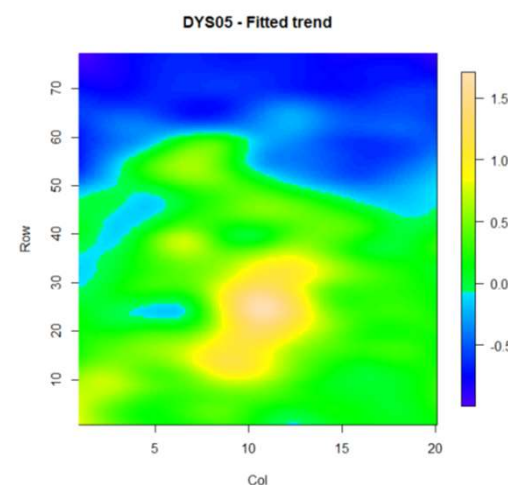# Spatial Analysis of Field Trials

Luis Fernando Delgado Munoz

luis.delgado@cgiar.org

Faisalabad, November 2023

@BiovIntCIAT_eng
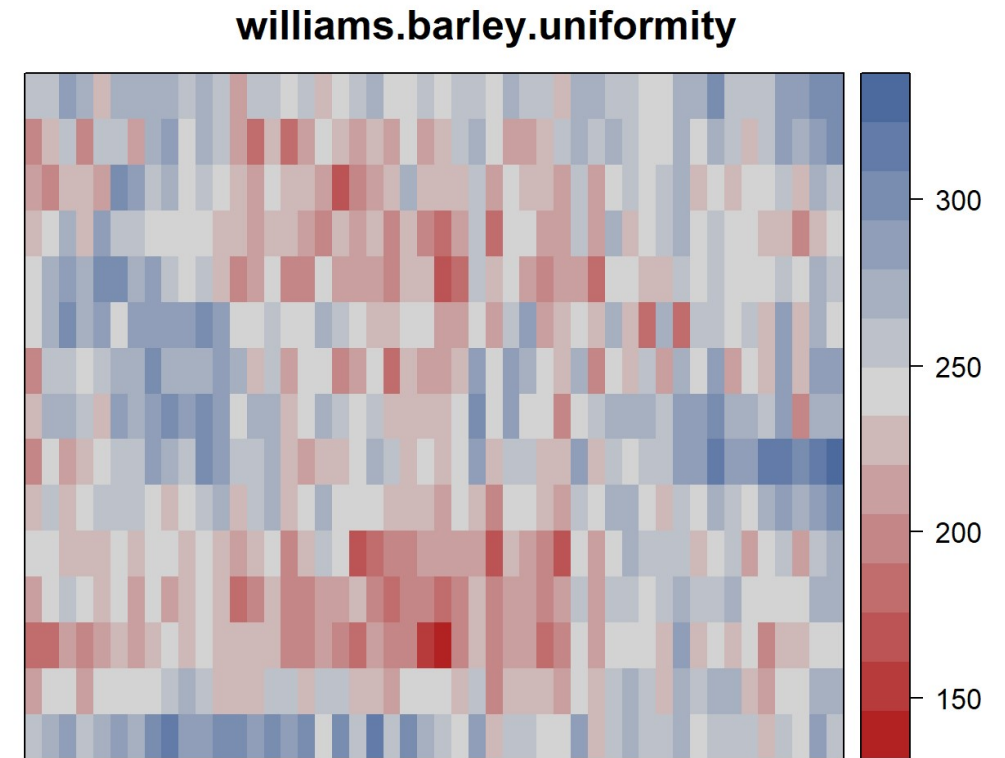@BiovIntCIAT_esp

#Alliance4Science

Spatial variation can be complex

Field variation often has irregular patterns than don't fully coincide with specific rows/columns.

How can we model this variation in

a more straightforward way?

williams.barley.uniformity

Models so far…

❑ Blocks (RCBD)

$$y_{ik} = \mu + b_k + G_i + \epsilon_{ik}$$

❑ Replicate and blocks within rep (alpha)

$$y_{ijk} = \mu + R_j + b_{k(j)} + G_i + \epsilon_{ijk}$$

❑ Row and column (row-column)

$$y_{ijk} = \mu + R_j + C_k + G_i + \epsilon_{ijk}$$

❑ Replicate, row and column (resolvable row-column)

$$y_{ijkm} = \mu + Rep_m + R_{j(m)} + C_{k(m)} + G_i + \epsilon_{ijkm}$$

In all these models, we assume that random effects are independent and come from a normal distribution

Alliance
Bioversity & CIAT   CGIAR

Causes of field spatial variation

- Field position.
- Soil conditions.
- Disease.
- Wildlife impacts
- The way the field was prepared before and during sowing or planting.

williams.barley.uniformity
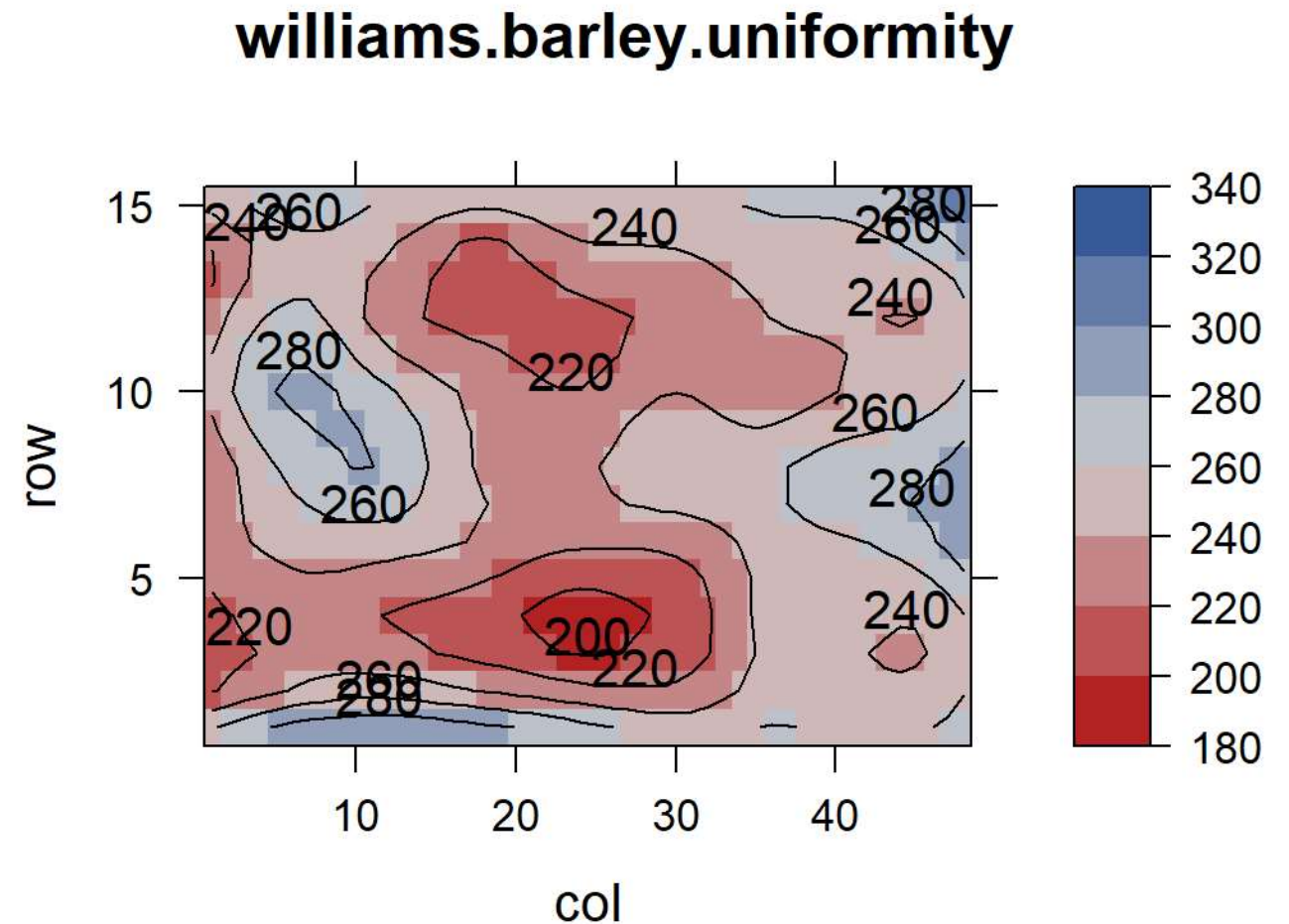
Sources of spatial variation in field trials

❑ Types of variation

- ○ Extraneous variation (rep + rep.row + rep.col)
  - Design (replicates, row, columns) + covariates

- ○ Global trend variation $(f(x_1) + f(x_2))$
  - Additive functions of spatial coordinates

- ○ Local trend variation (ξ)
  - Spatially dependent noise $AR1(x_1) \times AR1(x_2)$

x_1=variation in the direction of rows
x_2=variation in the direction of columns
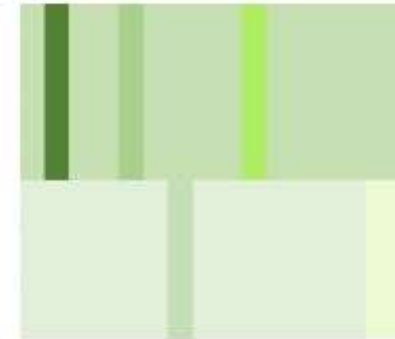
Uniformity trial (barley)

- **15 rows x 48 columns grid**
- **Variation across experiment due to uncontrolled conditions (="noise")○Minimum 1.490 ton/ha**
- **Maximum 2.290 ton/ha**
- **Variance = 0.084 (ton/ha)2**
- **Std = 0.289 ton/ha**

The use of uniformity data in the design and analysis of cotton and barley variety trials
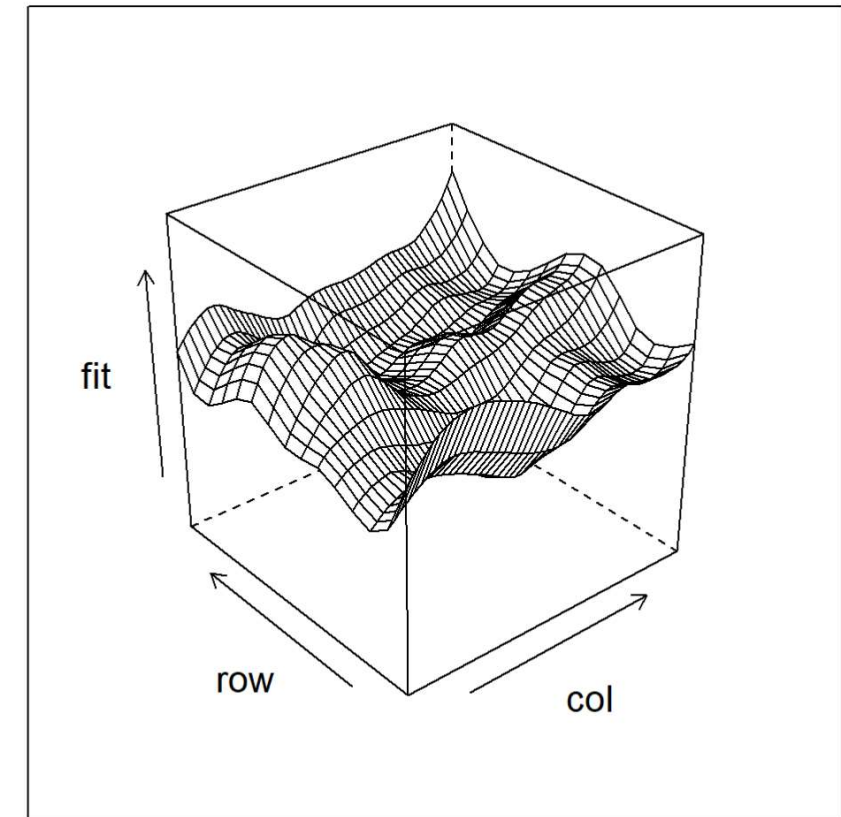
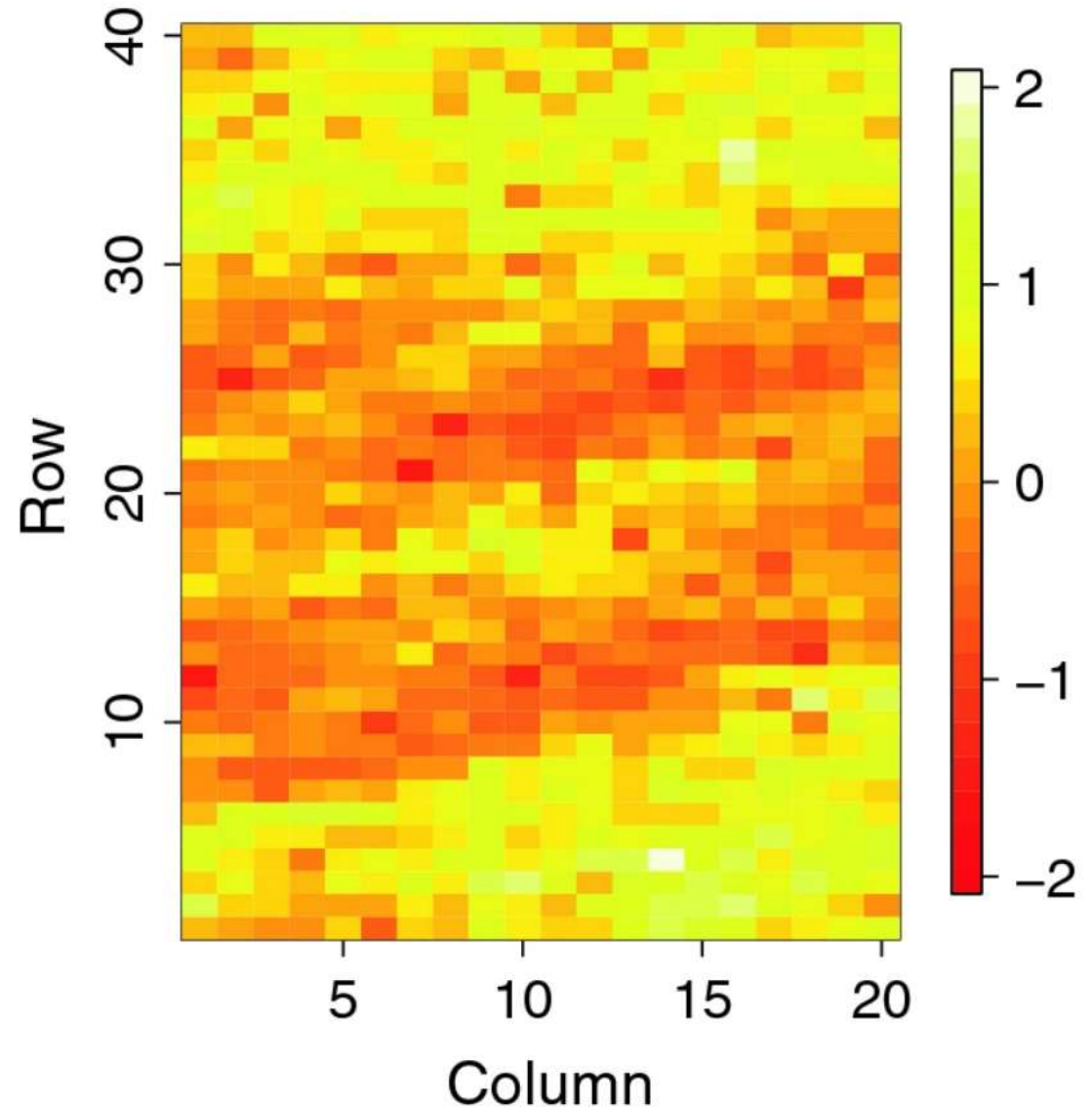ER Williams and DJ Luckett

williams.barley.uniformity

Is it reasonable to assume that plots are independent?

Spatial variation

- Mix model framework allows modelling of spatial effect.
- Decaying correlation (AR1) in both row and col direction

$$\begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

(c) AR1⊗AR1 model

Field trials: 'traditional' approach (AR x AR)

- Gilmour, Cullis & Verbyla, JABES, 1997

- Assume that residuals are not independent > they are spatially correlated

- Model local trend as correlated noise (AR x AR)> AR1; residual in plot 2 is correlated to the residual in plot 1 (adjacent plot)

- Example: autoregressive parameters $\rho_r = 0.56$ and $\rho_c = 0.75$ indicate that the correlation of residuals belonging to adjacent plots is 0.56 in the rows direction and 0.75 in the columns direction.
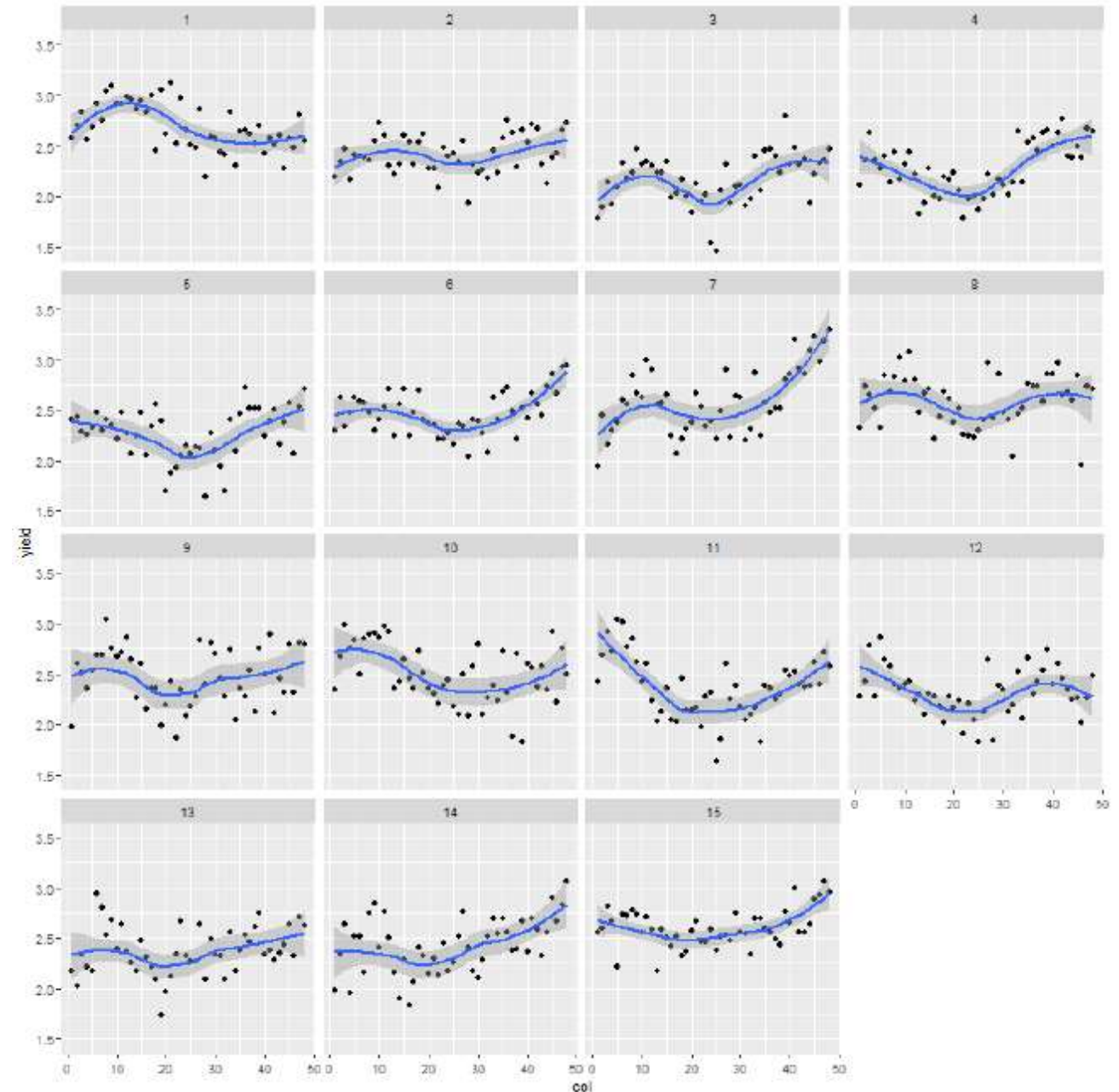
Challenging to use: 'traditional' approach (AR x AR)

- Limited flexibility
- Convergence problems due to many parameters
- Model selection difficult
- Interpretation of results (variograms) requires a lot of training
- No automated phenotypic analysis of series of trials.

Half-squared differences between pairs of residuals

Distance between columns

Distance between rows
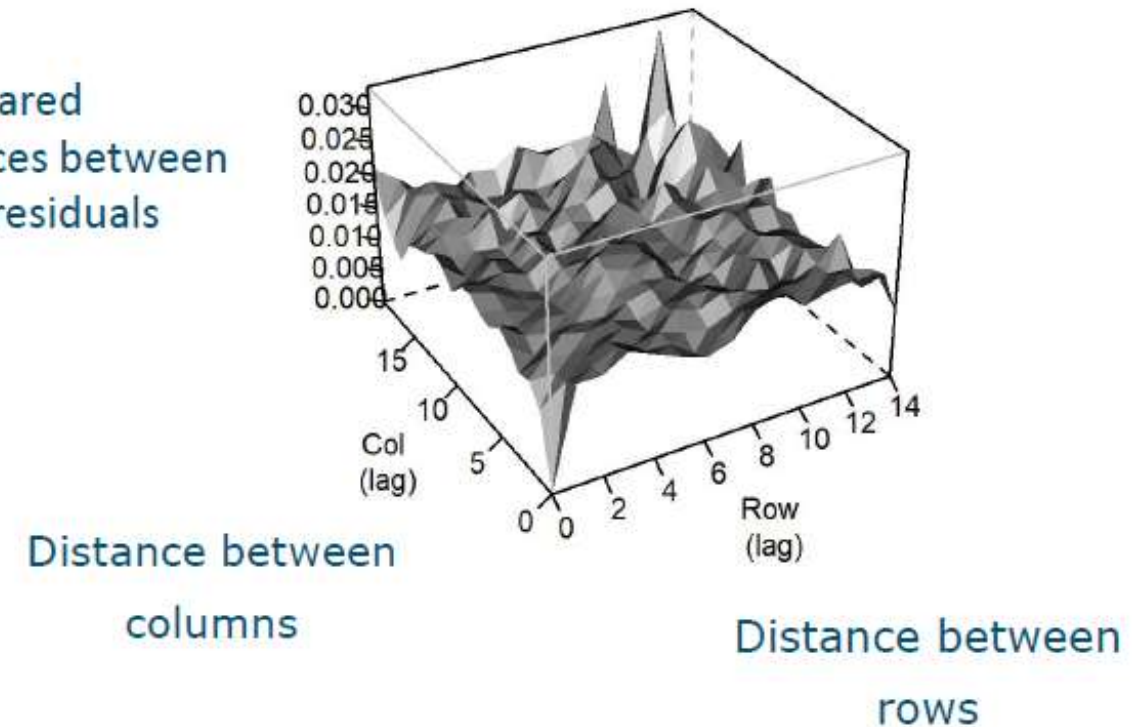
A proposal for spatial variation alternative to ARxARmodels

- Two-dimensional smooth function of spatial coordinates (interaction surface)

- Global + local variation

- Yield = genotype + rep + rep.row+ rep.col + $f(x1, x2)$ + ε
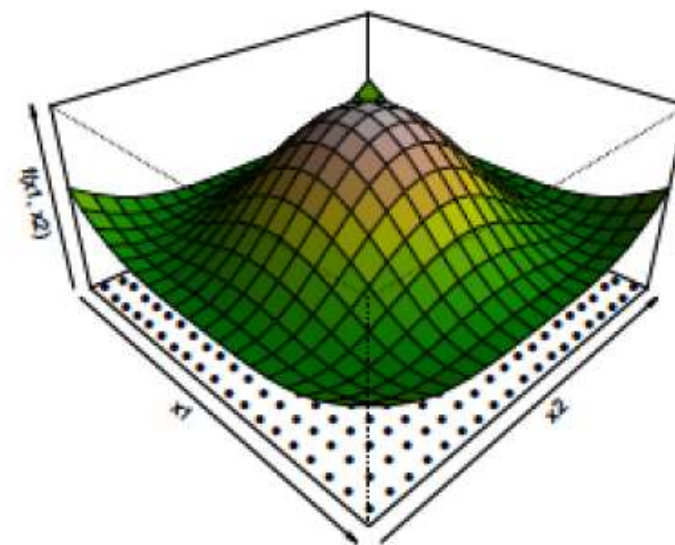
x_1=variation in the direction of rows
x_2=variation in the direction of columns

Correcting for spatial heterogeneity in plant breeding experiments with P-splines

María Xosé Rodríguez-Álvarez [a,b,*], Martin P. Boer [c],
Fred A. van Eeuwijk [c], Paul H.C. Eilers [d]

[a] BCAM — Basque Center for Applied Mathematics, Alameda de Mazarredo, 14. E-48009 Bilbao, Basque Country, Spain
[b] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[c] Biometris, Wageningen University & Research, Wageningen, The Netherlands
[d] Erasmus University Medical Centre, Rotterdam, The Netherlands

Package 'SpATS'

November 7, 2022

Type Package

Title Spatial Analysis of Field Trials with Splines

Version 1.0-18

Date 2022-11-07

Imports stats, grDevices, graphics, fields, spam, data.table, methods

Description Analysis of field trial experiments by modelling spatial trends using two-dimensional Penalised spline (P-spline) models.

License GPL

NeedsCompilation no

Author Maria Xose Rodriguez-Alvarez [aut, cre],
Martin Boer [aut],
Paul Eilers [aut],
Fred van Eeuwijk [ctb]

Maintainer Maria Xose Rodriguez-Alvarez <mxrodriguez@uvigo.es>

Repository CRAN

Date/Publication 2022-11-07 11:40:19 UTC

SpATSmodel: ANOVA-type decomposition of phenotypic variation

- Instead of modelling the correlation between rows and columns, variation is decomposed with splines.

- Splines are function defined piecewise by polynomials.

- The total variation is decomposed as the variation due to rows, columns and their interaction.

main smooth trends

linear-by-smooth interactions

smooth-by-smooth interaction

Rodríguez-Álvarez et al. (2016)
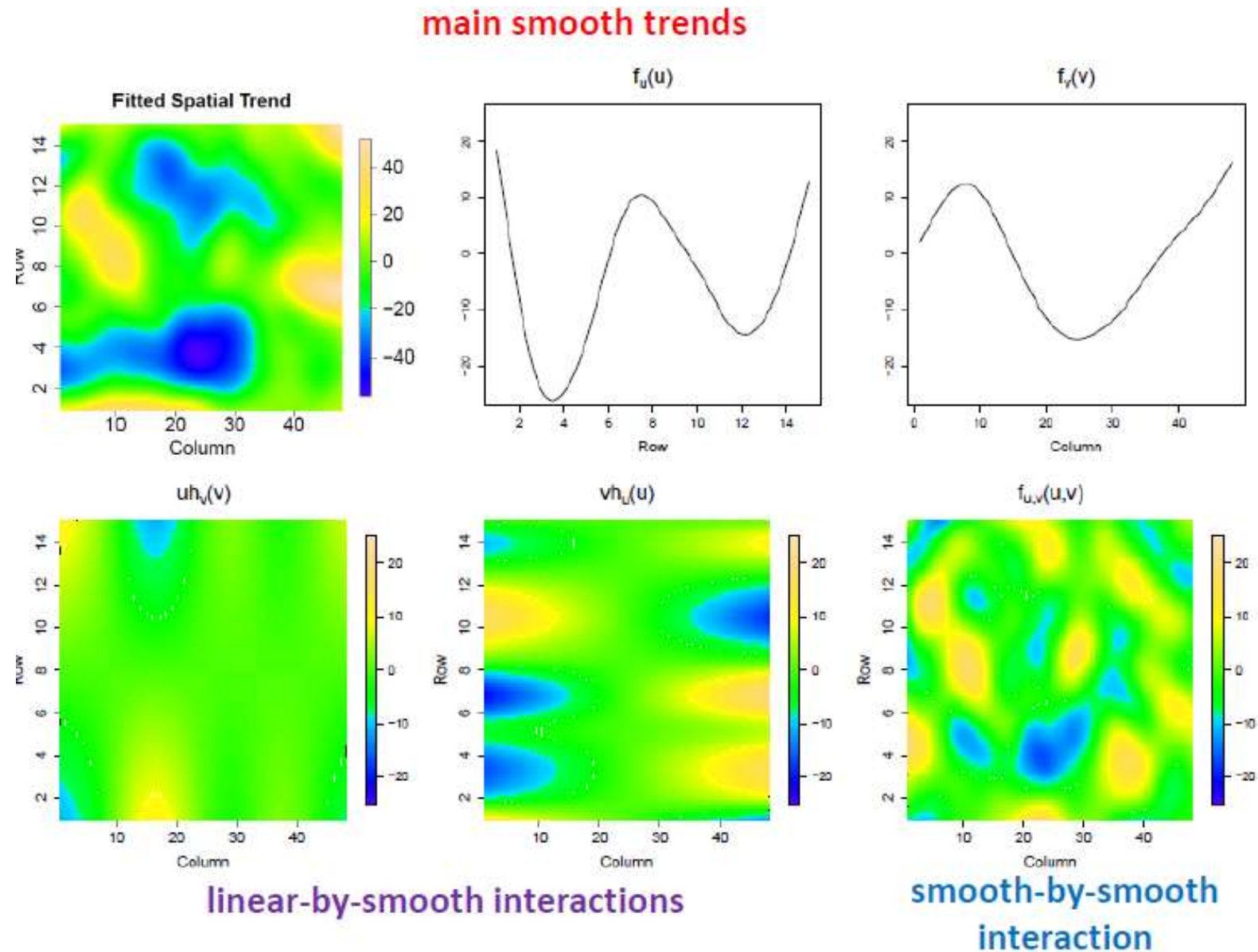
Wheat in Chile (Ladoet al, 2013)

- 384 wheat genotypes sown in a field of 40 rows and 20 columns (800 plots)

- Genotypes allocated to the plots, according to an alpha design
  - 20 incomplete blocks
  - Each block with 20 genotypes
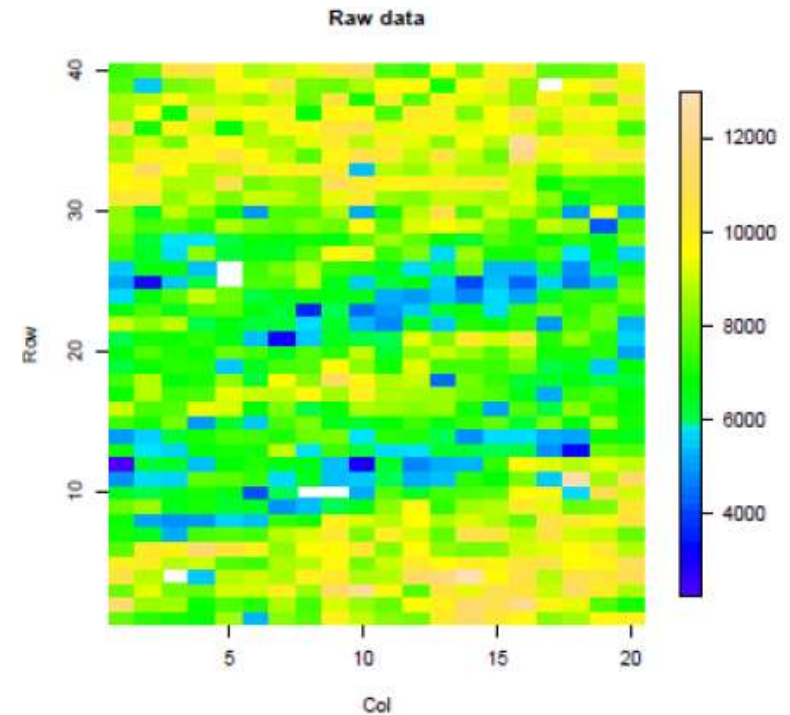 - Unequal replication:
    - 382 genotypes were sown twice
    - 2 genotypes were sown 18 times

- 6 experimental units had a missing genotype due to crop failure

Raw data

```
SpATS(response = "Yield",
genotype = "Geno", genotype.as.random = TRUE,
spatial = ~ PSANOVA(x1, x2, nseg = c(15,30), nest.div = 2),
random = ~ C + R, fixed = NULL, data = data)
```
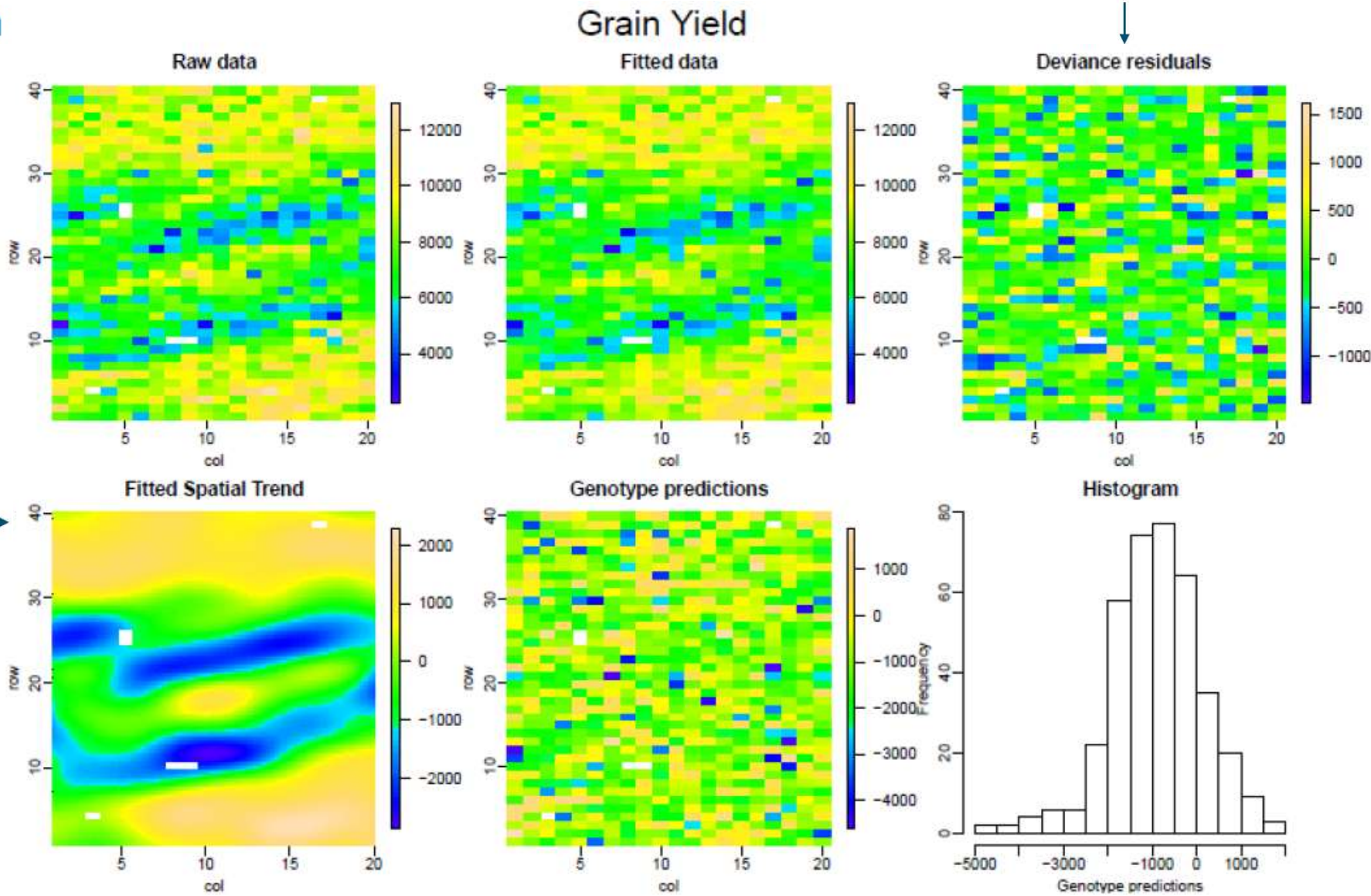
To assess the importance of each term, compare their range of variation

Yield variation that is not due to genetics and does not depend on the row and column positions

Grain Yield

Yield variation along rows and columns that is not due to genetics

Effective dimensions: a measure of the variation complexity.

Effective dimensions for the smooth part quantify the contribution of that term to the spatial variation.

E.g. variation along rows is more important than along columns $(f(x2)>f(x1))$

Fitted Spatial Trend

```
Dimensions:
              Effective    Model    Nominal    Ratio    Type
Intercept         1.0         1          1     1.00       F
Geno            264.3       383        382     0.69       R
C                 6.2        20         19     0.32       R
R                14.3        40         39     0.37       R
x1                1.0         1          1     1.00       S
x2                1.0         1          1     1.00       S
x1:x2             1.0         1          1     1.00       S
f(x1)             0.2        16         16     0.01       S
f(x2)             9.1        31         31     0.29       S
f(x1):x2          0.9         8          8     0.11       S
x1:f(x2)          1.0        16         16     0.07       S
f(x1):f(x2)      62.4       128        128     0.49       S

Total           362.4       646        643     0.56
Residual        431.6
Nobs              794

Type codes: F 'Fixed'    R 'Random'    S 'Smooth/Semiparametric'
```
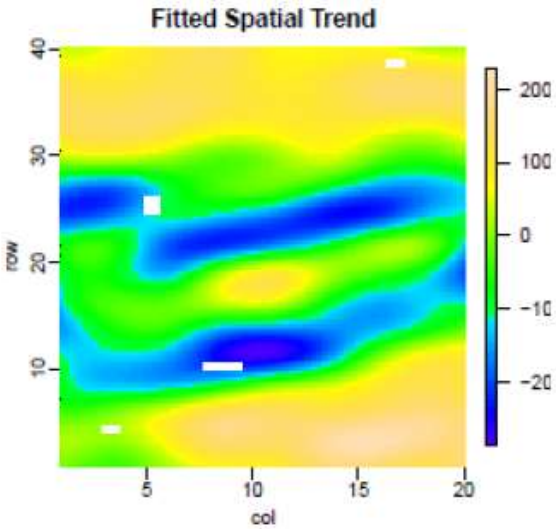
Summary spatial analysis

- Main sources of spatial field variation (that are not due to the genotype):
  - Extraneous field variation (specific rows or columns)
  - Global trend variation (trends along rows or columns)
  - Local field variation (spots)

- Spatial field variation can be separated from genetic effects with:
  - Autoregressive models
  - Spline models (SpATS)

- Both types of analysis lead to equivalent results, but SpATSmight be simpler to use (less convergence and model selection problems)

- Models are fitted twice (StatgenSTA library)
  - Genotypes as random to extract heritability
  - Genotypes as fixed to extract the BLUEs (adjusted means) for further analysis

Can Spatial Modeling Substitute for Experimental Design in Agricultural Experiments?

Alejandra Borges, Agustín González-Reymundez, Oswaldo Ernst, Mónica Cadenazzi, José Terra, and Lucía Gutiérrez★

ABSTRACT

One of the most critical aspects of agricultural experimentation is the proper choice of experimental design to control field heterogeneity, especially for large experiments. However, even with complex experimental designs, spatial variability may not be properly controlled if it occurs at scales smaller than blocks. Therefore, modeling spatial variability can be beneficial, and some studies even propose spatial modeling instead of experimental design. Our goal was to evaluate the effects of experimental design, spatial modeling, and a combination of both under real field conditions using GIS and simulating experiments. Yield data from cultivars was simulated using real spatial variability from a large uniformity trial of 100 independent locations and different sizes of experiments for four experimental designs: completely randomized design (CRD), randomized complete block design (RCBD), α-lattice incomplete block design (ALPHA), and partially replicated design (PREP). Each realization was analyzed using different levels of spatial correction. Models were compared by precision, accuracy, and the recovery of superior genotypes. For moderate and large experiment sizes, ALPHA was the best experimental design in terms of precision and accuracy. In most situations, models that included spatial correlation were better than models with no spatial correlation, but they did not outperformed better experimental designs. Therefore, spatial modeling is not a substitute for good experimental design.

Thank you.