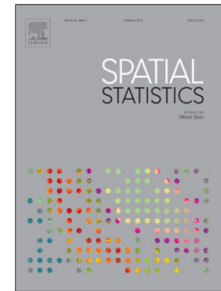


## Accepted Manuscript

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

María Xosé Rodríguez -Álvarez, Martin P. Boer, Fred A. van Eeuwijk, Paul H.C. Eilers



PII: S2211-6753(17)30107-0  
DOI: <https://doi.org/10.1016/j.spasta.2017.10.003>  
Reference: SPASTA 267

To appear in: *Spatial Statistics*

Received date : 24 March 2017  
Accepted date : 10 October 2017

Please cite this article as: Rodríguez -Álvarez M.X., Boer M.P., van Eeuwijk F.A., Eilers P.H.C., Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spatial Statistics* (2017), <https://doi.org/10.1016/j.spasta.2017.10.003>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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

María Xosé Rodríguez - Álvarez<sup>1,2</sup>, Martin P. Boer<sup>3</sup>,  
Fred A. van Eeuwijk<sup>3</sup>, Paul H. C. Eilers<sup>4</sup>

<sup>1</sup> BCAM - Basque Center for Applied Mathematics  
Alameda de Mazarredo, 14. E-48009 Bilbao, Basque Country, Spain  
[mxrodriguez@bcamath.org](mailto:mxrodriguez@bcamath.org)

<sup>2</sup> IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

<sup>3</sup> Biometris, Wageningen University & Research, Wageningen, the Netherlands

<sup>4</sup> Erasmus University Medical Centre, Rotterdam, the Netherlands

October 10, 2017

## Abstract

An important aim of the analysis of agricultural field experiments is to obtain good predictions for genotypic performance, by correcting for spatial effects. In practice these corrections turn out to be complicated, since there can be different types of spatial effects; those due to management interventions applied to the field plots and those due to various kinds of erratic spatial trends. This paper explores the use of two-dimensional smooth surfaces to model random spatial variation. We propose the use of anisotropic tensor product P-splines to explicitly model large-scale (global trend) and small-scale (local trend) spatial dependence. On top of this spatial field, effects of genotypes, blocks, replicates, and/or other sources of spatial variation are described by a mixed model in a standard way. Each component in the model is shown to have an effective dimension. They are closely related to variance estimation, and helpful for characterising the importance of model components. An important result of this paper is the formal proof of the relation between several definitions of heritability and the effective dimension associated with the genetic component. The practical value of our approach is illustrated by simulations and analyses of large-scale plant breeding experiments. An R-package, SpATS, is provided.

Keywords: Effective dimension; Field trials; Heritability; Linear mixed model; Spatial analysis; Tensor product P-splines.

# 1 Introduction

Spatial variation is common in agricultural field trials, especially in plant breeding, where large numbers of genotypes are compared. Many factors combine to generate micro-environments that differ from plot to plot, strongly influencing yield and other traits. It is necessary to correct for them when estimating treatment and/or genotypic effects. A part of the spatial variation can be attributed to systematic effects, caused by the way the field was prepared before and during sowing or planting. A familiar example are row and column effects, caused by the movements of machines during ploughing, tilling and other procedures. It is relatively easy to add factors to a statistical model to account for them. Random spatial variation, such as fertility trends, is harder to model.

Roughly speaking, there are two ways to model trends: one based on spatial variance-covariance structures, the other on smoothing techniques. In the first case, the model contains a spatially correlated stochastic component. This is non-trivial, as correlation in two directions, in the directions of the rows and columns of the field, has to be modeled. To keep the effort manageable, several simplifications are usually made, and it has become standard to assume separability and stationarity (see e.g. [Zimmerman and Harville, 1991](#)). Important contributions in this area are the separable autoregressive model proposed by [Cullis and Gleeson \(1991\)](#) and extended in [Gilmour et al. \(1997\)](#), the separable linear variance model discussed in [Piepho and Williams \(2010\)](#), or the Bayesian model with a prior based on first differences along rows and columns model ([Besag and Higdon, 1999](#)). In contrast, smoothing methods model spatial trends explicitly. Their use in the agricultural context dates back to [Green et al. \(1985\)](#), and it has been extended, among others, by [Durban et al. \(2003\)](#) and [Verbyla et al. \(1999\)](#). The main applications have been to the one-dimensional case, i.e., through separate (or additive) smoothed trend effects along the rows and columns of the field. However, while these approaches have proved to be useful for modelling large-scale dependence (or global trend), they suffer from the limitation of not always being able to capture small-scale dependence (local trend). As a consequence of this limitation, the inclusion of spatially correlated components might still be necessary ([Gilmour et al., 1997](#); [Verbyla et al., 1999](#)).

As an alternative, this paper explores the use of two-dimensional smooth surfaces. We propose the use of tensor product P-splines ([Eilers and Marx, 2003](#)) to explicitly model both sources of spatial dependence. P-splines approach smoothing as penalised regression: a rich B-spline basis is combined with a penalty on (higher order) differences of the B-spline coefficients to avoid over-fitting, and estimation is based on penalised least squares. As will be seen, the mixed model representation of P-splines ([Currie and Durban, 2002](#); [Wand, 2003](#)) provides us

with a general framework for the analysis of field trials. It allows additional fixed and random components, such as genotypic effects or the correction for rows and columns. Besides, using nested B-spline bases (Lee et al., 2013) the computational effort of our approach, which we call SpATS (Spatial Analysis of field Trials with Splines), is moderate, even for large field trials. The model has attractive properties: (1) an explicit estimate of the spatial trend in the field is obtained; (2) estimation is stable and fast; (3) missing plots, even a large number of them, are easily handled; and (4) extension to a non-normal response, along the lines of the generalised linear model, is straightforward.

Our approach is similar to the work of [Taye and Njuho \(2008\)](#) and [Robbins et al. \(2012\)](#) in field trials, and that of [Cappa and Cantet \(2008\)](#) in forest research. We go further, using an anisotropic penalty in a mixed model framework. The P-spline ANOVA (PS-ANOVA) decomposition (Lee et al., 2013), leads to a model with five smooth spatial components each having a clear interpretation. Partial effective model dimensions, a well-known complexity measure in the smoothing context ([Hastie and Tibshirani, 1990](#)), play a crucial role in the estimation of variances, but they also provide an appealing summary of the relative importance of model components. We show the remarkable link between the generalised definitions of heritability proposed by [Cullis et al. \(2006\)](#) and [Oakey et al. \(2006\)](#) and the effective dimension of the genetic component. Software is provided in the R package ([R Core Team, 2016](#)) SpATS.

The rest of the paper is structured as follows. We start with a motivating example in Section 2. Section 3 presents background on P-splines in two dimensions, including their representation as mixed models. They form the basis for spatial models, which are presented in Section 4. Simulations comparing our SpATS model and those of [Gilmour et al. \(1997\)](#) and [Piepho and Williams \(2010\)](#) can be found in Section 5. Section 6 presents several applications to large-scale plant breeding trials. Conclusion and Discussion sections close the paper. Some technical details and extra simulations are available in (Web) Appendices, where we also describe the R-package that accompanies this paper.

## 2 Motivating example

Uniformity field trials are trials in which a single genotype or variety is evaluated. In practice, the interest of such field trials is that its statistical analysis can help understanding the different sources of spatial variation present in a field, and thus serve as guidance for the design and subsequent analyses when genetic effects are to be evaluated. This section presents a series of analyses of a set of barley uniformity data discussed in the paper by [Williams and Luckett \(1988\)](#). We focus here on presenting the big picture of SpATS, leaving the more technical details to subsequent sections.

In this experiment, plots were laid out in a 15 row by 48 column grid, and the phenotypic trait of interest was yield. Figure 1(a) depicts the raw data. Note that there is a rather complex spatial pattern, with patches presenting larger/smaller yield. Let  $y_i$  denotes the yield data (in kg per hectare divided by 10) obtained at plot  $i$  ( $i = 1, \dots, 720$ ), and  $u_i$  and  $v_i$  the row and column position respectively, both centred and scaled. A common strategy in the analysis of field trial experiments is to start with a model involving only (random) effects for rows and columns. Williams and Luckett (1988) applied this model to the barley data, and noticed the presence of a three-column cycle pattern, possibly due to the seeder used in the experiment. Thus, we consider the following model as starting point

$$\mathbf{y} = \mathbf{1}_n \beta_0 + \mathbf{x}_d \beta_d + \mathbf{Z}_r \mathbf{c}_r + \mathbf{Z}_c \mathbf{c}_c + \boldsymbol{\varepsilon}, \quad (1)$$

where  $n = 720$ ,  $\mathbf{y} = (y_1, \dots, y_n)^t$ ,  $\mathbf{1}_n$  is a column vector of ones of length  $n$ , and  $\mathbf{x}_d = (x_{d1}, \dots, x_{dn})^t$ , with  $x_{di} = I(v_i \text{ modulo } 3 = 1)$ . This component is responsible for capturing the aforementioned three-column cycle pattern. The vectors  $\mathbf{c}_r = (c_{r1}, \dots, c_{r15})^t$  and  $\mathbf{c}_c = (c_{c1}, \dots, c_{c48})^t$  are, respectively, the random effect coefficients for the rows and columns with  $\mathbf{c}_r \sim N(\mathbf{0}, \sigma_r^2 \mathbf{I}_{15})$  and  $\mathbf{c}_c \sim N(\mathbf{0}, \sigma_c^2 \mathbf{I}_{48})$  and associated design matrices  $\mathbf{Z}_r$  and  $\mathbf{Z}_c$ . Finally, the random error vector  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^t \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ . Figures 1(c) and 1(d) depict the empirical best linear unbiased predictors (BLUPs) for the row and column random factors, respectively. As can be observed, and especially for the rows, the BLUPs show a clear evidence of a pattern, indicating that the independence assumption does not hold. Besides, the residuals' spatial plot shown in Figure 1(b) suggests that the spatial pattern has not been completely captured, and thus a more complex statistical analysis is required. To that end, we add a smooth bivariate surface,  $f(u, v)$ , defined over the row and column positions, i.e.,

$$\mathbf{y} = f(\mathbf{u}, \mathbf{v}) + \mathbf{x}_d \beta_d + \mathbf{Z}_r \mathbf{c}_r + \mathbf{Z}_c \mathbf{c}_c + \boldsymbol{\varepsilon}, \quad (2)$$

where  $\mathbf{u} = (u_1, \dots, u_n)^t$ ,  $\mathbf{v} = (v_1, \dots, v_n)^t$ , and  $f(\mathbf{u}, \mathbf{v}) = (f(u_1, v_1), \dots, f(u_n, v_n))^t$ . Note that the intercept,  $\beta_0$ , is embedded as part of  $f(u, v)$ . To have a better understanding of the interpretation of  $f(\cdot, \cdot)$ , we can further decompose it in a nested-type ANOVA structure (PS-ANOVA, Lee et al., 2013)

$$f(\mathbf{u}, \mathbf{v}) = \underbrace{\mathbf{1}_n \beta_0 + \mathbf{u} \beta_1 + \mathbf{v} \beta_2 + \mathbf{u} \odot \mathbf{v} \beta_3}_{\text{Bilinear polynomial}} + \underbrace{f_u(\mathbf{u}) + f_v(\mathbf{v}) + \mathbf{u} \odot h_v(\mathbf{v}) + \mathbf{v} \odot h_u(\mathbf{u}) + f_{u,v}(\mathbf{u}, \mathbf{v})}_{\text{Smooth part}},$$

where  $\odot$  denotes the element-wise vector (matrix) product. There are now two components: the bilinear polynomial and the smooth part. The bilinear (or parametric) component includes

the intercept ( $\beta_0$ ), the linear trends along the row ( $\beta_1$ ) and column ( $\beta_2$ ) directions, as well as the linear interaction trend ( $\beta_3$ ). In addition, the smooth component is responsible for modelling the deviation from this compound linear trend. Here,

- $f_u(u)$  and  $f_v(v)$  are smooth trends along the rows (columns), identical for all columns (rows), i.e., main smooth effects.
- $vh_u(u)$  and  $uh_v(v)$  are linear-by-smooth interaction trends (or varying coefficient surface trends). For instance,  $uh_v(v)$  are linear trends in the rows ( $u$ ) but with slopes ( $h_v(v)$ ) that change smoothly along the columns (the same holds for  $vh_u(u)$ ).
- $f_{u,v}(u, v)$  is a smooth-by-smooth interaction trend jointly defined over the row and column directions.

The functions  $f_u$ ,  $f_v$ ,  $h_u$  and  $h_v$  are constructed with variations on one dimensional P-splines, while  $f_{u,v}$  is based on tensor product P-splines. This decomposition translates model (2) directly to a standard mixed model (Lee et al., 2013). In fact, for each of the smooth components the desirable amount of smoothing is computed using restricted maximum likelihood (REML, Patterson and Thompson, 1971). The technical details will be presented in Section 3 and Appendix A.

Figure 1(e) shows the estimated spatial trend across the field, i.e.  $\hat{f}(\cdot, \cdot)$ , but excluding the intercept. A nice property of our proposal is that it allows depicting the spatial trend in a grid finer than the number of rows and columns, facilitating interpretation of results. Note that we recover quite successfully the spatial variation observed in the raw data. The residuals' spatial plot and BLUPs for  $c_r$  and  $c_c$ , shown in Figures 1(f), 1(g) and 1(h) respectively, suggest that the spatial independence assumption for the error vector  $\epsilon$  might be appropriate, and that no trend is now present in the BLUPs. Figure 2 shows the bilinear and smooth components of the PS-ANOVA decomposition discussed above. Note that the estimated smooth functions defined over the rows ( $\hat{f}_u$ ) and columns ( $\hat{f}_v$ ) capture the trends observed in the BLUPs for the row and column analysis (model (1) and Figures 1(c) and 1(d)). When we compare the estimated linear-by-smooth interactions trends, we observe that the contribution of  $v\hat{h}_u(u)$  to the fitted spatial trend is stronger than that due to  $u\hat{h}_v(v)$ . Finally, the smooth-by-smooth interaction term recovers the local patches observed in the raw data, that the other components would not be able to capture.

### 3 Modelling surfaces by means of P-splines

This section provides background information on P-splines, their tensor products and equivalent mixed model formulations. We avoid technicalities (some are described in Appendix A),

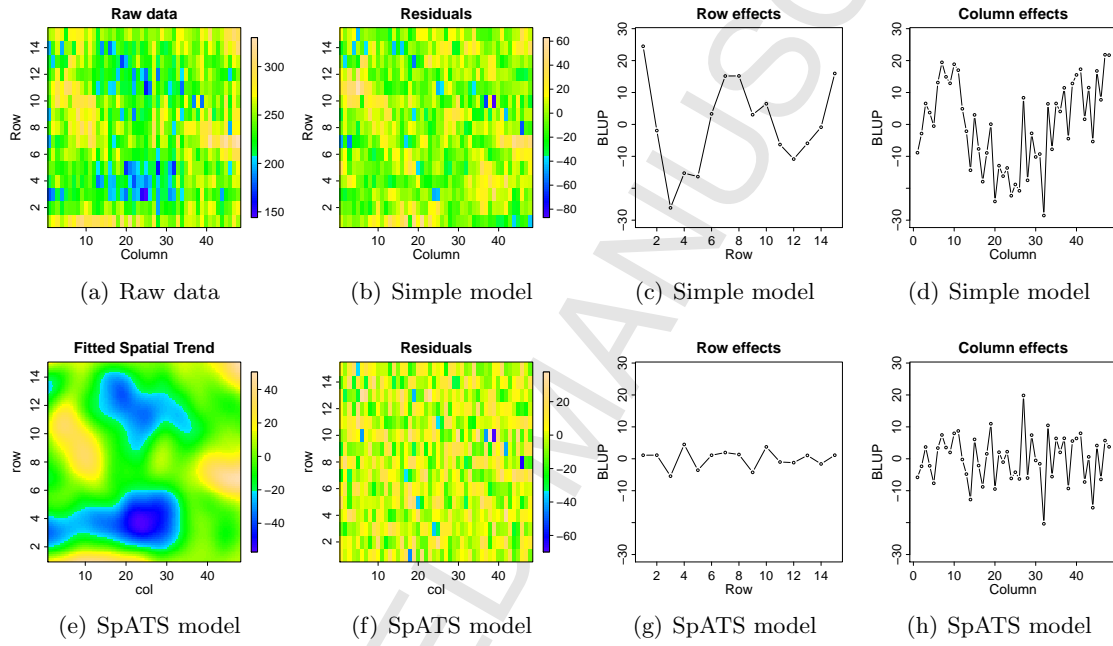


Figure 1: For the barley uniformity data: (a) Raw data, (b) and (f) residuals' spatial plot, (c) and (g) best linear unbiased predictions (BLUPs) for the row random factor, (d) and (h) BLUPs for the column random factor; and (e) contour plot of the estimated smooth spatial trend. *Simple model* refers the model including the row and column random factors and the correction for the three-column cycle (eqn. (1)) and *SpATS model* to that also including the smooth spatial trend (eqn. (2)).

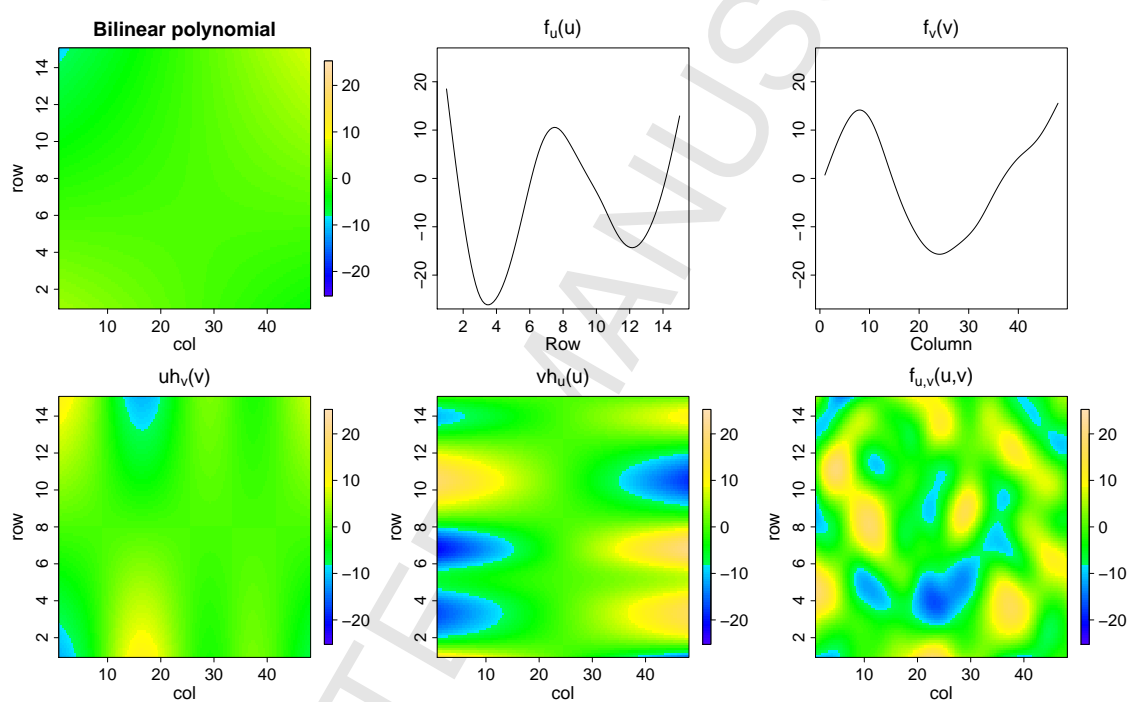


Figure 2: Bilinear and smooth components of the PS-ANOVA decomposition of the estimated spatial trend for the barley uniformity data.



and refer the interested reader to [Eilers et al. \(2015\)](#) (and references therein) for an extensive account of many aspects of P-splines.

For clarity's sake, we concentrate here on the modelling of the smooth bivariate surface of (2). A simplified model is thus considered

$$y_i = f(u_i, v_i) + \varepsilon_i, \text{ with } \varepsilon_i \sim N(0, \sigma^2). \quad (3)$$

Recall that  $y_i$  denotes the response variable, e.g. a phenotype in a trial, and  $u_i$  and  $v_i$  denote the positions, e.g. rows and columns ( $i = 1, \dots, n$ ). In the P-spline framework, the smooth bivariate function  $f(u_i, v_i)$  is modelled by the tensor product of B-spline bases ([Dierckx, 1993](#)). We form two B-spline bases,  $\hat{\mathbf{B}}$ , with  $\hat{b}_{il} = \hat{B}_l(u_i)$  and  $\check{\mathbf{B}}$ , with  $\check{b}_{ip} = \check{B}_p(v_i)$ , where  $\hat{B}_l(u_i)$  is the  $l$ -th B-spline function evaluated at  $u_i$  (and the same holds for  $\check{B}_p(v_i)$ ), and take  $f(u_i, v_i) = \sum_{l=1}^{\hat{m}} \sum_{p=1}^{\check{m}} \hat{B}_l(u_i) \check{B}_p(v_i) \alpha_{lp}$ , where  $\boldsymbol{\alpha} = (\alpha_{11}, \dots, \alpha_{1\check{m}}, \dots, \alpha_{\hat{m}\check{m}})^t$  is a vector of unknown regression coefficients of dimension  $(\hat{m}\check{m} \times 1)$ . Note that  $\hat{\mathbf{B}}$  and  $\check{\mathbf{B}}$  are matrices of order  $n \times \hat{m}$  and  $n \times \check{m}$  respectively, where  $\hat{m}$  and  $\check{m}$  are the number of the B-spline basis functions. Under this representation, model (3) can be expressed in matrix notation as

$$\mathbf{y} = \mathbf{B}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n), \quad (4)$$

where

$$\mathbf{B} = \check{\mathbf{B}} \square \hat{\mathbf{B}} = (\check{\mathbf{B}} \otimes \mathbf{1}_{\hat{m}}^t) \odot (\mathbf{1}_{\check{m}}^t \otimes \hat{\mathbf{B}}),$$

with  $\otimes$  denoting the Kronecker product and  $\square$  the ‘box’ product (the row-wise Kronecker product) ([Eilers et al., 2006](#)). To prevent over-fitting, difference penalties down the columns and along the rows are added, and estimation is based on penalised-likelihood methods ([Eilers and Marx, 2003](#)). In particular, for our application we exploit the formal similarity between P-splines and mixed models ([Currie and Durban, 2002](#); [Wand, 2003](#)), as it provides a general framework for the analysis of field trials. As described in [Appendix A](#), model (4) can be expressed as a mixed model

$$\mathbf{y} = \mathbf{X}_s \boldsymbol{\beta}_s + \mathbf{Z}_s \mathbf{c}_s + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n) \text{ and } \mathbf{c}_s \sim N(\mathbf{0}, \mathbf{G}_s),$$

where  $\mathbf{X}_s = [\mathbf{1}_n | \mathbf{u} | \mathbf{v} | \mathbf{u} \odot \mathbf{v}]$ ,  $\boldsymbol{\beta}_s = (\beta_0, \beta_1, \beta_2, \beta_3)^t$ ,  $\mathbf{Z}_s = [\mathbf{Z}_v | \mathbf{Z}_u | \mathbf{Z}_v \square \mathbf{u} | \mathbf{v} \square \mathbf{Z}_u | \mathbf{Z}_v \square \mathbf{Z}_u]$ , and  $\mathbf{c}_s = (\mathbf{c}_{s1}^t, \dots, \mathbf{c}_{s5}^t)^t$ , with  $\mathbf{c}_{sk}$  ( $k = 1, \dots, 5$ ) containing the elements of  $\mathbf{c}_s$  that correspond to the  $k$ -th block of  $\mathbf{Z}_s$ . Finally,  $\mathbf{G}_s = \bigoplus_{k=1}^5 \sigma_{sk}^2 \boldsymbol{\Lambda}_{sk}$ . The specific form of  $\mathbf{Z}_u$ ,  $\mathbf{Z}_v$ , and  $\boldsymbol{\Lambda}_{sk}$  is given in [Appendix A](#).

The previous reformulation of the tensor product P-spline model as a mixed model directly

provides the interesting PS-ANOVA decomposition discussed in Section 2 (Lee et al., 2013)

$$f(\mathbf{u}, \mathbf{v}) = \mathbf{X}_s \boldsymbol{\beta}_s + \mathbf{Z}_s \mathbf{c}_s$$

$$= \underbrace{\mathbf{1}_n \beta_0 + \mathbf{u} \beta_1 + \mathbf{v} \beta_2 + \mathbf{u} \odot \mathbf{v} \beta_3}_{\mathbf{X}_s \boldsymbol{\beta}_s} + \underbrace{f_v(\mathbf{v})}_{\mathbf{Z}_v \mathbf{c}_{s1}} + \underbrace{f_u(\mathbf{u})}_{\mathbf{Z}_u \mathbf{c}_{s2}} + \underbrace{\mathbf{u} \odot h_v(\mathbf{v})}_{[\mathbf{Z}_v \square \mathbf{u}] \mathbf{c}_{s3}} + \underbrace{\mathbf{v} \odot h_u(\mathbf{u})}_{[\mathbf{v} \square \mathbf{Z}_u] \mathbf{c}_{s4}} + \underbrace{f_{u,v}(\mathbf{u}, \mathbf{v})}_{[\mathbf{Z}_v \square \mathbf{Z}_u] \mathbf{c}_{s5}}.$$

The amount of smoothing for each function  $f_u$ ,  $f_v$ ,  $h_u$ ,  $h_v$  and  $f_{u,v}$  is determined by the variance parameter  $\sigma_{sk}^2$  associated with the corresponding random vector  $\mathbf{c}_{sk}$ . We note that the dimension of  $\mathbf{c}_{sk}$  ( $k = 1, \dots, 5$ ) depends on  $\hat{m}$  and  $\check{m}$ , the dimensions of the B-spline basis used for the tensor product. Web Appendix A contains a proposal to use smaller, nested, B-spline bases (Lee et al., 2013) that allow reducing the dimension associated with the smooth-by-smooth interaction term  $f_{u,v}$ , and, as a consequence, the computational effort.

## 4 Spatial models for field trials

The tensor product P-spline presented in Section 3 constitutes the basis for the analysis of agricultural field trials. On top of this spatial field, we build more complex models, to account for the genetic variation, the presence of block and/or replication effects, or other sources of spatial variation like those due to the way the field was prepared. From now on, we therefore consider the following linear mixed model

$$\mathbf{y} = \underbrace{\mathbf{X}_s \boldsymbol{\beta}_s + \mathbf{Z}_s \mathbf{c}_s}_{f(\mathbf{u}, \mathbf{v})} + \mathbf{X}_d \boldsymbol{\beta}_d + \mathbf{Z}_d \mathbf{c}_d + \boldsymbol{\varepsilon}, \text{ with } \mathbf{c}_s \sim N(\mathbf{0}, \mathbf{G}_s) \text{ and } \mathbf{c}_d \sim N(\mathbf{0}, \mathbf{G}_d), \quad (5)$$

where  $\mathbf{X}_s$ ,  $\mathbf{Z}_s$  and  $\mathbf{G}_s$  are defined in (15), (16) and (18) in Appendix A.  $\mathbf{X}_d$  and  $\mathbf{Z}_d$  represent column-partitioned matrices, associated respectively with extra fixed and random components, as for instance, row, column, replicate and/or genotypic effects. We assume that  $\mathbf{X}_d$  has full rank,  $\mathbf{Z}_d = [\mathbf{Z}_{d1} | \dots | \mathbf{Z}_{db}]$ , and  $\mathbf{c}_d = (\mathbf{c}_{d1}^t, \dots, \mathbf{c}_{db}^t)^t$ . Each  $\mathbf{Z}_{dk}$  corresponds to the design matrix of the  $k$ -th random component  $\mathbf{c}_{dk}$ , with  $\mathbf{c}_{dk}$  being a  $(m_{dk} \times 1)$  vector ( $k = 1, \dots, b$ ). We assume further that  $\mathbf{c}_s$  and  $\mathbf{c}_d$  are independent, and that  $\mathbf{G}_d = \bigoplus_{k=1}^b \mathbf{G}_{dk} = \bigoplus_{k=1}^b \sigma_{dk}^2 \boldsymbol{\Lambda}_{dk}$ , where  $\boldsymbol{\Lambda}_{dk}$  are diagonal matrices. To keep the notation simple, we rewrite model (5) as

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \mathbf{c} + \boldsymbol{\varepsilon}, \text{ with } \mathbf{c} \sim N(\mathbf{0}, \mathbf{G}) \text{ and } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n), \quad (6)$$

where  $\mathbf{X} = [\mathbf{X}_s | \mathbf{X}_d]$ ,  $\mathbf{Z} = [\mathbf{Z}_s | \mathbf{Z}_d] = [\mathbf{Z}_1 | \dots | \mathbf{Z}_q]$ ,  $\mathbf{c} = (\mathbf{c}_s^t, \mathbf{c}_d^t)^t = (\mathbf{c}_1^t, \dots, \mathbf{c}_q^t)^t$  ( $q = 5 + b$ ), and

$$\mathbf{G} = \text{blockdiag}(\mathbf{G}_s, \mathbf{G}_d) = \bigoplus_{k=1}^q \mathbf{G}_k = \bigoplus_{k=1}^q \sigma_k^2 \boldsymbol{\Lambda}_k. \quad (7)$$

#### 4.1 SpATS model estimation

Estimates of the fixed and random effect coefficients in model (6), for given values of the variance components, follow from standard mixed-model theory, and variance components can be obtained, as usual, by maximising the REML log-likelihood function. For the sake of brevity, the detailed description of the estimation procedure implemented in the R-package **SpATS** that accompanies this paper is presented in Web Appendix B. However, we present here the expression of the REML estimates of the variance components. The reason is the interesting link between the denominator of these expressions and the notion of effective dimension that is discussed in detail in Section 4.2.

As said, REML estimates of the variance components are obtained by maximising the REML log-likelihood function. Taking derivatives of the REML log-likelihood function with respect to the variance components  $\sigma_k^2$  ( $k = 1, \dots, q$ ), and equating the obtained expressions to zero (see Web Appendix B for more details), we obtain that

$$\hat{\sigma}_k^2 = \frac{\hat{\mathbf{c}}_k^t \mathbf{\Lambda}_k^{-1} \hat{\mathbf{c}}_k}{\text{ED}_k}, k = 1, \dots, q, \quad (8)$$

with

$$\text{ED}_k = \text{trace}(\mathbf{Z}_k^t \mathbf{Q} \mathbf{Z}_k \mathbf{G}_k).$$

Here  $\mathbf{Q} = \mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1}$  with  $\mathbf{V} = \mathbf{R} + \mathbf{Z} \mathbf{G} \mathbf{Z}^t$  and  $\mathbf{R} = \sigma^2 \mathbf{I}_n$ . An estimate of  $\sigma^2$  can also be easily obtained following the same reasoning

$$\hat{\sigma}^2 = \frac{\hat{\boldsymbol{\varepsilon}}^t \hat{\boldsymbol{\varepsilon}}}{\text{ED}_{\boldsymbol{\varepsilon}}}, \quad (9)$$

where  $\hat{\boldsymbol{\varepsilon}} = \mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}} - \mathbf{Z} \hat{\mathbf{c}}$  and

$$\text{ED}_{\boldsymbol{\varepsilon}} = \text{trace}(\mathbf{R} \mathbf{Q}) = n - \text{rank}(\mathbf{X}) - \sum_{k=1}^q \text{ED}_k. \quad (10)$$

#### 4.2 Effective dimensions

Denoting the denominator of (8) and (9) as  $\text{ED}_{\{\cdot\}}$  (for “effective dimension”) has not been done without purpose. In the smoothing context, the notion of effective dimension or effective degrees of freedom is well known (see, e.g., [Hastie and Tibshirani, 1990](#)). The effective dimension (denoted by ED) of a “smooth” model is defined as the trace of the so-called “hat” matrix  $\mathbf{H}$ , defined as  $\hat{\mathbf{y}} = \mathbf{H} \mathbf{y}$ . In this setting, ED can be interpreted as a measure of the complexity of the model: the larger the ED, the more complex (or the less smooth) the model

(see also [Eilers et al., 2015](#)).

In recent years, several new definitions and generalisations of the concept of effective dimension that are applicable to (generalised) linear mixed models have been proposed in the statistical literature (see, e.g., [You et al., 2016](#), and references therein). In almost all cases, the aim has been to provide a complexity measure that allows model comparison and selection (via, for instance, the Akaike's Information Criterion). However, we are more interested in obtaining a separate complexity measure for each component in model (6), that can give us insights about the contribution of that effect when explaining the response (phenotypic) variation. This issue has been already discussed by [Cui et al. \(2010\)](#). In that paper the authors define the effective dimension of a model's component as the trace of the ratio of the component modelled variance matrix to the total variance matrix.

In line with previous work in the smoothing context (e.g., [Hastie and Tibshirani, 1990](#); [Ruppert et al., 2003](#)), this paper considers defining the effective dimension of a model's component as the trace of the corresponding hat matrix. As shown in Web Appendix C, for the linear mixed model (6) we have that

$$\mathbf{H}\mathbf{y} = \hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\mathbf{c}} = \mathbf{H}_F\mathbf{y} + \mathbf{H}_R\mathbf{y} = \mathbf{H}_F\mathbf{y} + \sum_{k=1}^q \mathbf{H}_k\mathbf{y},$$

where  $\mathbf{H}_F = \mathbf{X}(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{V}^{-1}$  and  $\mathbf{H}_R = \mathbf{Z}\mathbf{G}\mathbf{Z}^t\mathbf{Q}$  are, respectively, the “hat” matrices of the fixed and the random parts of the model, and

$$\mathbf{H}_k = \mathbf{Z}_k\mathbf{G}_k\mathbf{Z}_k^t\mathbf{Q} \quad (k = 1, \dots, q),$$

is the “hat” matrix corresponding to the  $k$ -th random component  $\mathbf{c}_k$  (either a component of the PS-ANOVA spatial field or a “pure” random factor). The effective dimension for  $\mathbf{c}_k$  is then defined as the trace of  $\mathbf{H}_k$ . Web Appendix D shows the equivalence between this definition and that by [Cui et al. \(2010\)](#). Using trace properties, we have

$$\text{trace}(\mathbf{H}_k) = \text{trace}(\mathbf{Z}_k\mathbf{G}_k\mathbf{Z}_k^t\mathbf{Q}) = \text{trace}(\mathbf{Z}_k^t\mathbf{Q}\mathbf{Z}_k\mathbf{G}_k) = \text{ED}_k,$$

and the total effective dimension of model (6) is thus decomposed as the sum of independent

contributions

$$\begin{aligned} \text{ED} &= \text{trace}(\mathbf{H}) = \text{trace}(\mathbf{H}_F) + \text{trace}(\mathbf{H}_R) = \text{trace}(\mathbf{H}_F) + \sum_{k=1}^q \text{trace}(\mathbf{H}_k) \\ &= \text{rank}(\mathbf{X}) + \sum_{k=1}^q \text{ED}_k. \end{aligned}$$

Besides, result  $\hat{\boldsymbol{\varepsilon}} = \mathbf{H}_\varepsilon \mathbf{y} = \mathbf{RQy}$  (see Web Appendix C) also suggests defining the effective dimension for the residuals as the trace of  $\mathbf{H}_\varepsilon$ , and thus,  $\text{trace}(\mathbf{H}_\varepsilon) = \text{trace}(\mathbf{RQ}) = \text{ED}_\varepsilon$  (see also eqn. (10)). Note that

$$n = \text{rank}(\mathbf{X}) + \sum_{k=1}^q \text{ED}_k + \text{ED}_\varepsilon.$$

Thus, the number of observations  $n$  is partitioned into independent effective dimensions for the model's components and error. This is in concordance with the definition given by Cui et al. (2010) (see Web Appendix D). As extensively discussed by the authors,  $\text{ED}_\varepsilon$  avoids the problem that is posed by most of the traditional definitions for the residual effective dimension in the smoothing context, that is to say, that  $\text{ED} + \text{ED}_\varepsilon$  is not equal to the number of observations  $n$ .

For all components  $\text{ED}_k$  ( $k = 1, \dots, q$ ) will vary between 0 and  $(m_k - \zeta_k)$ , where  $\zeta_k$  is the number of zero eigenvalues of  $\mathbf{Z}_k^t \mathbf{Q} \mathbf{Z}_k \mathbf{G}_k$ . In Cui et al. (2010) the authors show that this upper bound,  $(m_k - \zeta_k)$ , can also be expressed as:  $\text{rank}([\mathbf{X}, \mathbf{Z}_k]) - \text{rank}(\mathbf{X})$ . The signal-to-noise ratio  $\sigma_k^2/\sigma^2$  modulates the value of  $\text{ED}_k$ : when  $\sigma_k^2/\sigma^2 \rightarrow 0$  then  $\text{ED}_k \rightarrow 0$ ; and when  $\sigma_k^2/\sigma^2 \rightarrow \infty$ , then  $\text{ED}_k \rightarrow (m_k - \zeta_k)$ . Arguably,  $\text{ED}_k$  can therefore be interpreted as a measure of the complexity of the corresponding component, which in turns gives a separate measure of its contribution. A value of zero would indicate that this component does not contribute to the response variability. What is more, when the genetic effect is included in model (6) as random, the associated (scaled) effective dimension corresponds to the generalised heritability proposed by Oakey et al. (2006). This equivalence emphasises interpreting  $\text{ED}_k$  as a measure of the shrinkage. A formal derivation and deeper discussion is provided in next section.

### 4.3 Heritability and effective dimension

To introduce the standard definition of heritability, let us start with the classical quantitative genetic model. Here,  $m_g$  genotypes, each replicated  $r$  times, are evaluated and no other model components (either spatial, fixed or random) are considered, that is,  $\mathbf{y} = \mathbf{1}_n \beta_0 + \mathbf{Z}_g \mathbf{c}_g + \boldsymbol{\varepsilon}$ , where  $\mathbf{Z}_g = \mathbf{I}_{m_g} \otimes \mathbf{1}_r$ ,  $\mathbf{c}_g \sim N(\mathbf{0}, \sigma_g^2 \mathbf{I}_{m_g})$  and  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ . Under this model, the standard

heritability measure, with respect to genotypic means, is defined as

$$H_s^2 = \frac{\sigma_g^2}{(\sigma_g^2 + \sigma^2/r)}. \quad (11)$$

That is to say, the standard heritability is the proportion of the total (phenotypic) variation that is attributable to the genetic component.

We turn now to the notion of effective dimension defined above. For the genetic effects  $\mathbf{c}_g$ , the associated effective dimension is  $\text{ED}_g = \text{trace}(\mathbf{Z}_g^t \mathbf{Q} \mathbf{Z}_g \mathbf{G}_g)$ , where  $\mathbf{G}_g = \sigma_g^2 \mathbf{I}_{m_g}$ . By eqn. (3.7) in [Harville \(1977\)](#), we have

$$\mathbf{Z}_g^t \mathbf{Q} \mathbf{Z}_g \mathbf{G}_g = \left( \frac{1}{\lambda} \mathbf{I}_{m_g} + \mathbf{Z}_g^T \mathbf{S} \mathbf{Z}_g \right)^{-1} \mathbf{Z}_g^T \mathbf{S} \mathbf{Z}_g,$$

where  $\lambda = \frac{\sigma_g^2}{\sigma^2}$  and  $\mathbf{S} = \mathbf{I}_n - \mathbf{1}_n(\mathbf{1}_n^t \mathbf{1}_n)^{-1} \mathbf{1}_n^t$ . It can be shown that  $\mathbf{Z}_g^T \mathbf{S} \mathbf{Z}_g$  has  $(m_g - 1)$  eigenvalues equal to  $r$  (the number of replicates), and 1 eigenvalue equal to zero. Moreover,  $\frac{1}{\lambda} \mathbf{I}_{m_g}$  has  $m_g$  eigenvalues all equal to  $\frac{1}{\lambda}$ . As a consequence

$$\text{ED}_g = \sum_{i=1}^{m_g} \lambda_i = \sum_{i=1}^{(m_g-1)} \left( \frac{1}{\lambda} + r \right)^{-1} r = (m_g - 1) \frac{\sigma_g^2}{(\sigma_g^2 + \sigma^2/r)},$$

where  $\lambda_i$  are the eigenvalues of  $\mathbf{Z}_g^t \mathbf{Q} \mathbf{Z}_g \mathbf{G}_g$ . Hence, there is direct link between the standard heritability measure  $H_s^2$  defined in (11) and the genetic effective dimension, with

$$H_s^2 = \frac{\text{ED}_g}{m_g - 1}.$$

When the statistical analysis of a field trial experiment involves modelling more sources of variation (as, e.g., spatial and/or extraneous variation), the standard definition of heritability given in (11) does not longer apply, and several generalisations have been proposed in the literature. For instance, [Cullis et al. \(2006\)](#) present a generalised definition of heritability – applicable whenever  $\mathbf{G}_g = \sigma_g^2 \mathbf{I}_{m_g}$  – based on the prediction error variance of genetic effects

$$H_c^2 = 1 - \frac{1}{\sigma_g^2} \sum_{i=1}^{m_g} \frac{\text{pev}(c_{gi})}{m_g},$$

where  $\text{pev}(c_{gi}) = \text{var}(\hat{c}_{gi} - c_{gi})$ . Using the equivalence

$$\mathbf{Z}_g^t \mathbf{Q} \mathbf{Z}_g \mathbf{G}_g = \mathbf{I}_{m_g} - \mathbf{G}_g^{-1} \mathbf{C}_{gg}^* = \mathbf{I}_{m_g} - \frac{1}{\sigma_g^2} \mathbf{C}_{gg}^*,$$

where  $\mathbf{C}^* = \mathbf{C}^{-1}$  (see Web Appendix B) and  $\mathbf{C}_{gg}^*$  is that partition of  $\mathbf{C}^*$  corresponding to  $\mathbf{c}_g$ , and noting that  $\mathbf{C}_{gg}^*$  corresponds to  $\text{Cov}(\hat{\mathbf{c}}_g - \mathbf{c}_g)$  (i.e., the prediction error covariance matrix for the genetic effects), the generalised heritability measure proposed by Cullis et al. (2006) can also be expressed in terms of the genetic effective dimension

$$H_c^2 = 1 - \frac{1}{\sigma_g^2} \sum_{i=1}^{m_g} \frac{pev(c_{gi})}{m_g} = \frac{\text{trace}(\mathbf{I}_{m_g} - \frac{1}{\sigma_g^2} \mathbf{C}_{gg}^*)}{m_g} = \frac{\text{trace}(\mathbf{Z}_g^t \mathbf{Q} \mathbf{Z}_g \mathbf{G}_g)}{m_g} = \frac{\text{ED}_g}{m_g}.$$

In Oakey et al. (2006) a more general definition of heritability is presented that can be used regardless of the structure of genetic covariance matrix  $\mathbf{G}_g$ . As can be seen in eqn. (7) of that paper, the definition the authors propose turns out to be

$$H_g^2 = \frac{\text{ED}_g}{m_g - \zeta_g}. \quad (12)$$

As said before, the denominator of (12) represents the upper bound of the genetic effective dimension, and  $H_g^2$  can therefore be interpreted as a shrinkage factor. The generalised heritability proposed by Cullis et al. (2006) would be a special case, ignoring the number of zero eigenvalues.

## 5 Simulation study

This section presents the results of a simulation study performed to evaluate the behaviour of our SpATS model under controlled scenarios, and its comparison with the separable autoregressive (AR $\times$ AR) model proposed by Gilmour et al. (1997). In the context of single-trial experiments, this proposal has become the standard modelling strategy, specially among applied breeders, and therefore it has been chosen as the benchmark model. However, as extensively discussed by Piepho et al. (2015), the AR $\times$ AR model may present severe convergence problems, and the authors recommend as an alternative the use of the linear variance (LV) models proposed by Williams (1986) and extended in two dimensions by Williams et al. (2006) and Piepho and Williams (2010). Thus, comparisons with the separable LV $\times$ LV model also have been included in this study. All simulations were done using the R-packages SpATS and ASrem1-R. Extra simulation results can be found in Web Appendix E. The R-code used can be downloaded from <http://bitbucket.org/mxrodriguez/spats-paper>.

To evaluate the SpATS approach under an a priori unfavourable scenario, data was generated assuming a separable Gaussian AR process of order 1 in the row and column directions.

The following model was considered

$$\mathbf{y} = \mathbf{Z}_g \mathbf{c}_g + \underbrace{\boldsymbol{\xi} + \boldsymbol{\varepsilon}}_{\boldsymbol{\epsilon}}, \quad (13)$$

where  $\mathbf{c}_g$  denotes the genotypic effects, with  $\mathbf{c}_g \sim N(\mathbf{0}, \sigma_g^2 \mathbf{I}_{m_g})$ .  $\boldsymbol{\xi}$  is a spatially dependent random vector with  $\text{cov}(\xi_l, \xi_p) = \sigma_s^2 \rho_r^{|u_l - u_p|} \rho_c^{|v_l - v_p|}$ , where  $\rho_r$  and  $\rho_c$  are the autocorrelation parameters for row and column, respectively. Finally,  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$  and  $\boldsymbol{\epsilon} = \boldsymbol{\xi} + \boldsymbol{\varepsilon}$ . We note that in the geostatistics literature,  $\boldsymbol{\varepsilon}$  is usually referred to as measurement error or nugget effect.

Different levels of genetic variation ( $\sigma_g^2 \in \{0.25; 1; 4\}$ ) and autocorrelation ( $\rho_r = \rho_c \in \{0.1; 0.5; 0.9; 0.99\}$ ) were studied. In all cases,  $\sigma^2 = \sigma_s^2 = 1$  and  $m_g = 100$  genotypes, each replicated twice, were allocated to the plots following an alpha design, in blocks of size 10. The R-package `agricolae` was used for that purpose. The field layout thus consisted of 10 blocks (rows) and 20 columns.

We fitted our SpATS model including the PS-ANOVA spatial field and the genetic random factor. For [Gilmour et al. \(1997\)](#)'s approach we considered a model including the genetic random factor, the AR $\times$ AR process and the nugget ( $\boldsymbol{\varepsilon}$ ). Note that this model corresponds to the true model generating the data (see (13)). Finally, for the separable linear variance model we included the row and column random factors (which are part of the baseline model), the genetic random factor, the LV $\times$ LV process (see eqn. (10) in [Piepho and Williams \(2010\)](#)) and the nugget. The procedure was repeated a total of 1000 times. For SpATS, cubic B-spline bases of dimension  $\hat{m} = 13$  and  $\hat{m} = 23$  were assumed for the row and column positions, respectively, and nested B-spline bases, with half the dimension, were used. To measure performance, the discrepancy between the BLUPs for the genotypic effects and the corresponding true (simulated) quantities was measured in terms of the empirical version of the global root mean squared error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{100} \sum_{i=1}^{100} (\hat{c}_{gi} - c_{gi})^2}.$$

As far as the REML estimates of  $\sigma_g^2$ ,  $\sigma^2$  and  $\sigma_s^2$  is concerned, the behaviour was evaluated in terms of the bias.

Table 1 shows the results for those runs for which the AR $\times$ AR model converged. If we focus on the genetic random factor, all approaches behave similarly for all scenarios considered. As expected, the RMSE decreases as the genotypic signal increases. A similar performance of all approaches is also observed for  $\sigma_g^2$ . The conclusion is that, with respect to genotype



predictions, both SpATS and the LV×LV model perform as well as the true model.

The interpretation of the results for  $\sigma^2$ ,  $\sigma_s^2$  and  $ED_s$  as well as comparisons among all approaches require a more detailed analysis. First note that when both  $\rho_r$  and  $\rho_c \rightarrow 1$ , then  $\text{Cov}(\boldsymbol{\xi}) \rightarrow \sigma_s^2 \mathbb{I}_{200}$ , where  $\mathbb{I}_n$  is a  $n \times n$  matrix of ones. In this case, for each realisation of the stochastic process (13), the spatially dependent random error  $\boldsymbol{\xi}$  would be a constant vector, and no spatial variation would be therefore present. On the other hand, when both  $\rho_r$  and  $\rho_c \rightarrow 0$  then  $\text{Cov}(\boldsymbol{\xi}) \rightarrow \sigma_s^2 \mathbf{I}_{200}$ . Thus,  $\boldsymbol{\xi}$  is confounded with the nugget or independent error  $\boldsymbol{\epsilon}$ , and  $\text{Cov}(\boldsymbol{\epsilon}) \rightarrow (\sigma_s^2 + \sigma^2) \mathbf{I}_{200}$ . To some extent, it would also imply that no spatial variation is present and that all is measurement error. As pointed out by Piepho et al. (2015), these two extreme cases may cause convergence problems when fitting the AR×AR model. Our results are in concordance with those previous findings, and we refer to Piepho et al. (2015) for a more comprehensive discussion. Besides, it could also explain the systematic bias (of opposite sign but similar magnitude, except for very large autocorrelation) in the estimates of  $\sigma_s^2$  and  $\sigma^2$  provided by the AR×AR model. The AR×AR model tends to underestimate  $\sigma^2$  (negative bias), and this is counteracted by overestimating  $\sigma_s^2$ . This effect is especially remarkable for medium and low autocorrelations.

With respect to SpATS, for large autocorrelation, it provides unbiased estimates of  $\sigma^2$ . When the autocorrelation decreases, the bias increases. It is worth remembering that in this case  $\text{Cov}(\boldsymbol{\epsilon}) \rightarrow (\sigma_s^2 + \sigma^2) \mathbf{I}_{200}$  (all is measurement error). In all simulations  $\sigma_s^2 = 1$ , and the bias in  $\hat{\sigma}^2$  approaches this value as  $\rho_r$  and  $\rho_c$  decrease. Thus, our SpATS model is correct when estimating the variance associated with the measurement error. The estimated effective dimension for the PS-ANOVA spatial field also reflects the expected performance: for low and large autocorrelations (small spatial variation) the model provides smaller  $ED_s$  than for medium values. Finally, it is interesting to note that both SpATS and the LV×LV model present a very similar behaviour for all measures of performance considered. As indicated in the discussion of Piepho and Williams (2010), the LV×LV model bears close relationships to the first-differencing model in rows and columns proposed by Besag and Higdon (1999), which in turn is closely related to our SpATS model. We discuss this connection in more detail in Section 8.

## 6 Applications

In this section we analyse two data sets, one of a field trial on wheat conducted in Chile and discussed in the paper by Lado et al. (2013), and the other on sugar beet from a big field in France.

Table 1: Numerical results associated with the simulation study when the underlying simulated model follows an AR×AR process. For the REML estimates of  $\sigma_g^2$ ,  $\sigma^2$  and  $\sigma_s^2$  the results show the bias. For the genotypic BLUPs, the  $\log_{10}(\text{RMSE})$  is shown.  $\text{ED}_s$  denotes the effective dimension associated with the PS-ANOVA spatial field (excluding the intercept). In all cases, averages and standard deviations are presented. The results correspond to those runs for which the AR×AR model converged.

$\sigma_g^2$	$\rho_c = \rho_r$	Model	Conv. (%)	Genotypic RMSE	$\sigma_g^2$	$\sigma^2$	$\sigma_s^2$	$\text{ED}_s$
0.25	0.99	SpATS	100.0	-0.380 (0.031)	-0.015 (0.127)	-0.019 (0.149)	-	5.91 (2.33)
		AR×AR	57.1	-0.377 (0.032)	-0.021 (0.133)	-0.480 (0.450)	-0.329 (0.537)	-
		LV×LV	100.0	-0.380 (0.032)	-0.016 (0.129)	-0.037 (0.151)	-	-
	0.9	SpATS	100.0	-0.355 (0.029)	-0.008 (0.135)	0.086 (0.162)	-	11.23 (3.46)
		AR×AR	83.0	-0.356 (0.030)	-0.012 (0.138)	-0.081 (0.237)	0.375 (1.442)	-
		LV×LV	100.0	-0.370 (0.031)	-0.017 (0.138)	0.006 (0.165)	-	-
	0.5	SpATS	100.0	-0.345 (0.031)	-0.001 (0.178)	0.566 (0.247)	-	13.12 (5.24)
		AR×AR	97.1	-0.346 (0.031)	-0.015 (0.176)	-0.222 (0.485)	0.244 (0.493)	-
		LV×LV	100.0	-0.344 (0.031)	-0.010 (0.185)	0.544 (0.260)	-	-
	0.1	SpATS	100.0	-0.340 (0.032)	0.000 (0.189)	0.903 (0.273)	-	5.55 (2.48)
		AR×AR	83.0	-0.338 (0.033)	-0.011 (0.195)	-0.513 (0.746)	0.510 (0.730)	-
		LV×LV	100.0	-0.338 (0.033)	-0.006 (0.197)	0.891 (0.272)	-	-
1	0.99	SpATS	100.0	-0.231 (0.030)	-0.081 (0.182)	-0.018 (0.153)	-	5.67 (2.18)
		AR×AR	56.3	-0.228 (0.030)	-0.082 (0.192)	-0.565 (0.434)	-0.277 (0.516)	-
		LV×LV	100.0	-0.233 (0.031)	-0.076 (0.180)	-0.043 (0.152)	-	-
	0.9	SpATS	100.0	-0.211 (0.030)	-0.075 (0.198)	0.084 (0.180)	-	10.44 (3.34)
		AR×AR	78.8	-0.213 (0.030)	-0.078 (0.194)	-0.123 (0.272)	0.370 (1.440)	-
		LV×LV	100.0	-0.214 (0.030)	-0.071 (0.196)	0.001 (0.173)	-	-
	0.5	SpATS	100.0	-0.161 (0.030)	-0.079 (0.260)	0.595 (0.269)	-	11.62 (4.73)
		AR×AR	93.6	-0.167 (0.030)	-0.085 (0.253)	-0.264 (0.520)	0.302 (0.503)	-
		LV×LV	100.0	-0.162 (0.030)	-0.082 (0.261)	0.576 (0.279)	-	-
	0.1	SpATS	100.0	-0.148 (0.030)	-0.085 (0.281)	0.920 (0.300)	-	5.43 (2.33)
		AR×AR	84.6	-0.147 (0.030)	-0.083 (0.291)	-0.512 (0.795)	0.544 (0.722)	-
		LV×LV	100.0	-0.149 (0.030)	-0.088 (0.278)	0.906 (0.297)	-	-
4	0.99	SpATS	100.0	-0.158 (0.030)	-0.306 (0.305)	-0.026 (0.154)	-	5.52 (2.05)
		AR×AR	57.8	-0.155 (0.031)	-0.309 (0.334)	-0.619 (0.409)	-0.217 (0.720)	-
		LV×LV	100.0	-0.162 (0.030)	-0.297 (0.305)	-0.049 (0.152)	-	-
	0.9	SpATS	100.0	-0.130 (0.031)	-0.298 (0.336)	0.085 (0.182)	-	9.49 (3.19)
		AR×AR	74.7	-0.133 (0.032)	-0.303 (0.336)	-0.184 (0.317)	0.425 (1.620)	-
		LV×LV	100.0	-0.135 (0.030)	-0.302 (0.336)	-0.001 (0.179)	-	-
	0.5	SpATS	100.0	-0.061 (0.031)	-0.312 (0.421)	0.604 (0.279)	-	10.14 (4.23)
		AR×AR	90.1	-0.069 (0.031)	-0.321 (0.407)	-0.355 (0.540)	0.395 (0.548)	-
		LV×LV	100.0	-0.062 (0.031)	-0.314 (0.415)	0.589 (0.286)	-	-
	0.1	SpATS	100.0	-0.043 (0.031)	-0.319 (0.444)	0.916 (0.303)	-	5.20 (2.05)
		AR×AR	82.5	-0.043 (0.031)	-0.302 (0.537)	-0.585 (0.720)	0.615 (0.678)	-
		LV×LV	100.0	-0.046 (0.031)	-0.323 (0.438)	0.890 (0.298)	-	-

## 6.1 Sugar beet data, France

The field experiment on sugar beet was performed in France, in 2011. The results will be presented for the trait  $\alpha$ -amino nitrogen content (in millimole per litre). All material was kindly provided by the breeding company SESVanderHave (Tienen, Belgium).

The field experiment consisted of 31 trials, with a total of 2411 plots and 1095 genotypes. All the trials contained four common checks, the other 1091 genotypes were observed in one of the trials. The trials were different in size, with the most common layout an alpha design with 36 genotypes and two replicates. Four of the trials were unreplicated, 26 trials had an alpha design with two replicates, and one trial had an alpha design with three replicates.

The field layout is shown in Figure 3(a). The 31 trials are represented by a different colour. Note that the layout of the field is irregular, that there are no trials in the lower right corner (white areas). Nonetheless, no data manipulation (e.g., filling in the missing values to make the layout regular) is required to use SpATS.

For this experiment, we assumed a model including random factors for rows ( $\mathbf{c}_r$ ) columns ( $\mathbf{c}_c$ ) and trials ( $\mathbf{c}_t$ ). The candidate genotypes were also included in the model as random ( $\mathbf{c}_g$ ) but the check genotypes were included as fixed ( $\beta_d$ ). To be more precise, and using the same notation as in Section 2, the following mixed model was fitted

$$\mathbf{y} = f(\mathbf{u}, \mathbf{v}) + \mathbf{X}_d \beta_d + \mathbf{Z}_g \mathbf{c}_g + \mathbf{Z}_r \mathbf{c}_r + \mathbf{Z}_c \mathbf{c}_c + \mathbf{Z}_t \mathbf{c}_t + \boldsymbol{\varepsilon}.$$

where  $\mathbf{X}_d$  is the  $(n \times 4)$  design matrix assigning observations to check genotypes and  $\mathbf{Z}_g$  is the  $(n \times 1091)$  design matrix associated with candidate genotypes, with  $\mathbf{c}_g \sim N(\mathbf{0}, \sigma_g^2 \mathbf{I}_{1090})$ . The rows of  $\mathbf{X}_d$  and  $\mathbf{Z}_g$  corresponding, respectively, to candidate and check genotypes have all their elements equal to zero. For the rest of the random factors, we assumed  $\mathbf{c}_r \sim N(\mathbf{0}, \sigma_r^2 \mathbf{I}_{26})$ ,  $\mathbf{c}_c \sim N(\mathbf{0}, \sigma_c^2 \mathbf{1}_{113})$  and  $\mathbf{c}_t \sim N(\mathbf{0}, \sigma_t^2 \mathbf{I}_{31})$ . For the tensor-product P-spline, basis dimensions of 29 and 53 were chosen for the row and column positions, respectively, and nested bases, with half the dimension, were used. The model has 1789 coefficients, and 2411 observations, and the fitting process needed around 18 seconds (3.00GHz Intel Core i7 processor computer with 12GB of RAM and Windows® 7 operating system).

Figure 3 depicts the raw yield data, the fitted values, the residuals, the fitted spatial trend (i.e., the PS-ANOVA component but excluding the intercept), and the genotypic BLUPs. As can be observed, the fitted spatial trend is successful in recovering the complex spatial pattern across the field, and the residual plot suggest that the spatial independence assumption for the error vector  $\boldsymbol{\varepsilon}$  might be appropriate. Table 2 shows the model and (estimated) effective dimensions associated with each of the five smooth components of the PS-ANOVA spatial trend, each random factor and the error term. Note that the estimation procedure presented

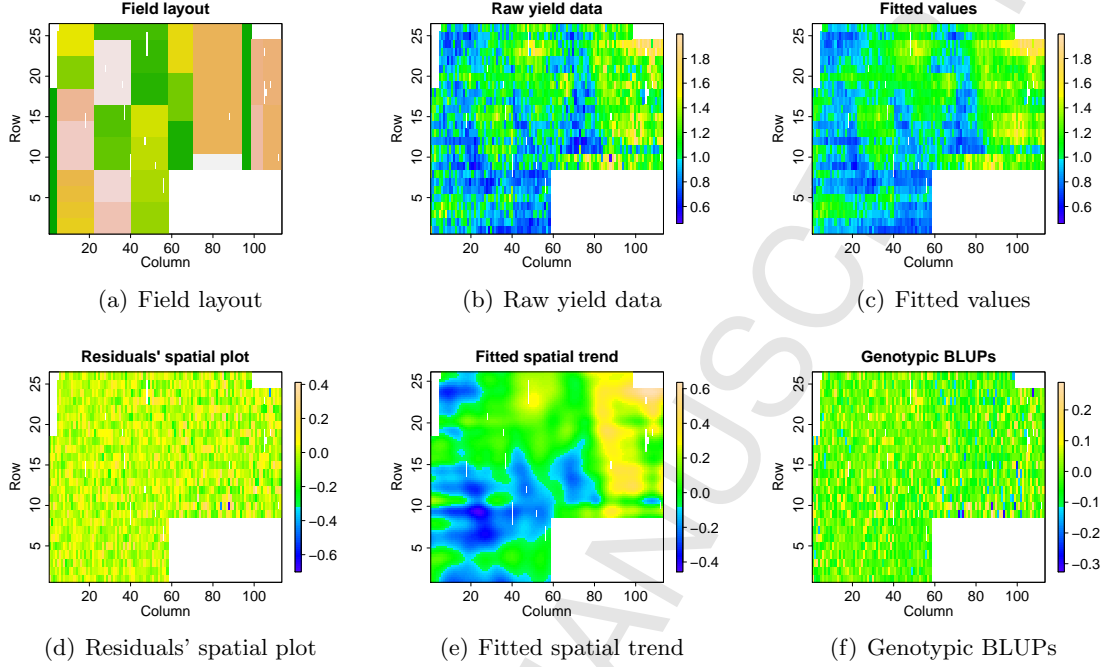


Figure 3: For the sugar beet experiment in France: field layout, raw yield data, fitted values (including all terms, fixed and random), residuals' spatial plot, fitted spatial trend and genotypic BLUPs.

in Section 4 provides, in some cases, effective dimensions that are exactly zero, meaning that this model component does not contribute or have an impact on the trait of interest. If we focus on the genetic signal, we have an effective dimension of about 590.5. In this case, it is easy to show that  $\text{rank}([\mathbf{X}, \mathbf{Z}_g]) - \text{rank}(\mathbf{X}) = 1090$ . Equivalently, there is only one eigenvalue equal to zero, which implicitly induces a zero-mean constraint on the BLUPs for  $\mathbf{c}_g$ . As a consequence, an estimate of the generalised heritability is  $\hat{H}_g = 590.5 / (1091 - 1) = 0.54$  (see eqn. (12)).

## 6.2 Wheat data, Chile

The study consisted of 384 advanced lines from two breeding programs in Chile and Uruguay. We focus here on the experiments conducted in 2011, all located in the same environment. The lines were evaluated under two different levels of water supply: mild water stress (MWS) and fully irrigated (FI) conditions. The traits of interest were grain yield (GY), thousand kernel weight (TKW), number of kernels per spike (NKS) and days to heading (DH). The experimental design was an alpha design in two replicates, each replicate containing 20 in-

Table 2: For the sugar beet experiment in France: model and effective dimensions associated with the PS-ANOVA spatial trend, each random factor, and the error term. The letter  $u$  denotes the row position,  $v$  the column position, and  $\mathbf{c}_r$ ,  $\mathbf{c}_c$ ,  $\mathbf{c}_t$  and  $\mathbf{c}_g$  the row, column, trial and genetic line random factors, respectively.

Dimensions	Model components									
	$f_u(u)$	$f_v(v)$	$vh_u(u)$	$uh_v(v)$	$f_{u,v}(u, v)$	$\mathbf{c}_r$	$\mathbf{c}_c$	$\mathbf{c}_t$	$\mathbf{c}_g$	$\varepsilon$
Model	27	51	27	51	364	26	113	31	1091	2411
Effective	0.0	12.4	21.1	0.0	120.1	18.3	13.1	22.8	590.5	1604.5

complete blocks of size 20. Incomplete blocks were placed in the row direction, as well as the replicates, placed on top of each other. The dimension of the trial was therefore 40 rows  $\times$  20 columns (800 plots). More details can be found in [Lado et al. \(2013\)](#).

In this paper we re-analyse the phenotypic data using SpATS. For each trait (GY, TKW, NKS and DH) and water condition (MWS and FI) the following model was assumed

$$\mathbf{y} = f(\mathbf{v}, \mathbf{u}) + \mathbf{Z}_g \mathbf{c}_g + \mathbf{Z}_r \mathbf{c}_r + \mathbf{Z}_c \mathbf{c}_c + \varepsilon,$$

with  $\mathbf{c}_g \sim N(\mathbf{0}, \sigma_g^2 \mathbf{I}_{384})$ ,  $\mathbf{c}_r \sim N(\mathbf{0}, \sigma_r^2 \mathbf{I}_{40})$  and  $\mathbf{c}_c \sim N(\mathbf{0}, \sigma_c^2 \mathbf{I}_{20})$ . For the tensor product P-splines, basis dimensions of 43 and 23 were assumed for the row and column positions, respectively, and, as usual, we used nested bases, with half the dimension. Each model has about 802 coefficients and 800 observations (there are missing values); the fitting processes needed between 2 and 10 seconds.

Figures 4 and 5 depict, for respectively MWS and FI water conditions, the raw data, the fitted spatial trends and the spatial plot of the residuals for each trait. Table 3 shows the effective dimensions related to all model components. Note that some  $ED_k$  are zero or close to zero. For instance, for the DH and FI condition, the spatial variation is mainly modelled by the smooth trend over the rows and the row random factor, with the rest of components having a rather low or even null impact. All results suggest that for both the MWS and FI conditions, GY is the trait presenting the largest spatial variation, and DH the one with the lowest. This is in concordance with the estimated heritability presented in Table 3, with the largest and the lowest heritability having been obtained for DH and GY, respectively. It should be noted that, in all cases, the heritability estimate obtained using SpATS is larger than the broad sense heritability reported in the paper by [Lado et al. \(2013\)](#).

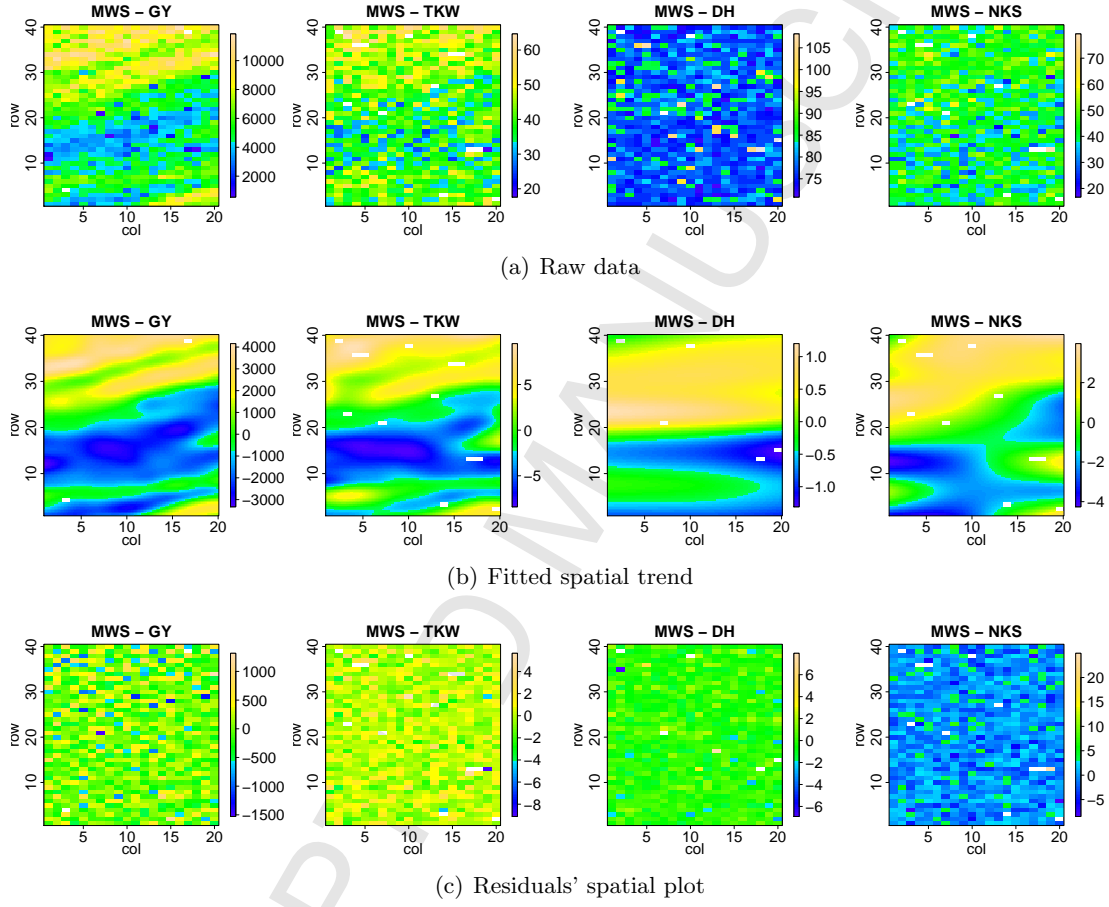


Figure 4: Raw data, fitted spatial trend and residuals' spatial plot for the Chilean wheat data (Lado et al., 2013), for each trait (GY: Grain Yield; TKW: thousand kernel weight; DH: days to heading; NKS: number of kernels per spike) and mild water stress (MWS) condition.

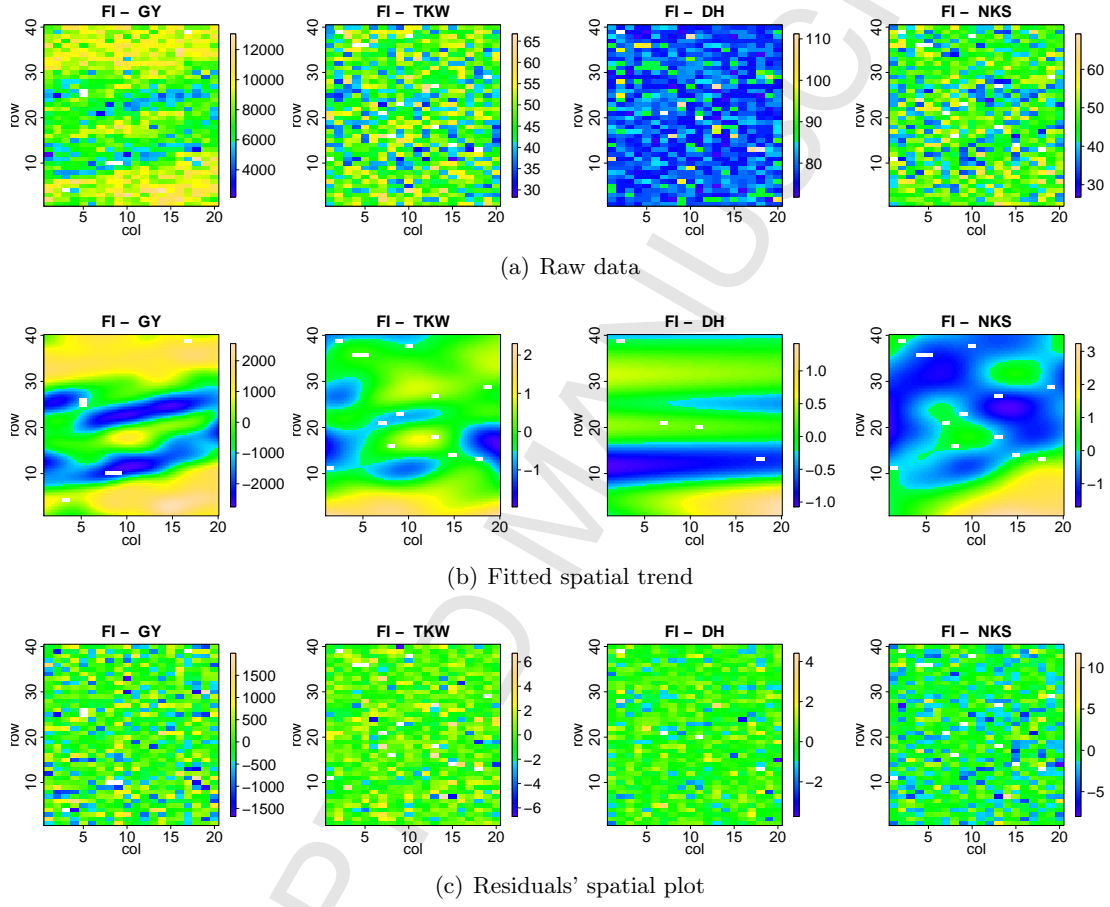


Figure 5: Raw data, fitted spatial trend and residuals' spatial plot for the Chilean wheat data (Lado et al., 2013), for each trait (GY: Grain Yield; TKW: thousand kernel weight; DH: days to heading; NKS: number of kernels per spike) and fully irrigated (FI) condition.

Table 3: For the Chilean wheat data: Effective dimensions associated with the spatial trend, the row and column random factors, the genetic random factor, and the residuals. The generalised heritability is also shown. The letter  $u$  denotes the row position,  $v$  the column position, and  $c_r$  and  $c_c$  the row and column random factors, respectively. The results shown are for Santa Rosa in 2011 and for each trait (GY: Grain Yield; TKW: thousand kernel weight; DH: days to heading; NKS: number of kernels per spike) and condition (MWS: Mild water stress; FI: Fully irrigated).

		Condition and Trait							
		MWS				FI			
		GY	TKW	DH	NKS	GY	TKW	DH	NKS
Effective dimension spatial components ( $ED_k$ )	$f_u(u)$	8.8	8.5	6.6	2.4	10.2	2.9	7.2	1.3
	$f_v(v)$	0.0	0.7	0.5	0.0	0.6	0.7	0.0	0.0
	$vh(u)$	16.8	16.7	1.1	7.6	8.4	0.4	1.4	0.5
	$uh_v(v)$	0.0	0.0	0.0	1.8	1.6	0.0	0.1	0.7
	$f_{u,v}(u, v)$	90.6	49.5	0.2	0.6	64.9	17.9	0.0	12.2
	$c_r$	20.2	5.2	3.7	12.5	13.7	17.8	7.7	23.7
	$c_c$	6.9	0.4	5.6	3.6	6.3	0.0	0.0	0.3
	Total	143.3	81.0	17.7	28.5	105.7	39.7	16.4	38.8
$ED_g$		281.8	357.4	368.5	318.3	264.5	354.4	373.4	326.6
$H_g^2$		0.74	0.94	0.96	0.84	0.69	0.94	0.98	0.86
$ED_\epsilon$		368.9	344.6	404.8	438.2	419.8	388.9	402.2	417.6
Number of observations		798	787	795	789	794	787	796	787



## 7 Conclusion

This paper presents a new approach to the modelling of field trials in plant breeding, called SpATS. It has many useful properties.

- The spatial effect is modelled explicitly as a surface, using anisotropic tensor-product P-splines. Graphical display of this surface is of considerable value when interpreting a trial.
- The surface is decomposed into a number of fixed and random components, to allow the use of mixed model technology to automatically determine the amount of smoothing.
- The (fixed or random) effects of genotypes, rows and columns, blocks and other known systematic pattern can be included as usual.
- The diagnostic value of the effective dimension of model components is demonstrated and connected to classical measures of heritability.
- The model has been applied to many hundreds of field trials, showing fast and robust operation without any need for user interaction.
- Thanks to the use of nested B-splines, very large fields can be analysed in moderate amounts of time.
- Missing observations, even many of them, and possibly in large patches, are handled gracefully and automatically.
- The observations do not have to be located on a rectangular grid.
- An R package is available. It performs the computations and presents informative and attractive tables and graphs.

SpATS is an attractive alternative to classical analyses of field trials that model spatial variation as correlated noise. In those classical analysis, human intervention is required and a painful model selection process, by a skilled user, is needed. Our simulations document the instability of  $AR \times AR$  models.  $LV \times LV$  worked much better, showing a similar performance to SpATS. However, especially for larger fields,  $LV \times LV$  is much slower and it does not provide an explicit estimate of the spatial surface.

## 8 Discussion

The modelling strategy followed in this paper represents a continuation of previous work by [Green et al. \(1985\)](#) and [Besag and Higdon \(1999\)](#), among others. In particular, the connection between SpATS and the first-differencing model in rows and columns proposed by [Besag and Higdon \(1999\)](#) becomes evident under the Bayesian formulation of P-splines ([Lang and Brezger, 2004](#); [Kneib, 2006](#)). In contrast to the proposal by [Besag and Higdon \(1999\)](#), SpATS relies on B-splines which considerably reduces the computational cost, and allows the analysis of very large fields and of non-lattice data. In addition, SpATS, as presented in the paper, is based on second order difference penalties for the tensor product P-splines. These second order differences are responsible for the presence in the PS-ANOVA spatial field of the bilinear polynomial (except the intercept) and the linear-by-smooth interaction trends. However, SpATS permits the use of other penalty orders, as e.g., first order differences. Comparisons among different penalty orders is a subject for further research.

Looking to the future, we see many interesting challenges. An obvious generalisation is the modelling of non-Gaussian data like counts or fractions. SpATS is based on penalised linear regression, framed in the mixed model framework. In principle, the theory of generalised linear mixed models applies, with a linear predictor composed from fixed and random effects. The iterative re-weighted REML algorithm essentially fits a linear mixed model to a “working variable” (see, e.g., [Schall, 1991](#); [Engel and Keen, 1994](#)). This lends itself well to the estimation of effective dimensions and variances. The theory will be presented elsewhere, but we mention that the R-package SpATS that accompanies this paper already allows a non-Gaussian response.

Partial effective dimensions are key elements of our model. They are needed to estimate variances, but they also summarise the importance of each component in a single number with strong intuitive appeal. Although in the smoothing literature there is growing attention for effective dimensions, this is not yet the case for mixed models. The paper by [Cui et al. \(2010\)](#) represents an important contribution. In that paper, the authors propose a new conception of components’ effective dimension for mixed effects models that can be loosely interpreted as the fraction of response variation attributed to individual components. As shown in Web Appendix D, it turns out that the definition of partial effective dimensions used in this paper (borrowed from the smoothing context) and that of [Cui et al. \(2010\)](#) are equivalent. We also show the strong connection between the generalised heritability proposed by [Cullis et al. \(2006\)](#) and [Oakey et al. \(2006\)](#) and the genetic effective dimension. All these results support moving the interpretation of the results obtained by fitting a mixed model from estimated variance components to estimated effective dimensions. Partial effective dimensions furnish a

common scale that allows comparing the contribution of each random/smooth model component (and the error term) to the observed phenotypic variation. This subject deserves further research. An immediate question is how things work out with non-Gaussian responses.

A current line of research is focused on the extension of SpATS to the analysis of multi-environment trials (METs, trials carried out in multiple environments or contexts, see, e.g., [Smith et al., 2009](#)). In the analysis of METs we are faced with several challenges. First, the modelling of a different spatial trend for each environment, and second the need to account for the (possible) interaction between the genotype and the environment. The first challenge may be approached by including a separate P-spline surface for each trial. The second challenge is related to the inclusion of correlated random effects (whenever the genotype effect is treated as random).

Traditionally, in field trial experiments the traits of interest are only measured at the end of the study (or a rather low number of times). However, with the arrival of high-throughput phenotyping platforms, long time-series measurements of traits are nowadays available. This requires a spatio-temporal perspective ([Fahlgren et al., 2015](#); [Cobb et al., 2013](#)) for the analysis of field trials. An exciting field for further study is thus the generalisation of the methodology presented in this paper to the analysis of high-throughput data. Further work is also needed to develop computationally efficient estimation procedures in the aforementioned settings.

Our calculations were done with the R-package SpATS that can be downloaded from <https://cran.r-project.org/package=SpATS>. It is written in pure R, only using the package `fields` for attractive display of results. Thus one can easily convert the software, if desired, to other languages like Python, Matlab<sup>®</sup> or Julia. Also mixed model packages like `asreml-R`, `nlme` and `lme4` can be used, or the PROC MIXED procedure in SAS<sup>®</sup>.

The R-code used for the simulation studies presented in Section 5 and Web Appendix E, and the analyses of the Chilean wheat trial discussed in Section 6 can be downloaded from <http://bitbucket.org/mxrodriguez/spats-paper>.

## Acknowledgements

This research was supported by the Spanish Ministry of Economy and Competitiveness MINECO grant MTM2014-55966-P and BCAM Severo Ochoa excellence accreditation SEV-2013-0323, and by the Basque Government through the BERC 360 2014-2017. We thank SESVanderHave for providing the sugar beet data. We are grateful to Cajo ter Braak, María Durbán, Dae-Jin Lee and Julio Velazco for useful discussions. We are also grateful to the editor and the two peer referees for their constructive comments of the paper.

## Supplementary Materials

Web Appendices A – F are available with this paper.

## Appendix A P-splines and mixed models

The two-dimensional P-spline model (4) can be estimated by minimising the penalised least squares objective function (Eilers and Marx, 2003)

$$S^* = \|\mathbf{y} - \mathbf{B}\boldsymbol{\alpha}\|^2 + \boldsymbol{\alpha}^t \mathbf{P} \boldsymbol{\alpha}, \quad (14)$$

where  $\mathbf{P} = \hat{\lambda}(\mathbf{I}_{\hat{m}} \otimes \hat{\mathbf{D}}^t \hat{\mathbf{D}}) + \check{\lambda}(\check{\mathbf{D}}^t \check{\mathbf{D}} \otimes \mathbf{I}_{\check{m}})$  is the penalty matrix,  $\hat{\lambda}$  and  $\check{\lambda}$  are the smoothing parameters acting, respectively, along the  $u$  (rows) and  $v$  (columns) directions, and  $\hat{\mathbf{D}}$  and  $\check{\mathbf{D}}$  are matrices that form differences of order  $d_u$  and  $d_v$  respectively (in this paper we use second order differences, i.e.,  $d_u = d_v = 2$ ).

The minimizer of (14), given  $\hat{\lambda}$  and  $\check{\lambda}$ , is  $\hat{\boldsymbol{\alpha}} = (\mathbf{B}^t \mathbf{B} + \mathbf{P})^{-1} \mathbf{B}^t \mathbf{y}$ . The only tuning mechanism for smoothness is the strength of the penalty  $\mathbf{P}$ , i.e the value of the smoothing parameters  $\check{\lambda}$  and  $\hat{\lambda}$ . The number of B-splines will be purposely chosen so large as to get over-fitting without a penalty. Accordingly, a critical issue is setting the right value for the smoothing parameters, which we like to see determined by the data. In this paper, we rely on the connection between P-splines and mixed models (Currie and Durban, 2002; Wand, 2003). In this approach, the smooth functions are treated as sums of fixed and random components, and the smoothing parameters are replaced by ratios of variances which are estimated by REML. We present here the main ideas, and refer to Lee et al. (2013) (and references therein) for a more detailed description.

Let  $\hat{\mathbf{D}}^t \hat{\mathbf{D}} = \mathbf{U}_u \mathbf{E}_u \mathbf{U}_u^t$  and  $\check{\mathbf{D}}^t \check{\mathbf{D}} = \mathbf{U}_v \mathbf{E}_v \mathbf{U}_v^t$  be the eigenvalue decomposition of the marginal penalties  $\hat{\mathbf{D}}^t \hat{\mathbf{D}}$  and  $\check{\mathbf{D}}^t \check{\mathbf{D}}$  respectively. Here  $\mathbf{U}_j$  denotes the matrix of eigenvectors and  $\mathbf{E}_j$  the diagonal matrix of eigenvalues ( $j = u, v$ ). For second order differences,  $\mathbf{E}_u$  and  $\mathbf{E}_v$  contain  $\hat{m} - 2$  and  $\check{m} - 2$  non-zero values respectively. Let us also denote by  $\tilde{\mathbf{U}}_j$  and  $\tilde{\mathbf{E}}_j$  the sub-matrices corresponding to the non-zero eigenvalues. Setting

$$\mathbf{X}_s = [\mathbf{X}_v \square \mathbf{X}_u] \equiv [\mathbf{1}_n \mid \mathbf{u} \mid \mathbf{v} \mid \mathbf{u} \odot \mathbf{v}], \quad (15)$$

$$\mathbf{Z}_s = [\mathbf{Z}_v \square \mathbf{X}_u \mid \mathbf{X}_v \square \mathbf{Z}_u] \equiv [\mathbf{Z}_v \mid \mathbf{Z}_u \mid \mathbf{Z}_v \square \mathbf{u} \mid \mathbf{v} \square \mathbf{Z}_u \mid \mathbf{Z}_v \square \mathbf{Z}_u]. \quad (16)$$

where  $\mathbf{X}_u = [\mathbf{1}_n \mid \mathbf{u}]$ ,  $\mathbf{X}_v = [\mathbf{1}_n \mid \mathbf{v}]$ ,  $\mathbf{Z}_u = \check{\mathbf{B}} \tilde{\mathbf{U}}_u$  and  $\mathbf{Z}_v = \hat{\mathbf{B}} \tilde{\mathbf{U}}_v$ , the penalised least squares problem (14) becomes

$$S^* = \|\mathbf{y} - \mathbf{X}_s \boldsymbol{\beta}_s - \mathbf{Z}_s \mathbf{c}_s\|^2 + \mathbf{c}_s^t \tilde{\mathbf{P}} \mathbf{c}_s, \quad (17)$$

where  $\tilde{\mathbf{P}} = \text{blockdiag}(\check{\lambda}\tilde{\mathbf{E}}_v, \hat{\lambda}\tilde{\mathbf{E}}_u, \check{\lambda}\tilde{\mathbf{E}}_v, \hat{\lambda}\tilde{\mathbf{E}}_u, \check{\lambda}\tilde{\mathbf{E}}_v \otimes \mathbf{I}_{\hat{m}-2} + \hat{\lambda}\mathbf{I}_{\check{m}-2} \otimes \tilde{\mathbf{E}}_u)$  is the new penalty matrix. Each block in  $\tilde{\mathbf{P}}$  corresponds to a block in  $\mathbf{Z}_s$ . This reformulation provides the ANOVA-type decomposition discussed in Section 2. The block-structure of both  $\mathbf{X}_s$  and  $\mathbf{Z}_s$  implies

$$\begin{aligned} f(\mathbf{u}, \mathbf{v}) &= \mathbf{X}_s \boldsymbol{\beta}_s + \mathbf{Z}_s \mathbf{c}_s \\ &= \underbrace{\mathbf{1}_n \beta_0 + \mathbf{u} \beta_1 + \mathbf{v} \beta_2 + \mathbf{u} \odot \mathbf{v} \beta_3}_{\mathbf{X}_s \boldsymbol{\beta}_s} + \underbrace{f_v(\mathbf{v})}_{\mathbf{Z}_v \mathbf{c}_{s1}} + \underbrace{f_u(\mathbf{u})}_{\mathbf{Z}_u \mathbf{c}_{s2}} + \underbrace{\mathbf{u} \odot h_v(\mathbf{v})}_{[\mathbf{Z}_v \square \mathbf{u}] \mathbf{c}_{s3}} + \underbrace{\mathbf{v} \odot h_u(\mathbf{u})}_{[\mathbf{v} \square \mathbf{Z}_u] \mathbf{c}_{s4}} + \underbrace{f_{u,v}(\mathbf{u}, \mathbf{v})}_{[\mathbf{Z}_v \square \mathbf{Z}_u] \mathbf{c}_{s5}}, \end{aligned}$$

where  $\mathbf{c}_{sk}$  ( $k = 1, \dots, 5$ ) contains the elements of  $\mathbf{c}_s$  that correspond to the  $k$ -th block of  $\mathbf{Z}_s$ , i.e.,  $\mathbf{c}_s = (\mathbf{c}_{s1}^t, \dots, \mathbf{c}_{s5}^t)^t$ .

The solution to (17), for given  $\check{\lambda}$  and  $\hat{\lambda}$ , corresponds to the empirical best linear unbiased estimator (BLUE) for the  $(4 \times 1)$  vector  $\boldsymbol{\beta}_s$ , and the BLUPs for the  $((\hat{m}\check{m} - 4) \times 1)$  vector  $\mathbf{c}_s$  under the linear mixed model

$$\mathbf{y} = \mathbf{X}_s \boldsymbol{\beta}_s + \mathbf{Z}_s \mathbf{c}_s + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n) \text{ and } \mathbf{c}_s \sim N(\mathbf{0}, \mathbf{G}_s),$$

where  $\mathbf{G}_s = \sigma^2 \tilde{\mathbf{P}}^{-1}$ . Note that, despite the five smooth components,  $\mathbf{G}_s$  involves only two variance parameters  $\check{\sigma}^2 = \sigma^2/\check{\lambda}$  and  $\hat{\sigma}^2 = \sigma^2/\hat{\lambda}$ . In fact, the same variance parameters control the smoothness of both main effects and interaction terms. This prevents the use of standard mixed model software for estimation, since  $\mathbf{G}_s$  has its last block depending on  $\check{\sigma}^2$  and  $\hat{\sigma}^2$ , but in a non-linear way. Specialised algorithms and software have been presented (e.g., Rodríguez-Álvarez et al., 2015), but this paper adopts the so-called PS-ANOVA model (Lee et al., 2013), allowing our proposal to be implemented using standard mixed model estimation procedures. Specifically, in Lee et al. (2013) the ANOVA-type decomposition is further exploited, and the authors propose to use a different variance component for each smooth component. For ease of notation, let  $\boldsymbol{\Lambda}_{s1}^{-1} = \boldsymbol{\Lambda}_{s3}^{-1} = \tilde{\mathbf{E}}_v$ ,  $\boldsymbol{\Lambda}_{s2}^{-1} = \boldsymbol{\Lambda}_{s4}^{-1} = \tilde{\mathbf{E}}_u$ , and  $\boldsymbol{\Lambda}_{s5}^{-1} = \tilde{\mathbf{E}}_v \otimes \mathbf{I}_{\hat{m}-2} + \mathbf{I}_{\check{m}-2} \otimes \tilde{\mathbf{E}}_u$ . For the PS-ANOVA model the precision matrix is then defined as  $\mathbf{G}_s^{-1} = \text{blockdiag}\left(\frac{1}{\sigma_{s1}^2} \boldsymbol{\Lambda}_{s1}^{-1}, \frac{1}{\sigma_{s2}^2} \boldsymbol{\Lambda}_{s2}^{-1}, \frac{1}{\sigma_{s3}^2} \boldsymbol{\Lambda}_{s3}^{-1}, \frac{1}{\sigma_{s4}^2} \boldsymbol{\Lambda}_{s4}^{-1}, \frac{1}{\sigma_{s5}^2} \boldsymbol{\Lambda}_{s5}^{-1}\right)$ , and thus the covariance matrix  $\mathbf{G}_s$  is a linear function of variance parameters

$$\mathbf{G}_s = \bigoplus_{k=1}^5 \sigma_{sk}^2 \boldsymbol{\Lambda}_{sk} = \bigoplus_{k=1}^5 \mathbf{G}_{sk} = \text{blockdiag}(\mathbf{G}_{s1}, \mathbf{G}_{s2}, \mathbf{G}_{s3}, \mathbf{G}_{s4}, \mathbf{G}_{s5}), \quad (18)$$

where  $\mathbf{G}_{sk} = \sigma_{sk}^2 \boldsymbol{\Lambda}_{sk}$  ( $k = 1, \dots, 5$ ). In other words, here the tensor product P-spline mixed model is represented as the sum of 5 sets of mutually independent Gaussian random components  $\mathbf{c}_{sk}$ , each depending on one variance  $\sigma_{sk}^2$  ( $k = 1, \dots, 5$ ).

## References

- Besag, J. and D. Higdon (1999). Bayesian analysis of agricultural field experiments. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 61, 691–746.
- Cappa, E. P. and R. J. C. Cantet (2008). Bayesian estimation of a surface to account for a spatial trend using penalized splines in an individual-tree mixed model. *Canadian Journal of Forest Research* 37, 2677–2688.
- Cobb, J. N., G. DeClerck, A. Greenberg, R. Clark, and S. McCouch (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics* 126(4), 867–887.
- Cui, Y., J. S. Hodges, X. Kong, and B. P. Carlin (2010). Partitioning degrees of freedom in hierarchical and other richly-parameterized models. *Technometrics* 52, 124–136.
- Cullis, B., A. Smith, and N. Coombes (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics* 11(4), 381–393.
- Cullis, B. R. and A. C. Gleeson (1991). Spatial analysis of field experiments - an extension to two dimensions. *Biometrics* 47, 1449–1460.
- Currie, I. D. and M. Durban (2002). Flexible smoothing with P-splines: a unified approach. *Statistical Modelling* 2(4), 333–349.
- Dierckx, P. (1993). *Curve and Surface Fitting with Splines*. Oxford: Clarendon Press.
- Durban, M., C. A. Hackett, J. W. McNicol, A. C. Newton, W. T. B. Thomas, and I. D. Currie (2003). The practical use of semiparametric models in field trials. *Journal of Agricultural, Biological, and Environmental Statistics* 8(1), 48–66.
- Eilers, P. H. C., I. D. Currie, and M. Durban (2006). Fast and compact smoothing on large multidimensional grids. *Computational Statistics and Data Analysis* 50, 61–76.
- Eilers, P. H. C. and B. D. Marx (2003). Multidimensional calibration with temperature interaction using two-dimensional penalized signal regression. *Chemometrics and Intelligence Laboratory Systems* 66, 159–174.
- Eilers, P. H. C., B. D. Marx, and M. Durban (2015). Twenty years of P-splines. *SORT* 39(2), 149–186.

- Engel, B. and A. Keen (1994). A simple approach for the analysis of generalized linear mixed models. *Statistica Neerlandica* 48(1), 1–22.
- Fahlgren, N., M. A. Gehan, and I. Baxter (2015). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology* 24, 93 – 99.
- Gilmour, A. R., B. R. Cullis, and A. P. Verbyla (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological and Environmental Statistics* 2, 269–293.
- Green, P., C. Jennison, and A. Seheult (1985). Analysis of field experiments by least squares smoothing. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 47, 299–315.
- Harville, D. A. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *Journal of the American Statistical Association* 72(358), 320–338.
- Hastie, T. J. and R. J. Tibshirani (1990). *Generalized additive models*. London: Chapman & Hall.
- Kneib, T. (2006). *Mixed model based inference in structured additive regression*. Ludwig-Maximilians-Universität München.
- Lado, B., I. Matus, A. Rodríguez, L. Inostroza, J. Poland, F. Belzile, A. del Pozo, M. Quincke, M. Castro, and J. von Zitzewitz (2013). Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3 (Genes, Genomes, Genetics)* 3(12), 2105–2114.
- Lang, S. and A. Brezger (2004). Bayesian p-splines. *Journal of Computational and Graphical Statistics* 13(1), 183–212.
- Lee, D.-J., M. Durban, and P. H. C. Eilers (2013). Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested basis. *Computational Statistics and Data Analysis* 61, 22–37.
- Oakey, H., A. Verbyla, W. Pitchford, B. Cullis, and H. Kuchel (2006). Joint modeling of additive and non-additive genetic line effects in single field trials. *Theoretical and Applied Genetics* 113(5), 809–819.
- Patterson, H. D. and R. Thompson (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58(3), 545–554.

- Piepho, H. P., J. Möhring, M. Pflugfelder, W. Hermann, and E. R. Williams (2015). Problems in parameter estimation for power and AR(1) models of spatial correlation in designed field experiments. *Communications in Biometry and Crop Science* 10(1), 3–16.
- Piepho, H. P. and E. R. Williams (2010). Linear variance models for plant breeding trials. *Plant Breeding* 129, 1–8.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Robbins, K. R., J. E. Backlund, and K. D. Schnelle (2012). Spatial corrections of unreplicated trials using a two-dimensional spline. *Crop Science* 52, 1138–1144.
- Rodríguez-Álvarez, M. X., D.-J. Lee, T. Kneib, M. Durban, and P. H. C. Eilers (2015). Fast smoothing parameter separation in multidimensional generalized P-splines: the sap algorithm. *Statistics and Computing* 25, 941–957.
- Ruppert, D., M. P. Wand, and R. Carroll (2003). *Semiparametric Regression*. Cambridge University Press.
- Schall, R. (1991). Estimation in generalized linear models with random effects. *Biometrika* 78(4), 719–727.
- Smith, A., B. Cullis, and R. Thompson (2009). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57, 1138–1147.
- Taye, G. and P. M. Njuho (2008). Smoothing fertility trends in agricultural field experiments. *Statistics* 42, 275–289.
- Verbyla, A. P., B. R. Cullis, M. G. Kenward, and S. J. Welham (1999). The analysis of designed experiments and longitudinal data by smoothing splines (with discussion). *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 48, 269–312.
- Wand, M. P. (2003). Smoothing and mixed models. *Computational Statistics* 18(2), 223–249.
- Williams, E. R. (1986). A neighbour model for field experiments. *Biometrika* 73(2), 279–287.
- Williams, E. R., J. A. John, and D. Whitaker (2006). Construction of resolvable spatial row-column designs. *Biometrics* 62(1), 103–108.



- Williams, E. R. and D. J. Lockett (1988). The use of uniformity data in the design and analysis of cotton and barley variety trials. *Australian Journal of Agricultural Research* 39, 545–554.
- You, C., S. Müller, and J. T. Ormerod (2016). On generalized degrees of freedom with application in linear mixed models selection. *Statistics and Computing* 26, 199–210.
- Zimmerman, D. L. and D. A. Harville (1991). A random field approach to the analysis of field-plot experiments and other spatial experiments. *Biometrics* 47, 223–239.