



breedR: statistical methods for forest genetic resources analysis

Facundo Munoz, Leopoldo Sanchez Rodriguez

► To cite this version:

Facundo Munoz, Leopoldo Sanchez Rodriguez. breedR: statistical methods for forest genetic resources analysis. Trees for the future: plant material in a changing climate, Nov 2014, Tulln, Austria. 13 p. hal-02801127

HAL Id: hal-02801127

<https://hal.inrae.fr/hal-02801127>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Facundo Muñoz & Leopoldo Sanchez

breedR: statistical methods for forest genetic resources analysis

3rd Annual Meeting T4F

Tulln – Austria 17-21 November 2014

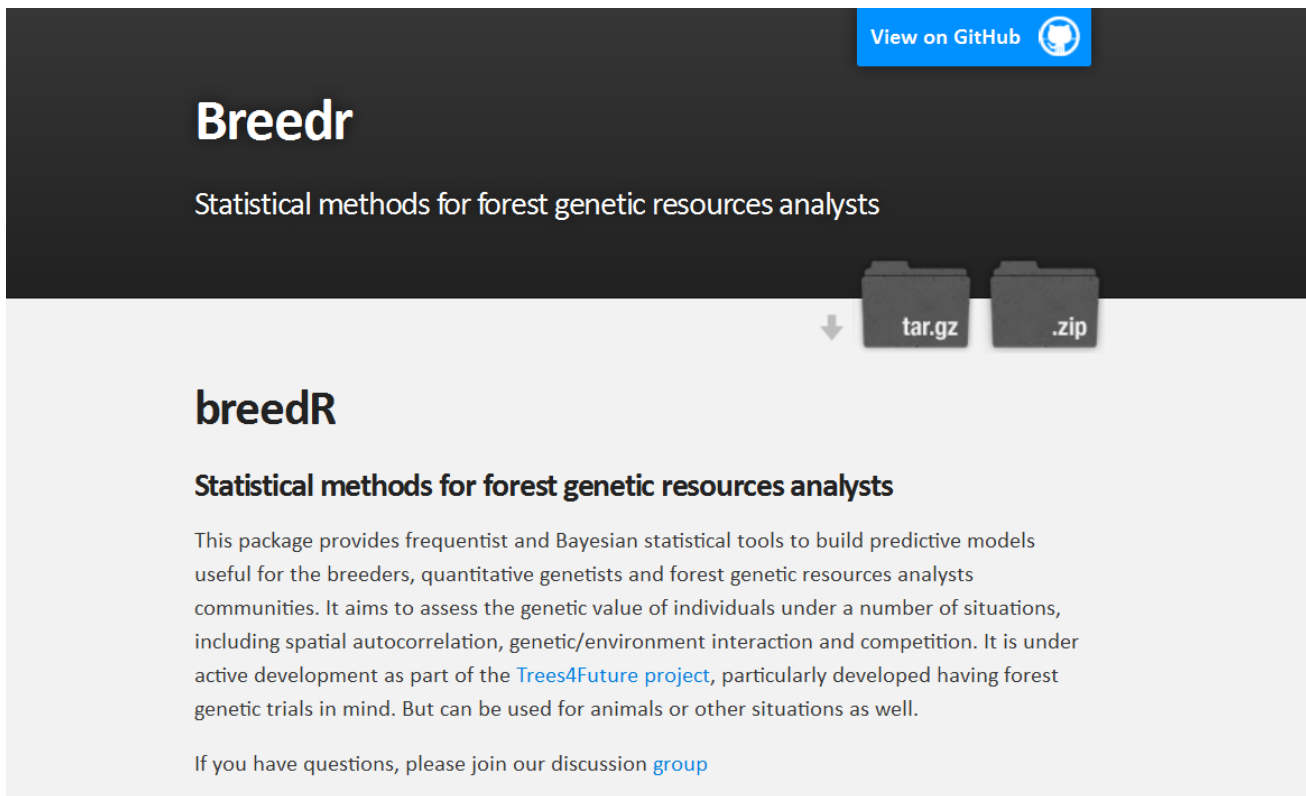
vector_y
150
100
50

resp_sp
150
100
50
0
-50
-100



Where to get BreedR?

- BreedR general web page: <http://famuvie.github.io/breedR/>
- BreedR Forge: <https://github.com/famuvie/breedR>



The screenshot shows the GitHub repository page for 'BreedR'. At the top right, there is a blue button that says 'View on GitHub' with the GitHub logo. The main heading is 'Breedr' in a large, bold, white font. Below it, the subtitle reads 'Statistical methods for forest genetic resources analysts'. In the center, there are two folder icons labeled 'tar.gz' and '.zip', with a downward arrow pointing to them. Below the folder icons, the word 'breedR' is written in a bold, black font. Underneath, the subtitle 'Statistical methods for forest genetic resources analysts' is repeated. The main body of text describes the package as providing frequentist and Bayesian statistical tools for building predictive models for forest genetic resources analysts. It mentions the 'Trees4Future project' and provides a link to the discussion group.

Breedr

Statistical methods for forest genetic resources analysts

tar.gz .zip

breedR

Statistical methods for forest genetic resources analysts

This package provides frequentist and Bayesian statistical tools to build predictive models useful for the breeders, quantitative genetists and forest genetic resources analysts communities. It aims to assess the genetic value of individuals under a number of situations, including spatial autocorrelation, genetic/environment interaction and competition. It is under active development as part of the [Trees4Future project](#), particularly developed having forest genetic trials in mind. But can be used for animals or other situations as well.

If you have questions, please join our discussion [group](#)



- R-package implementing **statistical models** specifically suited for forest genetic resources evaluation.
- Ultimately Mixed Models, flexible, easy to implement and use
- **BreedR** acts as an **interface** which provides the means to:
 1. **Combine** any number of explanatory factors into these models
 2. Compute automatically **incidence** and **covariance matrices** (from explanations to data) from a few input parameters
 3. **Fit** the model
 4. Plot data and results, and perform **model diagnostics**



- Mixed model:

$$y = \mu + \beta x + Zu + \varepsilon$$

- Progeny testing:
 - ✓ estimation of variance components
 - ✓ prediction of breeding values

essential toolkit

Spatial
statistics

G×E
analysis

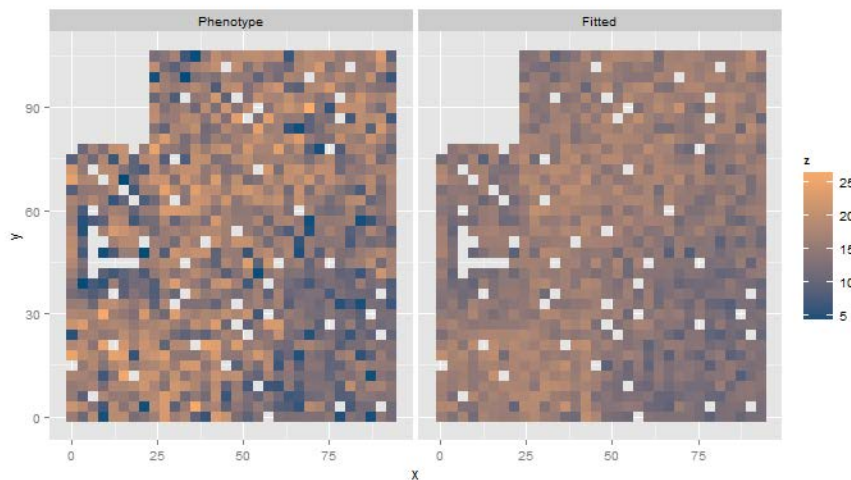
Mixed-model
framework

Pedigree-
free
evaluation

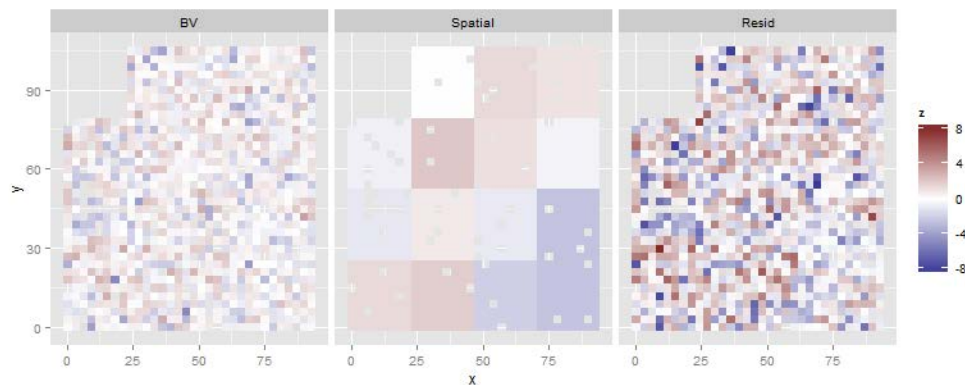
Competition
evaluation



BreedR demo

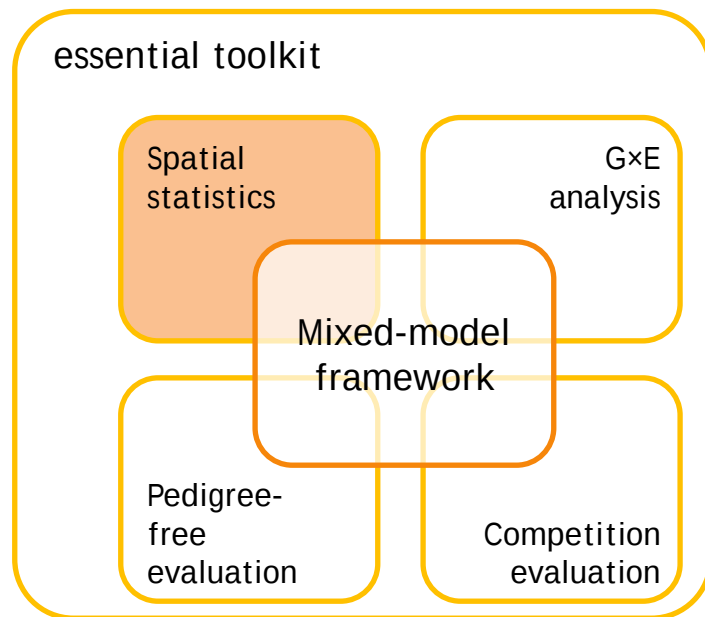


output example with
globulus data set



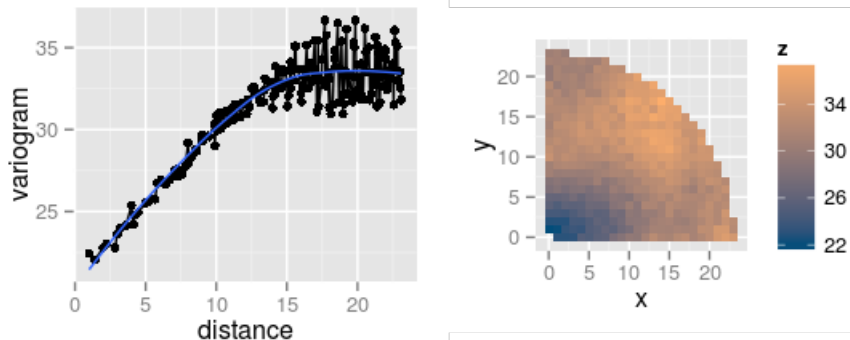


- **Spatial models**
- trials often occupy heterogeneous environments, partially accounted for by devising a priori block designs
- alternatives (BreedR) propose accounting for heterogeneity a posteriori together with the fitting of the evaluation model
- gain in flexibility whenever there is faulty design either by conception or by ageing of trial
- take the form of individual-tree mixed models

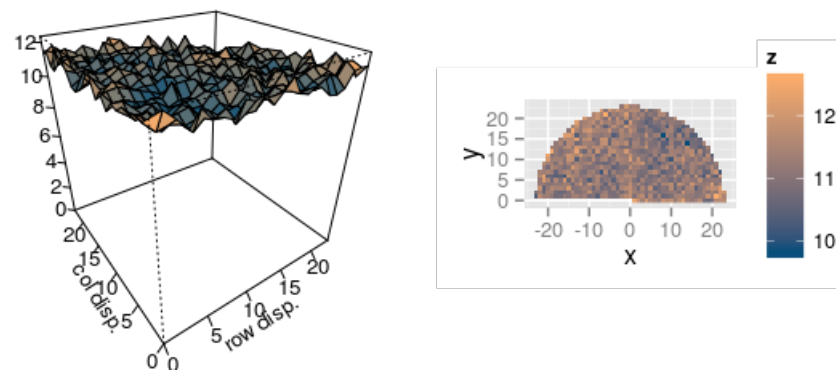
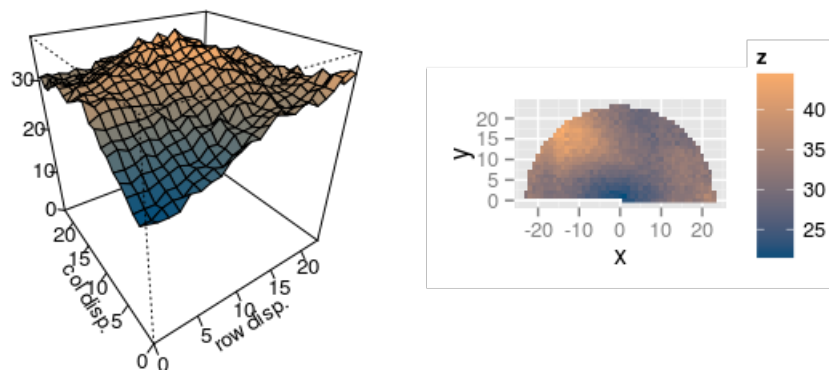
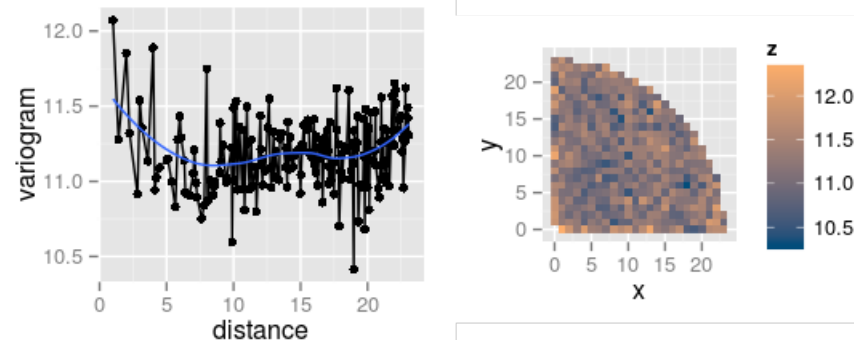




without spatial adjustments



with spatial adjustments



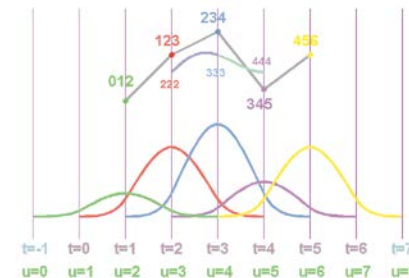
Variograms: diagnosis of spatially distributed patterns of residual variation



B-Splines model

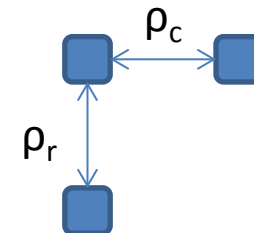
- A continuous and smooth spatial surface built from a linear combination of basis functions, whose coefficients are modelled as a random effects

Cubic B-Spline Basis

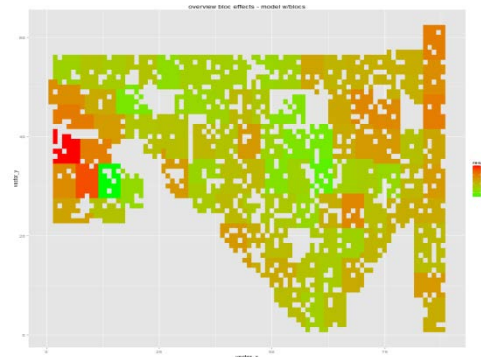


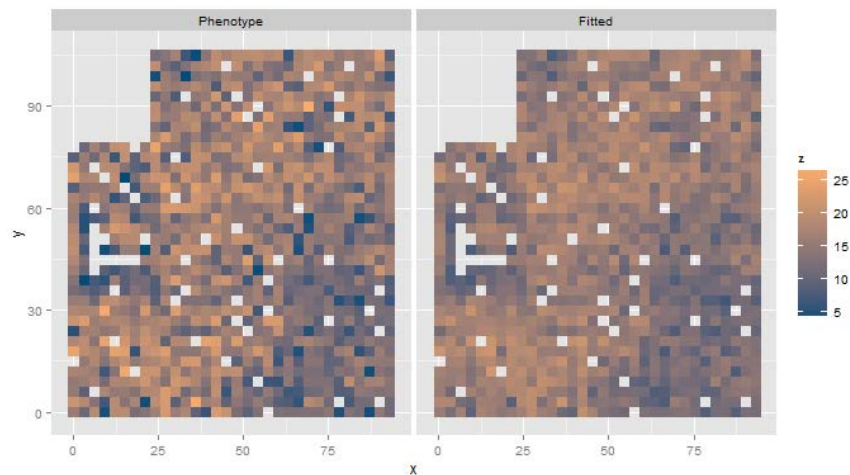
Autoregressive model

- A separable kronecker product of First order Autoregressive processes on the rows and the columns

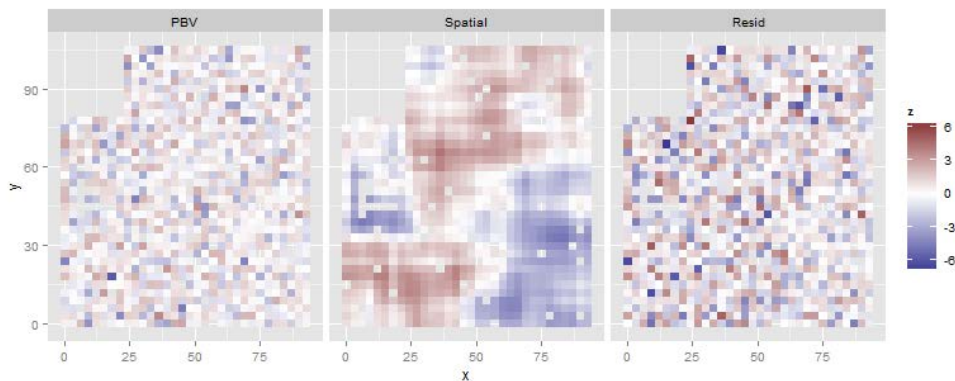


and old good Blocks models!



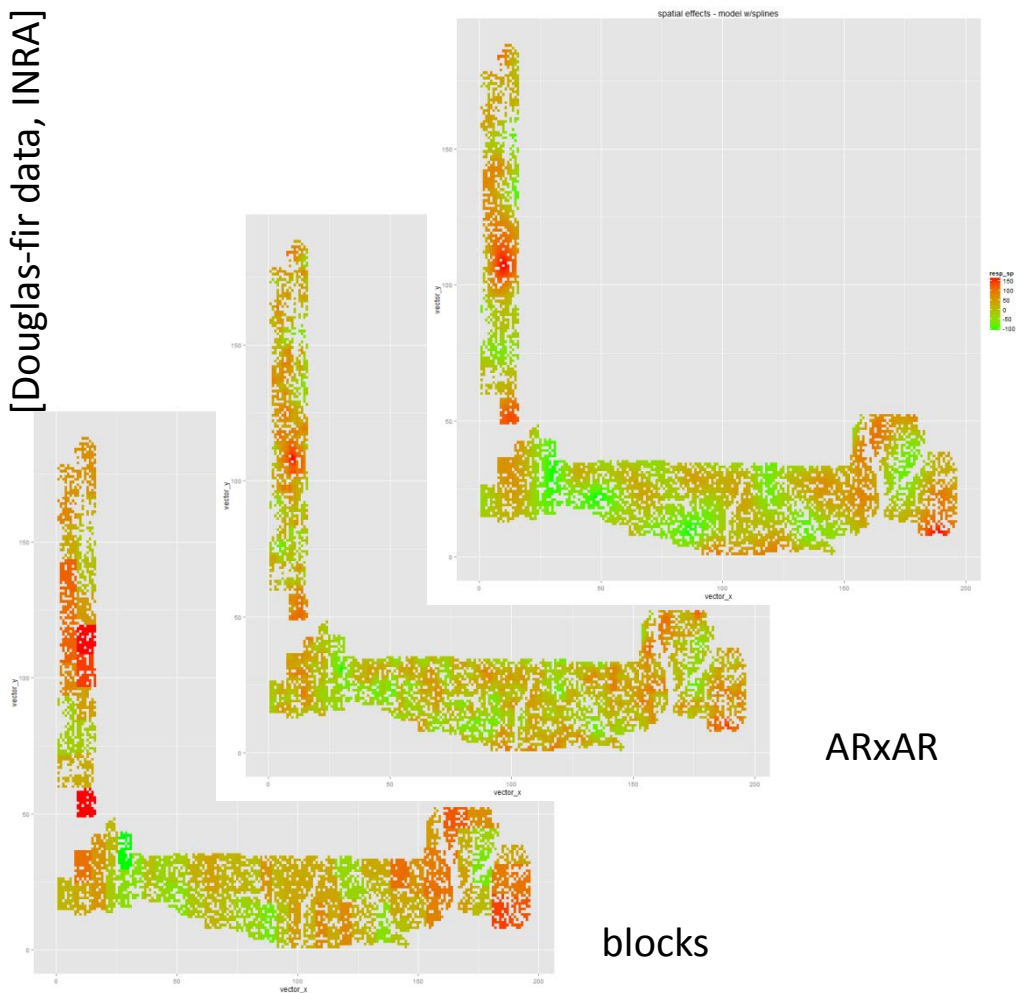


output example with
globulus data set and
spatial models





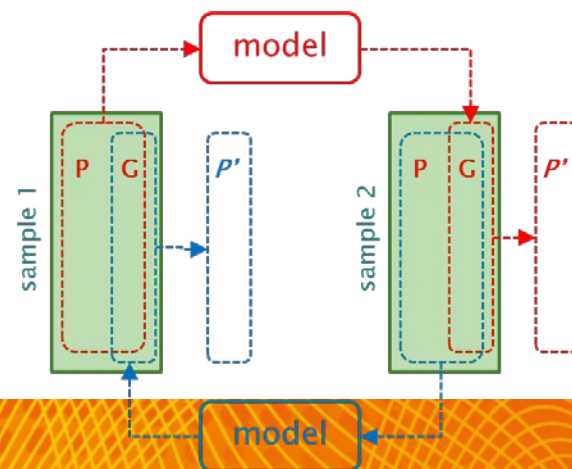
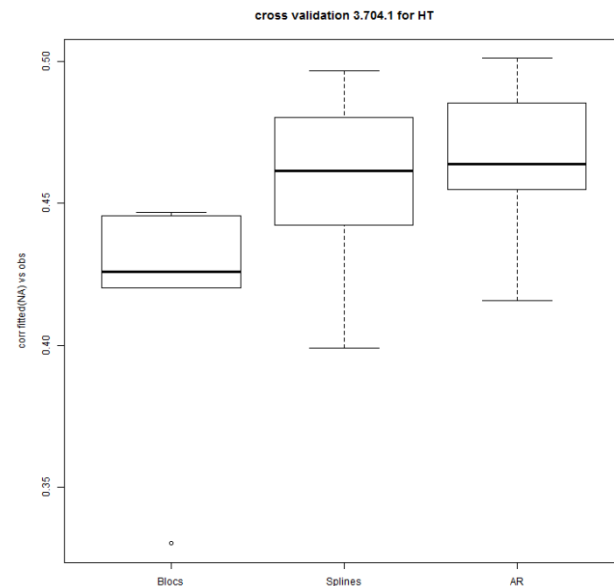
[Douglas-fir data, INRA]



Splines

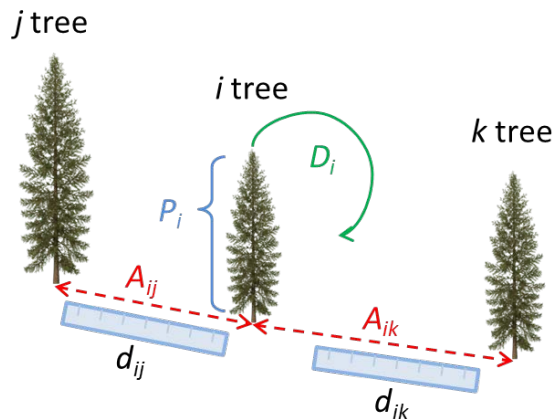
$$AR \times AR$$

blocks





- Competition models
- Competition reflects the impairing interplay of closely growing trees, often when local resources are limiting



- Competition: an antagonism between direct and associative effects, that results ultimately into impaired phenotypes

essential toolkit

Spatial
statistics

G×E
analysis

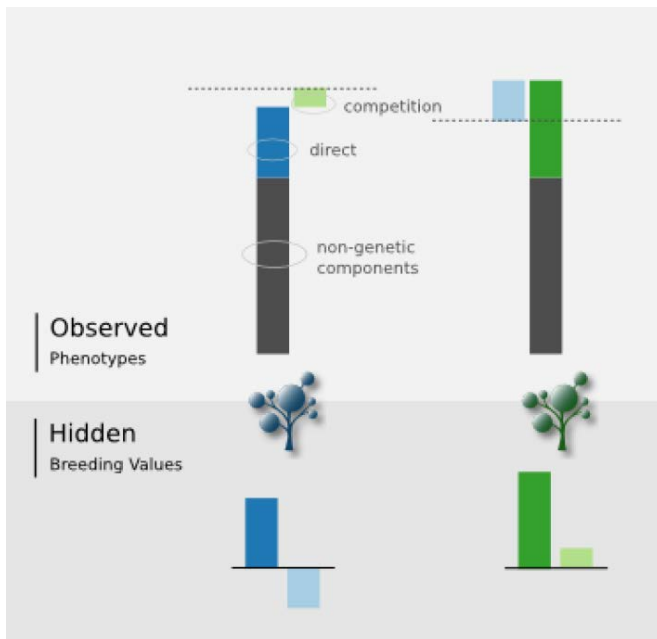
Mixed-model
framework

Pedigree-
free
evaluation

Competition
evaluation



- Competition models



essential toolkit

Spatial
statistics

G×E
analysis

Mixed-model
framework

Pedigree-
free
evaluation

Competition
evaluation



- Other models developments to come in 2015
- multiple traits, multiple sites
- G by E
- longitudinal data
- Selection tools (index selection)
- Bayesian inference

