nspope/radish

**1.** IBR inference is as “simple” as fitting a nonlinear model via ML.

# github.com/nspope/radish

**1**. IBR inference is as “simple” as fitting a nonlinear model via ML.

**2**. Reverse algorithmic differentiation + gradient-accelerated optimization = seconds/minutes to fit these models.

# github.com/nspope/radish

**1**. IBR inference is as “simple” as fitting a nonlinear model via ML.

**2**. Reverse algorithmic differentiation + gradient-accelerated optimization = seconds/minutes to fit these models.

**3**. Influence/leverage diagnostics are a useful byproduct.

# github.com/nspope/radish

**In progress:**

- Stan hooks (Bayesian inference)
- Unsupervised estimation
- Formula interface for categorical spatial covariates
- Approximate cross-validation
- Confidence intervals on per-cell conductance