


Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model

Julio G. Velazco^{1,2}  · María Xosé Rodríguez-Álvarez^{3,4} · Martin P. Boer¹ · David R. Jordan⁵ · Paul H. C. Eilers⁶ · Marcos Malosetti¹ · Fred A. van Eeuwijk¹

Received: 2 November 2016 / Accepted: 18 March 2017 / Published online: 3 April 2017
© The Author(s) 2017. This article is an open access publication

Abstract

Key message A flexible and user-friendly spatial method called SpATS performed comparably to more elaborate and trial-specific spatial models in a series of sorghum breeding trials.

Abstract Adjustment for spatial trends in plant breeding field trials is essential for efficient evaluation and selection of genotypes. Current mixed model methods of spatial analysis are based on a multi-step modelling process where global and local trends are fitted after trying several candidate spatial models. This paper reports the application of a novel spatial method that accounts for all types of continuous field variation in a single modelling step by fitting a smooth surface. The method uses two-dimensional P-splines with anisotropic smoothing formulated in the mixed model framework, referred to as SpATS model.

We applied this methodology to a series of large and partially replicated sorghum breeding trials. The new model was assessed in comparison with the more elaborate standard spatial models that use autoregressive correlation of residuals. The improvements in precision and the predictions of genotypic values produced by the SpATS model were equivalent to those obtained using the best fitting standard spatial models for each trial. One advantage of the approach with SpATS is that all patterns of spatial trend and genetic effects were modelled simultaneously by fitting a single model. Furthermore, we used a flexible model to adequately adjust for field trends. This strategy reduces potential parameter identification problems and simplifies the model selection process. Therefore, the new method should be considered as an efficient and easy-to-use alternative for routine analyses of plant breeding trials.

Communicated by Christine A Hackett.

✉ Fred A. van Eeuwijk
fred.vaneeuwijk@wur.nl

- ¹ Biometris, Wageningen University and Research, P.O. Box 16, 6700 AA Wageningen, The Netherlands
- ² Department of Plant Breeding, National Institute of Agricultural Technology (INTA), B2700WAA EEA Pergamino, Buenos Aires, Argentina
- ³ BCAM, Basque Center for Applied Mathematics, Bilbao, Spain
- ⁴ IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
- ⁵ Queensland Alliance for Agriculture and Food Innovation, Hermitage Research Facility, The University of Queensland, Warwick, QLD 4370, Australia
- ⁶ Erasmus University Medical Centre, Rotterdam, The Netherlands

Introduction

Efficient phenotypic and genomic selection schemes in plant breeding programs rely on accurate assessment of the phenotypic performance of genotypes in field experiments (Qiao et al. 2004; Lado et al. 2013; Bernal-Vasquez et al. 2014; Sarker and Singh 2015). Plant breeding trials usually involve a large number of test entries covering large areas where spatial variation is likely to be an obstacle to reliable prediction of genetic values. This is particularly challenging in early generation variety trials conditioned by the use of limited replication of genetic material.

A number of sophisticated experimental designs, such as those enabling the recovery of inter-block information (Yates 1940; Patterson et al. 1978; John and Williams 1995) or partially replicated designs (Cullis et al. 2006; Williams et al. 2014), have been developed to correct for

part of the field trend. However, efficient approaches to account for more complex environmental variation require complementing experimental designs with appropriate models of analysis (Basford et al. 1996; Qiao et al. 2000; Smith et al. 2002). Several spatial methods have been suggested to improve the precision of phenotyping. The most commonly used spatial models consider the correlation between residuals from neighboring plots to adjust for local trend or small-scale variation. These spatial methods include nearest neighbor analyses (Bartlett 1978; Wilkinson et al. 1983), and mixed model analyses using the first-order autoregressive (AR1) functions (Cullis and Gleeson 1991) or other spatial covariance structures (e.g. Zimmerman and Harville 1991; Piepho and Williams 2010). Polynomials have been also used on top of experimental design features to account for additive and non-additive trends along row and column directions (Edmondson 1993; Federer 1998). Fertility trends in early generation variety trials have been modelled by fitting one-dimensional cubic smoothing splines within blocks (Durbán et al. 2001). Durbán et al. (2003) applied semiparametric models for spatial analysis of field experiments and presented graphical and analytical model selection criteria.

Within the mixed model framework, Gilmour et al. (1997) proposed an elaborate procedure for spatial analysis of agricultural variety trials. Their approach starts by fitting a two-dimensional separable AR1 model by default to account for local trend. Eventually, extraneous variation resulting from trial management practices may be accommodated with additional model terms, while global trends reflecting large-scale variation across the field are modelled by one-dimensional polynomials or splines in the direction of rows and/or columns. The authors suggested a sequential model-fitting scheme to identify the most suitable spatial model. The procedure relies on graphical diagnostic tools and requires several modelling choices to be tried. Stefanova et al. (2009) extended this modelling process by including more formal diagnostics to facilitate model selection. However, the above-mentioned approach is not without limitations. First and foremost, the proposed multi-step procedures may not be attractive for routine analysis of large series of trials, since it requires a high level of hands-on intervention. Furthermore, there exists a risk of over-fitting the spatial data when the number of candidate models involved in the model selection process increases. Finally, convergence failures due to parameter identification problems may occur when trying to fit different spatial terms simultaneously (Dutkowski et al. 2006; Müller et al. 2010; Piepho et al. 2015).

Multidimensional regression spline methods represent a flexible alternative to account for complex variation

structures. They allow the modelling of smooth multidimensional (or interaction) surfaces (e.g., Ruppert et al. 2003; Currie et al. 2006; Wood 2006). Regression splines are efficient curve-fitting functions composed of polynomial pieces, generally quadratic or cubic, that are joined at points called “knots”. An interesting method using splines is based on two-dimensional P-splines (2D P-splines) as proposed by Eilers and Marx (1996, 2003), and its formulation in the linear mixed model framework (Eilers 1999; Currie and Durbán 2002). P-splines combine regression splines and a roughness penalty, which is the key component. This penalization is tuned by one or more smoothing parameters that control the degree of smoothness of the fitted spatial surface to prevent over-fitting. The connection between P-splines and mixed models provides attractive advantages. It enables the use of efficient algorithms for inference and prediction. Furthermore, the optimal smoothing parameters are automatically estimated by restricted maximum likelihood (REML; Patterson and Thompson 1971) as ratios of variance components.

Some applications of 2D P-spline models have been reported for spatial analysis of field trials. Cappa and Cantet (2007) and Cappa et al. (2011, 2015) used these models within a Bayesian approach to account for global trends in forest genetic trials. These studies considered a single smoothing parameter that controls the smoothness of the spatial effects in the direction of both rows and columns, imposing isotropic smoothing. In agricultural experiments, Taye and Njuho (2008) proposed using P-splines in two dimensions to adjust for global trend and to model local variation with Papadakis and kriged covariates. The authors compared P-spline models assuming additive trends or interaction between trends and emphasized the importance of choosing between both model settings. A different approach to spatial analysis of field trials using 2D P-spline mixed models was recently proposed by Rodríguez-Álvarez et al. (2016a). They introduced a novel spatial model that adjusts for both global and local trends simultaneously. The authors called this model SpATS, an acronym for Spatial Analysis of field Trials with Splines. The new spatial method makes use of the P-spline ANOVA representation of the smooth surface according to Lee et al. (2013). The distinctive feature of the SpATS model is an attractive decomposition of the spatial surface into additive one-dimensional trends and two-dimensional interaction trends. Furthermore, the model assigns a different smoothing parameter to each spatial component, allowing for anisotropic smoothing. This parametrization enables a flexible modelling of the spatial surface, where each component has a straightforward interpretation.

For the present research, we considered a series of multi-environmental trials from a sorghum [*Sorghum bicolor* (L.) Moench] breeding program in eastern Australia. These trials belong to the initial stages of evaluations, where a large number of breeding lines (approximately 1000) were tested in each experiment using partially replicated designs. Furthermore, studies regarding the implications of performing spatial analysis in sorghum genetic trials are limited in the literature. Consequently, this data set serves to illustrate a situation when a flexible and efficient spatial analysis tool is specially required.

This paper reports an application of the SpATS mixed model to adjust for all types of field trend in early generation sorghum breeding trials. We use a one-step modelling approach to spatial analysis by fitting a general SpATS model to analyze the whole series of trials. This approach is assessed in comparison with more elaborate and trial-specific spatial models identified according to the method of Gilmour et al. (1997). Both methods are compared in terms of variance component estimates, the improvement of precision, and correlation of predicted genotypic effects. The new spatial model has been fitted using a tailor-made R package (R Development Core Team 2016) called SpATS (Rodríguez-Álvarez et al. 2016b), which is publicly available from CRAN (<https://cran.r-project.org/package=SpATS>).

Materials and methods

Data set

In this study, we used data from 21 sorghum breeding trials conducted at 12 different locations in eastern Australia between 2005 and 2008. The data set is part of the public germplasm enhancement program managed by the University of Queensland and Queensland's Department of Agriculture and Fisheries. A total of 3947 backcross recombinant inbred lines (BC-RILs) were evaluated as male parents in test-cross hybrid combinations with a single female tester. The BC-RILs were derived from crosses between an elite inbred line and a range of exotic sorghum lines. Detailed descriptions of the breeding population used in this paper can be found in Jordan et al. (2011) and Mace et al. (2013). The set of trials is considered to represent the target population of environments in the Australian sorghum cropping region.

Each trial was laid out as a rectangular array using resolvable p -rep designs (Cullis et al. 2006). Table 1 summarizes information related to the individual trials, including the field layout and the number of genotypes per location. Plots were 5 m wide along rows by 1.5 or 2 m long down the columns, with two rows of plants in each plot. The p -rep designs consisted of 30% of the test-cross hybrids having two replicates ($p=30\%$), while the remaining 70% of the genotypes were unreplicated. Across all trials, a total of ten commercial

Table 1 Description of experimental layout and mean values of grain yield (GY) and plant height (PH) for each trial in the sorghum breeding data set

Trial	Year	Location	Rows	Columns	Plots	Genotypes	Mean GY (t/ha)	Mean PH (cm)
BIL05	2005	Biloela	55	28	1540	1136	4.10	104
DAB05	2005	Dalby Box	76	20	1520	1167	3.05	91
DYS05	2005	Dysart	77	20	1540	1079	1.31	95
HER05	2005	Hermitage	77	20	1540	1202	7.65	113
JIM05	2005	Jimbour	44	20	880	682	4.60	105
BIL06	2006	Biloela	81	20	1620	1060	6.70	115
CEP06	2006	Cecil Plains	48	30	1440	953	3.24	95
DAB06	2006	Dalby	62	20	1240	823	2.00	101
GON06	2006	Goondiwindi	72	20	1440	957	6.58	117
HER06	2006	Hermitage	74	20	1480	1075	8.75	109
BIL07	2007	Biloela	86	20	1720	998	2.78	104
CLE07	2007	Clermont	34	40	1360	768	3.03	112
DYS07	2007	Dysart	44	40	1760	938	2.82	113
HER07	2007	Hermitage	70	25	1750	1012	5.39	104
BIL08	2008	Biloela	80	20	1600	1010	4.33	123
DAB08	2008	Dalby Box	64	20	1280	947	6.67	133
DAL08	2008	Dalby	62	20	1240	903	6.48	132
HER08	2008	Hermitage	80	20	1600	1012	–	133
KIL08	2008	Kilcummin	75	20	1500	899	3.36	131
LIV08	2008	Liverpool Plains	66	20	1320	980	10.06	125
SPR08	2008	Springsure	46	20	920	753	3.87	–

varieties were included as check entries with additional levels of replication. Allocation of the replicated test genotypes was based on an optimality measure determined by the average pairwise prediction error variance and assuming a pre-specified spatial model (Cullis et al. 2006). The search algorithm is constrained, so that the replicated hybrids occurred once in each half of the trial, which established two resolvable blocks in all the designs.

We illustrate the spatial analyses with two traits: grain yield (t/ha) and plant height (cm). Data were not available for grain yield at trial HER08 and for plant height at trial SPR08 (Table 1). The proportion of missing plots ranged between 3 and 29%.

The SpATS model

In this section, we present a brief description of the SpATS model; for a thorough treatment of the model specifications, we refer to the original study by Rodríguez-Álvarez et al. (2016a).

Consider that observations in each sorghum breeding trial were obtained from plots arranged as a rectangular grid, where plot positions are collected in vectors of row (\mathbf{r}) and column (\mathbf{c}) coordinates. Under the SpATS model, field trends are modelled by a smooth bivariate function of the spatial coordinates $f(\mathbf{r}, \mathbf{c})$ represented by 2D P-splines. As said, this technique optimizes the fitted surface by penalizing or shrinking the spatial effects. The magnitude of the penalization over the fitted trend is determined by the smoothing parameters. These terms control the balance between smoothness of the fitted surface and fidelity to the spatial data. For instance, larger values of the smoothing parameters result in smoother spatial gradients, while smaller values produce rougher fitted trends. Following the approach of Rodríguez-Álvarez et al. (2016a), additional terms were included in the SpATS model to account for other sources of environmental variation and genotype effects in our sorghum breeding trials.

Thus, the SpATS mixed model for each trial is given by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}_s\boldsymbol{\beta}_s + \mathbf{Z}_s\mathbf{s} + \mathbf{Z}_u\mathbf{u} + \mathbf{Z}_g\mathbf{g} + \mathbf{e}, \tag{1}$$

where the vector \mathbf{y} contains the phenotypic observations (grain yield or plant height) arrayed as rows within columns, $\boldsymbol{\beta}$ is a vector of fixed terms including the intercept, a check variety effect, and a resolvable block effect, and \mathbf{X} is the associated design matrix. The fixed (unpenalized) term $\mathbf{X}_s\boldsymbol{\beta}_s$ and the random (penalized) component $\mathbf{Z}_s\mathbf{s}$ form the mixed model expression of the smooth spatial surface, i.e., $f(\mathbf{r}, \mathbf{c}) = \mathbf{X}_s\boldsymbol{\beta}_s + \mathbf{Z}_s\mathbf{s}$, where the vector of random spatial effects \mathbf{s} has covariance matrix \mathbf{S} . The vector \mathbf{u} comprises the mutually independent sub-vectors of random row and column effects accounting for discontinuous field variation, with design matrix $\mathbf{Z}_u = [\mathbf{Z}_r | \mathbf{Z}_c]$ and covariance matrix $\mathbf{U} = \text{diag}(\sigma_r^2\mathbf{I}_r, \sigma_c^2\mathbf{I}_c)$. The vector \mathbf{g} contains the random

genotypic effects of test-cross hybrids and \mathbf{Z}_g is the associated design matrix. We assumed independent genotypic variance, i.e., $\mathbf{g} \sim N(\mathbf{0}, \mathbf{G})$, with $\mathbf{G} = \sigma_g^2\mathbf{I}_g$. The vector \mathbf{e} consist of spatially independent residuals with distribution $\mathbf{e} \sim N(\mathbf{0}, \sigma_e^2\mathbf{I})$. This term, also called nugget, represents the measurement error from each plot.

The SpATS model adopts the P-spline ANOVA (PS-ANOVA) formulation proposed by Lee et al. (2013) to represent the 2D P-splines in the mixed model framework. Detailed descriptions of the design matrices \mathbf{X}_s and \mathbf{Z}_s and the covariance matrix \mathbf{S} under this formulation are given in Lee et al. (2013) and Rodríguez-Álvarez et al. (2016a). In this paper, we present the main result of the PS-ANOVA parameterization, which is the decomposition of the smooth surface into a sum of linear components and univariate and bivariate smooth functions, such that

$$f(\mathbf{r}, \mathbf{c}) = \underbrace{\beta_{s1}\mathbf{r} + \beta_{s2}\mathbf{c} + \beta_{s3}\mathbf{r}\mathbf{c}}_{\mathbf{X}_s\boldsymbol{\beta}_s} + \underbrace{f_1(\mathbf{r}) + f_2(\mathbf{c}) + h_3(\mathbf{r})\mathbf{c} + rh_4(\mathbf{c}) + f_5(\mathbf{r},\mathbf{c})}_{\mathbf{Z}_s\mathbf{s}} \tag{2}$$

where the spatial surface is represented by: linear trends across the row (β_{s1}) and column (β_{s2}) positions and a linear interaction trend (β_{s3}); two main smooth trends across rows [$f_1(\mathbf{r})$] and columns [$f_2(\mathbf{c})$]; two linear-by-smooth interaction terms, where the slope of a linear trend along one covariate (\mathbf{c} or \mathbf{r}) is allowed to vary smoothly as function of the other covariate [$h_3(\mathbf{r})$ or $h_4(\mathbf{c})$, respectively]; and $f_5(\mathbf{r}, \mathbf{c})$ is the pure smooth-by-smooth interaction between column and row trends.

Under this representation, the vector of random spatial effects \mathbf{s} contains five mutually independent sub-vectors \mathbf{s}_k , with $k=1, \dots, 5$ referring to the additive and interaction random components in [2]. Then, the spatial covariance matrix \mathbf{S} is a direct sum of matrices \mathbf{S}_k , that is $\mathbf{S} = \text{blockdiag}(\mathbf{S}_1, \dots, \mathbf{S}_5)$, where each block \mathbf{S}_k depends on a specific smoothing parameter λ_{s_k} (see Rodríguez-Álvarez et al. 2016a for details). Within the mixed model framework, each smoothing parameter is determined by REML as the ratio between the residual variance and the corresponding variance of spatial effects, i.e., $\lambda_{s_k} = \sigma_e^2 / \sigma_{s_k}^2$. Therefore, the smoothness of the spatial surface is tuned by five distinct parameters, applying anisotropic smoothing. The parameterization provides the SpATS model with flexibility to account for both global trends and local variation in the field. Furthermore, the decomposition of $f(\mathbf{r}, \mathbf{c})$ enables a more explicit interpretation of the main patterns of spatial variation.

Implementation of the model

The SpATS model with anisotropic smoothing based on the PS-ANOVA approach by Lee et al. (2013) was fitted with the R package (R Development Core Team 2016) SpATS (Rodríguez-Álvarez et al. 2016b), which is publicly available from CRAN (<https://cran.r-project.org/package=SpATS>). The spatial surface in model [1] was fitted using cubic B-spline bases and second-order penalties, which are commonly used settings in the P-spline framework. Across trials, we used 11 and 31 equally spaced knots for the P-splines in the column and row directions, respectively. In this way, we set approximately one knot for every two rows or columns. Then, the spatial surface contains a total of 425 model parameters to be estimated. These quantities were chosen to provide enough flexibility to the spatial surface. Within the penalized smoothing context, the exact choice of the number of knots is not critical once a certain minimum number of knots is exceeded (Ruppert et al. 2003; Eilers et al. 2015). This number can be equal to the number of rows and columns, i.e., the number of data points in each dimension, or even more. The only limiting factor would be the computational time: the larger the number of knots, the larger the computational effort. It is important to remark that the use of a large number of knots provides flexibility, but in practice, the smoothing parameters are responsible for optimizing the fit to the data.

The estimation procedure implemented in the R package SpATS provides REML-based variance components and computes the empirical best linear unbiased estimates (BLUEs) of fixed effects and the empirical best linear unbiased predictors (BLUPs) of random effects. An important by-product of the procedure is that, for each random effect of the model, an associated effective dimension is computed. The practical implications of the latter concept are considered in the following sections.

The effective dimension of the fitted spatial surface

The effective model dimension (ED) or effective number of parameters of a model is a central concept within the P-spline methodology. It is a measure of complexity of the model components and is mainly a function of the smoothing parameters (Eilers et al. 2015). The effective dimension of a model is computed as the trace of the *hat* matrix \mathbf{H} . If we focus on the spatial part of the SpATS model (1), we have that

$$\tilde{f}(\mathbf{r}, \mathbf{c}) = \mathbf{X}_s \hat{\boldsymbol{\beta}}_s + \mathbf{Z}_s \tilde{\boldsymbol{\varepsilon}} = \mathbf{H}_\beta \mathbf{y} + \mathbf{H}_s \mathbf{y},$$

where \mathbf{H}_β is hat matrix of the fixed component with effective dimension $ED_\beta = \text{trace}(\mathbf{H}_\beta) = \text{rank}(\mathbf{H}_\beta)$, which is

always a constant. More importantly, the total effective dimension of the random (penalized) component of the spatial surface is $ED_s = \text{trace}(\mathbf{H}_s)$, where \mathbf{H}_s is known as the smoother matrix. In this context, the sum of the diagonal elements of \mathbf{H}_s expresses the number of parameters effectively involved in the modelling of the spatial surface. From the PS-ANOVA decomposition used in the SpATS model, we have that $\mathbf{H}_s = \sum_{k=1}^5 \mathbf{H}_{s_k}$, with $k=1, \dots, 5$ referring to the additive and interaction smooth components of the spatial trend as detailed in (2). Thus, we can decompose ED_s as the sum of partial effective dimensions associated with each spatial component:

$$ED_s = \text{trace}(\mathbf{H}_s) = \sum_{k=1}^5 \text{trace}(\mathbf{H}_{s_k}) = \sum_{k=1}^5 ED_{s_k}.$$

Specifically, when a smoothing parameter $\lambda_{s_k} = \sigma_e^2 / \sigma_{s_k}^2 \rightarrow \infty$, then $ED_{s_k} \rightarrow 0$; while for a value of $\lambda_{s_k} = \sigma_e^2 / \sigma_{s_k}^2 \rightarrow 0$, ED_{s_k} approaches the maximum value. The upper bound for ED_{s_k} is determined by the number of knots used to fit the smooth surface. Therefore, ED_{s_k} serves as a reverse indicator of the smoothness of the corresponding component, i.e., the higher the degree of smoothness (larger value of λ_{s_k}), the smaller the number of ED_{s_k} (see Rodríguez-Álvarez et al. 2016a for details).

Consequently, the total effective dimension ED_s can be interpreted as a measure of the magnitude of field variation, with larger values indicating more intense spatial patterns. In addition, the partial effective dimensions ED_{s_k} are indicative of the relative importance of each spatial component in (2). In this case, the magnitudes of specific ED_{s_k} will quantify the contribution of the main and interaction spatial trends to the fitted surface, reflecting the complexity of the spatial pattern.

Generalized heritability based on the genetic effective dimension

As previously mentioned, an effective dimension connected to each variance component of the SpATS model is computed. The effective dimension associated with the genotypic effects (ED_g) is of particularly interesting for plant breeding. $ED_g = \text{trace}(\mathbf{H}_g)$ is a measure of the degree of shrinkage imposed on genotypic effects, where \mathbf{H}_g is the hat matrix for genotypes. In this case, \mathbf{H}_g depends on the regularization parameter $\lambda_g = \lambda_g = \sigma_e^2 / \sigma_g^2$ and transforms the observations into predicted genotypic values, such that $\mathbf{H}_g \mathbf{y} = \mathbf{Z}_g \tilde{\mathbf{g}}$ (see Rodríguez-Álvarez et al. 2016a for details). Therefore, ED_g decreases as shrinkage of genotypic effects increases. Given the properties of the genetic effective dimension, Rodríguez-Álvarez et al. (2016a) proposed a novel expression of heritability:

$$H^2 = \frac{ED_g}{n_g - l}$$

where n_g is the number of genotypes and l is the number of zero eigenvalues of \mathbf{H}_g . The authors showed that this definition corresponds to the generalized heritability introduced by Oakey et al. (2006). Furthermore, in the specific situation when genotypic effects are assumed independent (i.e., ignoring pedigree/marker information), and by ignoring the zero eigenvalues, the following equivalence can be established:

$$H^2 = \frac{ED_g}{n_g} = 1 - \frac{\overline{PEV}}{\sigma_g^2},$$

where \overline{PEV} stands for average prediction error variance of genotype BLUPs.

Note that the right-hand term corresponds to the generalized heritability developed by Welham et al. (2010) and is also equivalent to the heritability given by Cullis et al. (2006). Given that our study does not incorporate a genetic relationship matrix, we can profit from the latter equivalence to perform a straightforward comparison between the heritability estimated by the SpATS model and that obtained from the standard mixed models.

Standard models

Under the standard mixed model framework, we started by fitting a non-spatial model. This baseline model included a random and independent genotypic effect for the test-cross hybrids, a fixed effect for check varieties, a fixed resolvable block effect accounting for the randomization design, and the spatially independent error term $e \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I})$. Then, the non-spatial model was extended by searching for the most appropriate spatial model for each case following the approach of Gilmour et al. (1997). The latter model is referred to as the *best standard spatial* (BSS) model.

The general representation of the BSS model can be formulated as

$$y = \mathbf{X}\beta + \mathbf{X}_s\beta_s + \mathbf{Z}_s s + \mathbf{Z}_u u + \mathbf{Z}_g g + \xi + e, \quad (3)$$

where $\mathbf{X}\beta$ contains the same fixed terms as the non-spatial model. The term $\mathbf{X}_s\beta_s$, in this case, may include linear trends aligned with rows and/or columns to account for global variation, while $\mathbf{Z}_s s$ contains the random part of one or two one-dimensional cubic smoothing splines indexed by row or column positions (see Verbyla et al. 1999 for details). This latter term accounts for additive non-linear global trends. $\mathbf{Z}_u u$, $\mathbf{Z}_g g$, and e are defined as in the SpATS model (1). Finally, ξ is the vector of spatially

correlated residuals modelling local trend, with distribution $\xi \sim N(\mathbf{0}, \mathbf{R})$. The matrix $\mathbf{R} = \sigma_\xi^2 [\text{AR1}(\rho_c) \otimes \text{AR1}(\rho_r)]$ represents the Kronecker product of first-order autoregressive processes across columns and rows, respectively, and σ_ξ^2 is the spatial residual variance.

Following Gilmour et al. (1997), the search for the BSS model was based on diagnostic graphics such as the sample variogram and related plots of residuals. Comparison between candidate models with the same fixed effects was assessed by the REML-likelihood ratio test (REML-LRT). Fixed spatial terms were included in the BSS model when judged significant according to Wald-F test. It is important to note that the BSS model for each trial and trait may represent a simplified version of the full model (3), where the reduced model results from omitting one or more superfluous spatial components.

The standard mixed models were fitted using the ASReml-R package (Butler et al. 2009).

Comparison of spatial methods

The SpATS model was compared with the non-spatial and the BSS models in terms of meaningful parameters for plant breeding application. The following estimates were considered for comparison:

- Genetic variance (σ_g^2) and spatially independent residual variance (σ_e^2).
- Generalized heritability. Estimated following Rodríguez-Álvarez et al. (2016a) for the SpATS model and according to Cullis et al. (2006) for the standard models. These measures are interpreted as broad-sense heritability, which serves as a descriptive measure of precision of a trial, i.e., of the ability to detect genotypic differences among test-cross means.
- Pearson correlations of predicted genotypic values between environments. Given that genotype-by-environment interaction has the same effect on the magnitudes of these correlations for the three models, any increase in their values relative to the non-spatial model will indicate the improvement of precision caused by the spatial models (Qiao et al. 2004; Müller et al. 2010). Only correlations between pairs of environments presenting at least 30 common genotypes were considered.
- Spearman rank correlations between predicted genotypic values from the different models in the same environment. Calculated to compare whether the ranking of genotypes obtained from SpATS and from the standard models differed.

Results

Spatial analysis with SpATS

We start with a detailed treatment of the spatial analysis using the SpATS model illustrated with two contrasting trials regarding the intensity and structure of spatial variability. Table 2 presents the ED_s of the univariate and bivariate spatial smooth components (see Eq. (2)), and their relative contribution to the fitted surface for grain yield in trials DYS05 and DAB08. The magnitudes of the total ED_s indicate that the spatial variation was more intense in DYS05. This is reflected by the higher ED_s or fitted parameters required to model the underlying field trend (111.2 ED_s in DYS05 vs 2.1 ED_s in DAB08). According to the partial ED_s , DYS05 also exhibited a higher complexity in the structure of the spatial surface, where the smooth-by-smooth interaction between trends accounted for most of the field variation (87% of the total ED_s). In contrast, the environmental trend at DAB08 was smoother and less complex as it presented a lower total ED_s and was mostly captured by main smooth effects across row positions. The zero values of ED_s associated with the linear-by-smooth interactions in DAB08 indicate that these terms were not necessary to model the spatial surface.

Figure 1 shows the graphical representations of the fitted spatial trend $f(\mathbf{r}, \mathbf{c})$ and the spatially independent residuals \mathbf{e} for the two example trials, as obtained from the SpATS package. Note that the pictures of the spatial trend use a finer grid than that of the field plots; the P-splines make their computation possible. The spatial surfaces display an irregular patchy pattern in DYS05 and a rather smooth gradient across the field in DAB08. The shape of an evident patch of fertility present in DYS05 was best modelled by considering interactions between column and row trends, as indicated by the partial ED_s (Table 2). Likewise, the previous interpretation of the spatial trend based on the ED_s in

DAB08 coincides with the plot of the fitted surface, which essentially exhibit a one-dimensional gradient across rows. The inspection of the plots of residuals suggests that the spatial patterns have effectively been removed in both trials by the 2D P-spline surface; hence, these residuals could be considered as true random noise. Other plots of residuals and formal tests could also be used to diagnose outliers, model assumptions, or remaining spatial trends after fitting the spatial model. For the latter purpose, an interesting alternative is the variogram computed from the independent residuals, as proposed by Piepho and Williams (2010). This nugget-based variogram can also be obtained with the SpATS package. The ranges of variation of grain yield data (in t/ha) explained by the fitted trends reflect the magnitude of spatial effects in each trial. The comparison between the scales of spatial and residual site variations provides a clear idea of the relative importance of field trends in these trials. For instance, the range of yield variability due to spatial trends in DYS05 was of similar magnitude to that caused by the spatially independent error, while the amount of variation resulting from the latter term was about tenfold the spatial variability in DAB08 (Fig. 1). Again, the higher relevance of spatial trends for trial DYS05 was also indicated by the total ED_s presented in Table 2.

In Table 3, we specify the spatial terms of the BSS models for grain yield in DYS05 and DAB08. The connection between these results and those from the analysis with the SpATS model (Table 2) is not straightforward, since the parameterization of both spatial models is different. Assuming that extraneous variations have been adjusted by both models, here, we stress the differences in modelling global and local trends. For instance, according to the standard spatial analysis, in DYS05, there was only a main global trend in the direction of rows, while the column and interaction trends detected by SpATS were apparently modelled as two-dimensional autocorrelated residuals by the BSS model. The main trend across row positions in DAB08 (see Table 2; Fig. 1) seems to be modelled, under the standard approach, by a small autocorrelation across rows and by a value of ρ_c close to 1. The latter autocorrelation suggests that the trend across columns is actually confounded with the random row effects (Piepho and Williams 2010; Piepho et al. 2015). Finally, the ratios of spatial variance to residual variance (σ_s^2/σ_e^2) were 2.0 for DYS05 and 0.3 for DAB08, indicating a higher intensity of spatial variation in the former trial (Dutkowski et al. 2002; Zas 2006). The latter results coincide with the interpretation based on the total effective dimensions of the spatial surfaces given in Table 2.

The effective dimensions associated with the fitted spatial trends (ED_s) for all trials and both traits are given in Fig. 2. For simplicity, the partial ED_s for the five smoothing terms of the SpATS model are grouped as: ED_s of the

Table 2 Spatial effective dimensions (ED_s) of the smooth surface components fitted by the SpATS model and its relative contribution (%) for grain yield in two example trials

Spatial smooth terms	DYS05		DAB08	
	ED_s	%	ED_s	%
Additive trends				
$f_1(\mathbf{r})$	3.0	3	1.4	67
$f_2(\mathbf{c})$	4.2	4	0.2	10
Interaction trends				
$h_3(\mathbf{r})\mathbf{c}$	1.9	2	0.0	0
$rh_4(\mathbf{c})$	5.5	5	0.0	0
$f_3(\mathbf{r}, \mathbf{c})$	96.6	87	0.5	24
Total	111.2	100	2.1	100

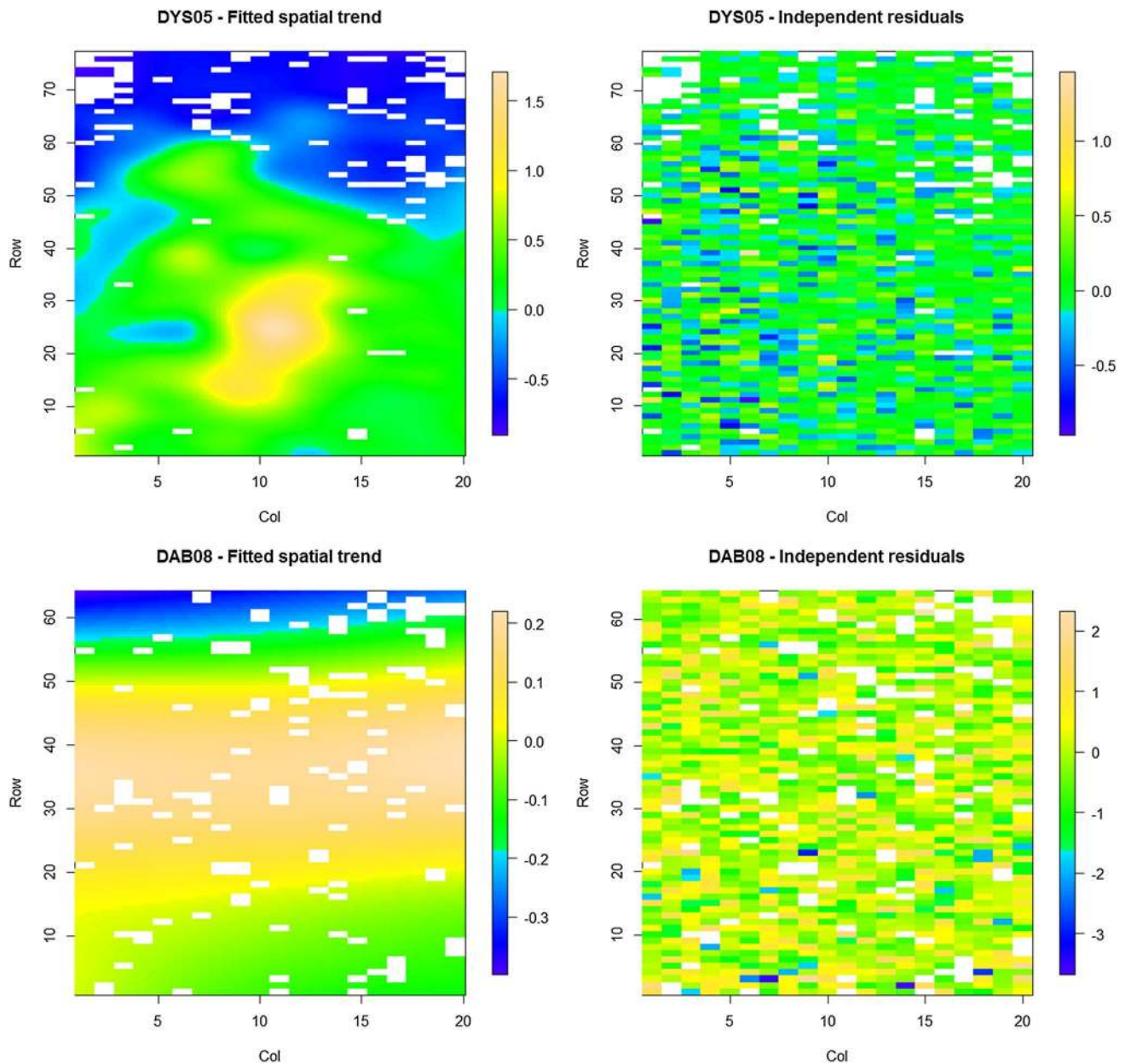


Fig. 1 Fitted spatial trend and spatially independent residuals from the SpATS model for grain yield in trials DYS05 (*top*) and DAB08 (*bottom*) plotted against row and column positions. Scales of grain yield variation expressed in t/ha

Table 3 Spatial terms, estimates of autocorrelations, and variance components for spatially dependent (σ_r^2) and independent residuals (σ_e^2) from the *best standard spatial* (BSS) models fitted to grain yield data in two example trials

Trial	BSS model ^a	ρ_r	ρ_c	σ_r^2	σ_e^2
DYS05	R + Spl(r) + AR1 \times AR1 + n	0.87	0.67	0.103	0.064
DAB08	R + AR1 \times AR1 + n	0.24	0.96	0.201	0.611

^aR: random row effects; Spl(r): cubic smoothing spline indexed by row positions; AR1 \times AR1: correlated residual modelled as two-dimensional first-order autoregressive process; n : spatially independent residual (or *nugget* variance)

additive smooth trends and ED_s of the interactions between trends. The intensity of spatial variation and the complexity of the fitted surfaces were highly variable across sites and traits. For instance, the environmental trends for grain yield at DYS05 and BIL05 or HER05 for plant height present a large number of ED_s and a significant contribution of the trend interaction terms, indicating strong and complex patterns of field variation. Others cases, such as DAB08 for grain yield and LIV08 for yield and plant height, show lower total ED_s , reflecting smoother spatial surfaces that were mainly described by additive one-dimensional trends.

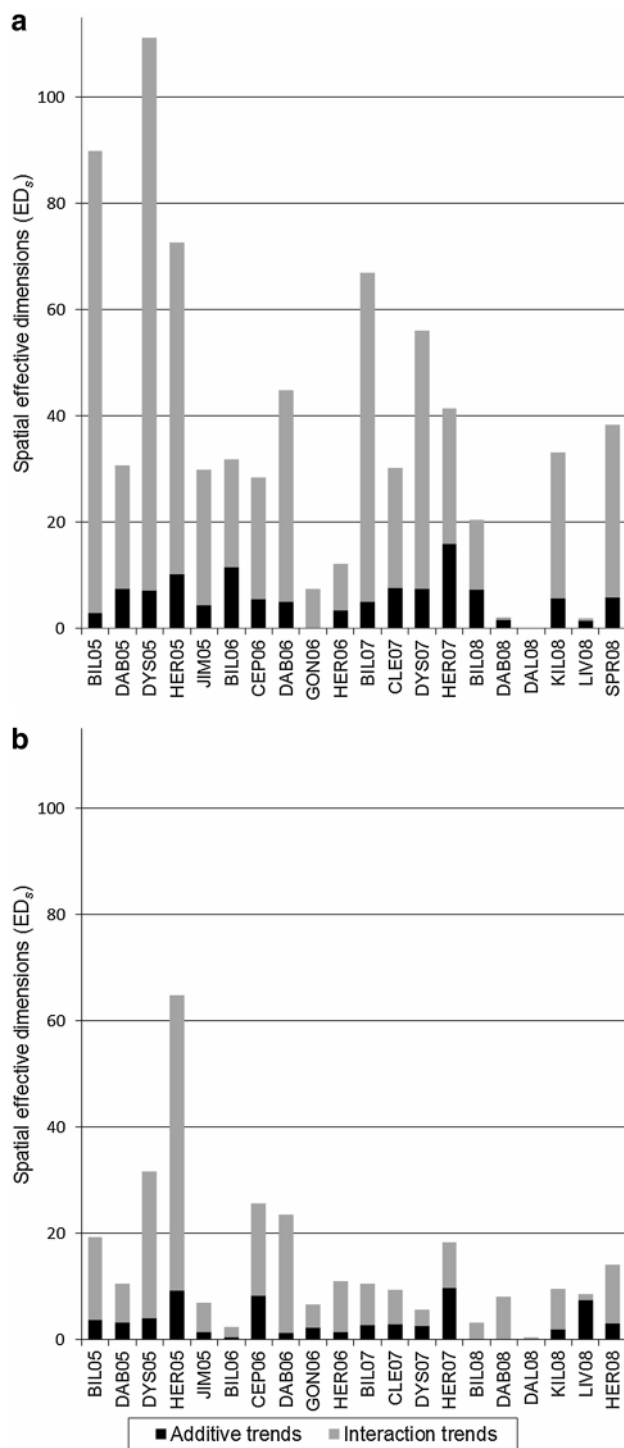


Fig. 2 Effective dimensions (ED_s) associated with additive and interaction trends of the spatial surfaces fitted by the SpATS model for grain yield (a) and plant height (b) in the sorghum breeding trials. Note: data were not available for grain yield at HER08 and for plant height at SPR08

In general, the intensity of spatial variation for grain yield was higher than for plant height, with median total ED_s of 31 and 10, respectively. In most instances, the smooth

trend interactions represented the major components of the spatial surface. This is reflected by the median ED_s associated with interaction effects, which were 82 and 79% of the total ED_s for yield and plant height, respectively. The latter results highlight the importance of modelling interactions between row and column trends and reveals complex structures of field variation in the sorghum data set.

Standard spatial analysis

A summary of the main features of the BSS models fitted to the sorghum data set is reported in Table 4. Details of the BSS model identified in each of the 20 trials for both traits are presented in Table 5. The results in Table 4 show that most of the trials required terms accounting for global trends, local variation, and nugget effect. Autocorrelations (ρ) along rows and columns were predominantly large and similar for both traits, as reflected by their median. Over 80% of the autocorrelation coefficients were larger than 0.60, indicating strong spatial variation that could be interpreted as a combination of large-scale gradients and patchy patterns according to the standard approach. When considering the models with nugget, the importance of the spatial variance relative to the spatially independent residual variance was generally higher for grain yield. The predominance of random noise in plant height measurements indicates that this trait was less influenced by spatial effects in the field. This is consistent with the generally lower effective dimensions of the spatial surfaces estimated for plant height (see Fig. 2). Furthermore, there was a strong positive correlation between the $\sigma_{\xi}^2/\sigma_e^2$ ratio and the total ED_s across the whole data set ($r=0.69$).

Comparison of SpATS and the standard method

The estimates of trial genetic variability from SpATS and the BSS models were generally similar for both traits (Fig. 3). Small differences were evident for grain yield at some environments, where the estimates increased or decreased from one model to the other without a clear tendency. More marked discrepancies were observed between the genetic variances from the non-spatial model and those from both spatial models for grain yield (not shown). This suggests that ignoring the adjustment for spatial trends in yield data can lead to either overestimating or underestimating the genetic variability.

The SpATS model and the BSS models reduced the spatially independent residual variance compared with error variance of the non-spatial model in both traits (Fig. 4). In general, the relative decreases in σ_e^2 were larger for grain yield, with the spatial models achieving a mean reduction by 49% for grain yield and by 22% for plant height. These reductions reflect the ability of both methodologies to

Table 4 Number of times the *best standard spatial* (BSS) models for the 20 trials included terms accounting for global and local trends, and median of estimated spatial parameters

	Grain yield	Plant height
Number of trials including:		
Global trend terms	15	11
Correlated residuals (AR1xAR1)	17	17
Nugget effect	17	14
Median of spatial parameters:		
ρ_r	0.82	0.87
ρ_c	0.73	0.69
Proportion (%) of correlated error ^a	52	25

^aRelative to the sum of correlated and independent residual variances

account for field variation not adjusted by the randomization-based model. Exceptionally, the adjustment of spatial trend for plant height caused a large decrease in σ_e^2 at trial HER05. Note that field trend in this case was particularly important, presenting the highest total ED_s for plant height and a major contribution of interaction effects (see Fig. 2). In general, the BSS models estimated smaller values of σ_e^2 compared to the SpATS model. The spatially independent component from SpATS and the BSS models represented,

on average, 66 and 60% of the residual variance from the non-spatial model, respectively.

Figure 5a shows the changes in the estimates of trial heritability from the non-spatial model to the SpATS model. The spatial method increased the heritability in most instances, with levels of improvement in precision being generally higher for grain yield. Not surprisingly, a remarkable increase in heritability was also achieved for plant height in HER05 after fitting trends with SpATS. Trial heritabilities estimated by both spatial methods were very consistent for plant height (Fig. 5b). However, more variation in the estimates was observed for grain yield, where similar or slightly higher heritabilities were obtained with the SpATS model in most trials. Finally, notice that heritabilities were, in general, lower for grain yield, which was the trait affected by stronger spatial variation (as inferred from the total ED_s in Fig. 2).

The Pearson correlations of genotype BLUPs between environments obtained from the two spatial methods were, on average, slightly higher than those obtained from the non-spatial model in both traits (Fig. 6). The mean correlations for grain yield increased from 0.04 to 0.10 and 0.09 after applying the BSS models and SpATS, respectively (Fig. 6a). For plant height, both spatial models caused a mean increase of 0.05 in the correlations, changing from

Table 5 Details of the *best standard spatial* (BSS) models in each trial for grain yield (GY) and plant height (PH)

Trial	BSS model for GY ^a	BSS model for PH ^a
BIL05	Lin(<i>r</i>) + AR1xAR1 + <i>n</i>	Lin(<i>r</i>) + AR1xAR1(<i>c</i>) + <i>n</i>
DAB05	R + AR1xAR1 + <i>n</i>	Lin(<i>c</i>) + AR1xAR1 + <i>n</i>
DYS05	R + Spl(<i>r</i>) + AR1xAR1 + <i>n</i>	Lin(<i>r</i>) + AR1xAR1 + <i>n</i>
HER05	Lin(<i>r</i>) + Spl(<i>c</i>) + AR1xAR1 + <i>n</i>	Spl(<i>r</i>) + Lin(<i>c</i>) + AR1xAR1 + <i>n</i>
JIM05	C + AR1xAR1 + <i>n</i>	R + AR1xAR1 + <i>n</i>
BIL06	Spl(<i>c</i>) + AR1xAR1 + <i>n</i>	C + AR1xAR1
CEP06	R + AR1xAR1 + <i>n</i>	Spl(<i>c</i>) + AR1xAR1 + <i>n</i>
DAB06	C + Spl(<i>r</i>) + AR1xAR1 + <i>n</i>	AR1xAR1 + <i>n</i>
GON06	Spl(<i>c</i>) + AR1 + <i>n</i>	R + C + AR1
HER06	R + Spl(<i>r</i>) + Lin(<i>c</i>)	R + C + Lin(<i>c</i>) + AR1xAR1 + <i>n</i>
BIL07	Lin(<i>c</i>) + AR1xAR1 + <i>n</i>	R + Spl(<i>c</i>)
CLE07	R + Lin(<i>r</i>) + AR1xAR1 + <i>n</i>	Lin(<i>c</i>) + AR1xAR1 + <i>n</i>
DYS07	Spl(<i>c</i>) + AR1xAR1 + <i>n</i>	Lin(<i>c</i>) + AR1xAR1 + <i>n</i>
HER07	R + Spl(<i>r</i>) + Spl(<i>c</i>) + AR1xAR1 + <i>n</i>	Lin(<i>r</i>) + AR1xAR1 + <i>n</i>
BIL08	R + C + Spl(<i>c</i>)	C
DAB08	R + AR1xAR1 + <i>n</i>	Lin(<i>c</i>) + AR1 + <i>n</i>
DAL08	R + C + Lin(<i>c</i>)	C
HER08	–	AR1
KIL08	R + Lin(<i>r</i>) + AR1xAR1 + <i>n</i>	AR1xAR1 + <i>n</i>
LIV08	R + C + AR1xAR1 + <i>n</i>	AR1xAR1 + <i>n</i>
SPR08	Spl(<i>c</i>) + AR1xAR1 + <i>n</i>	–

^aSpl(·): cubic smoothing spline indexed by row (*r*) or column (*c*) positions; Lin(·): linear regression on row (*r*) or column (*c*) positions. R: random row effects; C: random column effects; AR1 and AR1xAR1: correlated residuals modelled as one- and two-dimensional first-order autoregressive process, respectively; *n*: spatially independent residual (nugget effect). Note that all the models included a fixed block effect

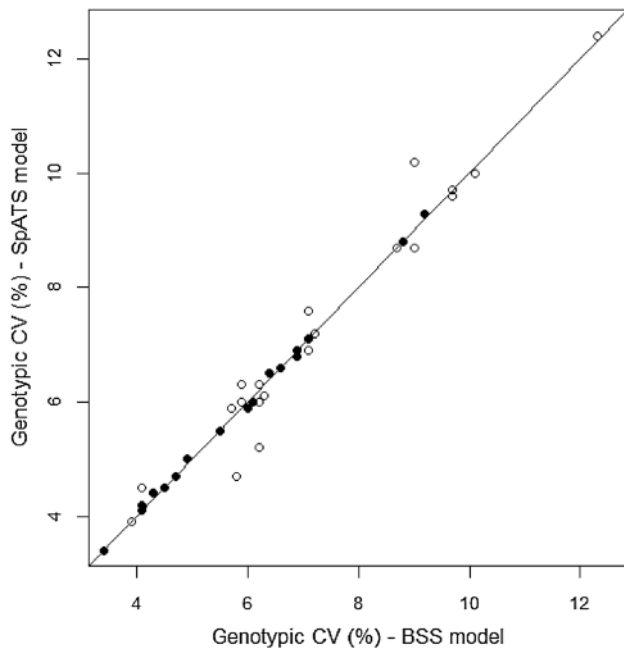


Fig. 3 Comparison of genotypic variability estimated by the BSS and the SpATS models, expressed as coefficient of variation (CV), for grain yield (open circle), and plant height (filled circle). The diagonal line indicates identical values

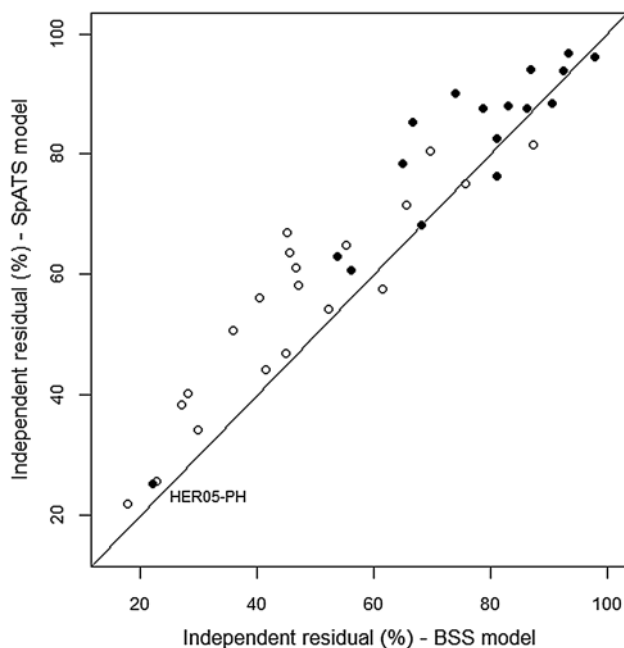


Fig. 4 Comparison of spatially independent residual variance (σ_e^2) from the BSS and the SpATS models, expressed as percentage (%) of the residual variance in the non-spatial model, for grain yield (open circle) and plant height (filled circle). The diagonal line indicates identical values. The labelled data point corresponding to plant height at trial HER05 (HER05-PH) is mentioned in the text

0.46 to 0.51 (Fig. 6b). At the same time, the spatial methods reduced the variation of estimated correlations for the latter trait. The higher mean correlations between environments in plant height reflect a lower influence of genotype-environment interaction.

For illustration purpose, Fig. 7 presents the BLUPs of genotype effects from SpATS and the BSS models for grain yield in the example trials DYS05 and DAB08. Differences in the rankings were small for both environments. However, changes in the order of genotypes were more evident at DYS05 (Fig. 7a), an environment where, as previously noted, the nature of spatial variation was more complex. The predicted rankings established by both spatial methods were also consistent for the rest of the data set, with mean Spearman correlations across trials of 0.970 for grain yield and 0.989 for plant height. As expected, the rankings of genotype were more dissimilar between SpATS and the non-spatial models. Rank correlations for yield between these models ranged from 0.500 at DYS05 (where $ED_s = 111.2$) to 0.926 at DAL08 (where $ED_s = 0.0$), with a mean value of 0.802. In the case of plant height, correlations were generally higher, varying from 0.767 at HER05 (where $ED_s = 64.8$) to 0.988 at DAL08 (where $ED_s = 0.4$) and a mean value of 0.944.

As suggested by one of the reviewers, we tried to implement a single-step model selection strategy with the standard method using the full model (3) across our data set. Convergence problems were evident in 8 out of 20 trials for grain yield and in 10 out of 20 trials for plant height. It was possible to decrease the rate of failure by relaxing convergence criteria. However, given that we know the full model was a misspecified one, tuning strategies should not be used to get convergence. The failures to converge reflected identifiability problems for the situation when the AR1xAR1 structure, global trend terms, and the nugget are included in the same model. In contrast, the SpATS model did not suffer from this confounding difficulty; the three types of field variation were fitted in a stable way.

Discussion

Spatial analysis with SpATS

This study presented the SpATS model as a suitable alternative to the standard spatial models for the adjustment of field trends in sorghum genetic trials. We reported a first application of the new spatial model to a real and extensive plant breeding field testing. This method fits a smooth surface to account for all sources of continuous environmental variation. The mixed model representation of SpATS features the joint modelling of additive one-dimensional trends plus interactions between trends in the

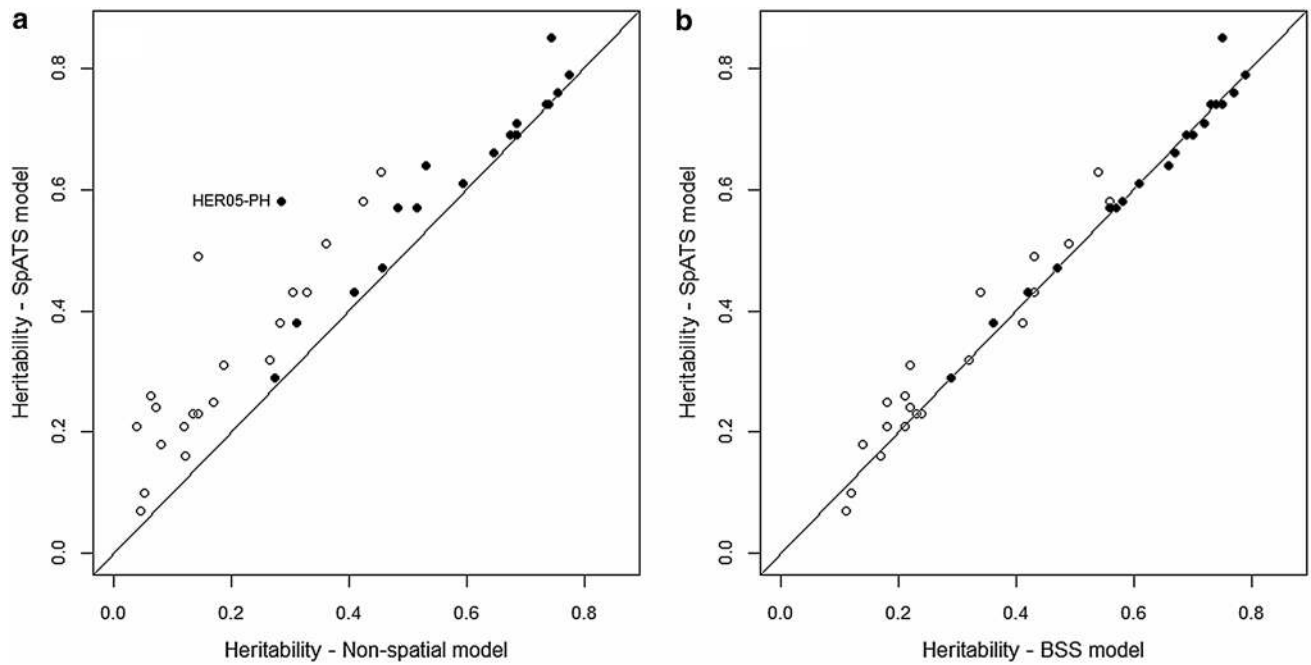


Fig. 5 Comparison between estimates of heritability from the non-spatial and SpATS models (**a**), and from SpATS and the BSS models (**b**) for grain yield (*open circle*) and plant height (*filled circle*). The

diagonal lines indicate identical values. The labelled data point corresponding to plant height at trial HER05 (HER05-PH) is mentioned in the text

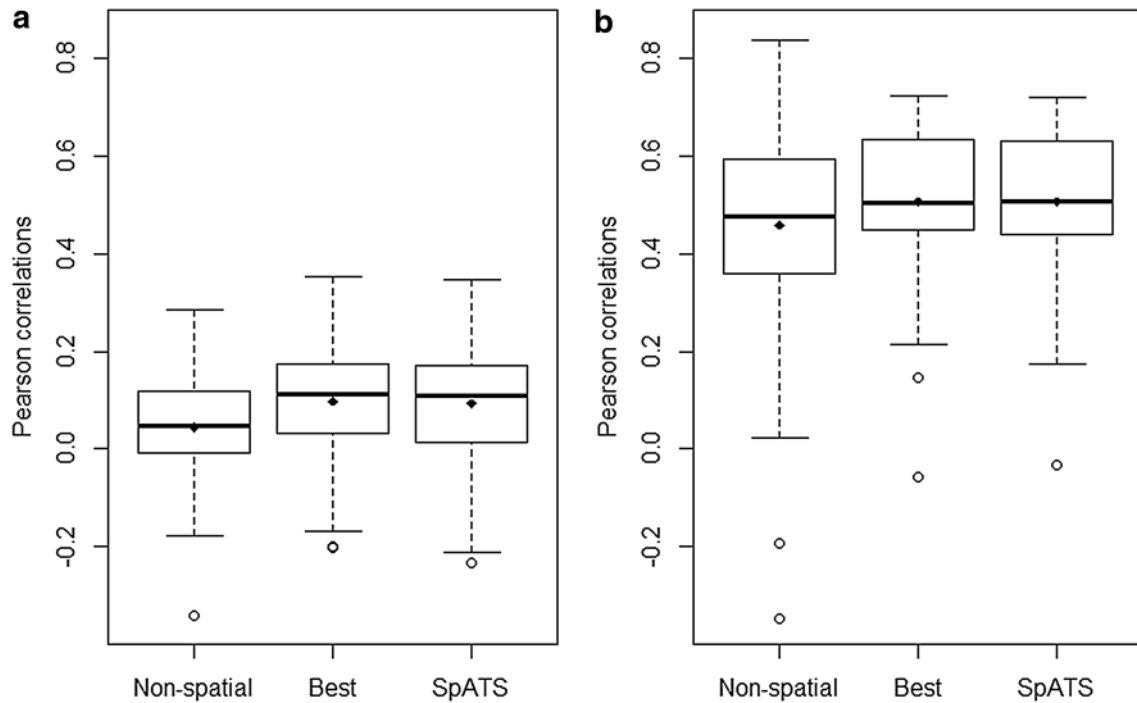


Fig. 6 Correlations of genotype BLUPs between environments from each model for grain yield (**a**) and plant height (**b**)

row and column directions. Moreover, the model specification assigns different degrees of smoothing to each additive and interaction effect by means of specific smoothing

parameters. These weighting terms, which are automatically tuned by REML-based variance components, shrink irrelevant effects to optimize the fit of the spatial surface.

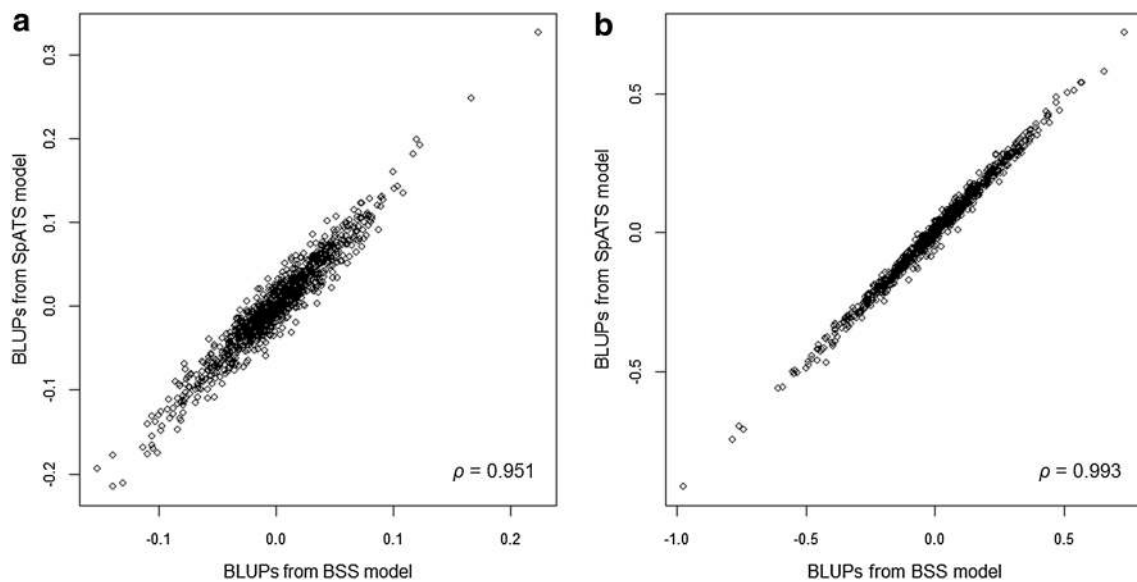


Fig. 7 Genotype BLUPs from the BSS model and the SpATS model and Spearman rank correlations (ρ) for grain yield in trials DYS05 (a) and DAB08 (b)

We have stressed the practical importance of the effective dimension of the model as an integral part of spatial analysis with SpATS. This study highlights how ED_s can be used to interpret the intensity and the structure of spatial variation. The ED_s is a very appealing tool to quantify the magnitude of spatial effects, reflecting the amount of smoothing of the spatial surface and allowing an easy identification of the main patterns of field heterogeneity. Furthermore, the genetic effective dimension was used to compute a generalized heritability in the context of analysis with SpATS (Rodríguez-Álvarez et al. 2016a). This novel expression of heritability is valid for more general situations commonly found in plant genetic trials, *e.g.*, when data are unbalanced and/or when residuals are spatially correlated. Equivalent definitions of generalized heritability were also proposed by Cullis et al. (2006) and Welham et al. (2010) in the context of standard mixed model analyses.

To subject the new method to a hard evaluation, we analysed large-scale sorghum breeding trials arrayed as partially replicated (*p*-rep) designs (Cullis et al. 2006). These experiments are characterised by the absence of the traditional blocking factors, allowing very few or no design features to be retained in the randomization-based model. Moreover, the use of partially replicated experiments assumes that field trend affecting unreplicated genotypes can be properly predicted by the spatial model (Payne 2006). Consequently, the analysis of *p*-rep designs requires the inclusion of spatial parameters as an essential add-on component for an efficient testing of genetic material. The results of our study demonstrate the effectiveness of SpATS

to account for spatial trend and predict adjusted genotypic values under these circumstances. The SpATS model adjusts a continuous surface across the whole field. A more refined modelling could consider a discontinuity in spatial trend by fitting a different surface within each block. Even though the former approach is more conservative, we consider that it reflects the structure of trial design and should be a realistic model for most commonly used experiments in plant breeding. Furthermore, the smoothness of the spatial surface fitted by SpATS is controlled by five different terms, providing enough flexibility for an appropriate fit of the spatial trend.

Comparison of SpATS and the standard method: parameterization

The SpATS model presents similarities with the full formulation of the standard spatial model (see Eq. 3). Both models contains two one-dimensional spline terms, each one fitted as the sum of a fixed linear trend and a random non-linear component. In addition, discontinuous spatial trends are accounted for by random row and column effects in both cases. However, there is a major difference between both models when accounting for the remaining spatial variation. The standard model fits a separable AR1 process, whereas SpATS uses P-spline interaction terms.

This difference in parameterization affects the way in which SpATS and the standard methods model field variation. Under the standard mixed model approach, gradients across the field can be adjusted by blocking factors and by one-dimensional polynomials or splines along rows and

columns. However, more complex two-dimensional gradients that do not align well with row and column directions are expected to affect field trials as well. The structure of spatial variation found in our data set (Fig. 2) and in the previous studies of agricultural and forest field trials demonstrate that fitting only additive gradients in one dimension may result in insufficient modelling of global trend (Federer 1998; Fu et al. 1999; Taye and Njuho 2008). It is in principle possible to extend the standard spatial model with additional fixed terms, like a linear \times linear interaction term (Federer 1998), and random terms, like a smoothing spline interaction term, but these extensions were never used under the standard approach and are prone to cause problems (Gilmour 2000). In this research, we showed that the SpATS model is able to account for intricate patterns of large-scale variation by explicitly modelling the interactions between global trends along rows and columns.

A common practice in spatial analysis is to fit an autoregressive model, originally proposed to adjust for local trend, and assume that it is flexible enough to also account for global trend (Zimmerman and Harville 1991; Dutkowski et al. 2006; Piepho et al. 2008). A more conservative approach considers that the underlying spatial correlation is likely to hold only within blocks and that large-scale trend is partially accounted by blocking factors (Williams et al. 2006; Piepho and Williams 2010). However, continuous non-stationary trends across the whole field can be better fitted by specific spatial terms in the model, as was suggested in the seminal paper by Gilmour et al. (1997). Following their approach, we found that additional terms accounting for global variation could not have been ignored in most of the sorghum trials and for both traits (see Table 4). Several studies using real and simulated data have shown that underfitting global trend may cause the variance of treatment differences to be underestimated (Zimmerman and Harville 1991; Brownie et al. 1993; Brownie and Gumpertz 1997). This false improvement in precision can be particularly negative in plant breeding trials as it reduces the efficiency of selection decisions. Furthermore, Brownie and Gumpertz (1997) reported that local trend is overestimated in presence of unaccounted large-scale trend. Given that global and local model terms are actually “competing” to fit part of the same spatial variation, the estimated covariance parameters will vary according to the global terms included in the spatial model. This inconsistency in the estimates of autocorrelations was also observed in our study during the search of the BSS models (not shown). The aforementioned situation raises the issue of parameter identification when both global and local trend are trying to be fitted. Therefore, spatial parameters should be interpreted with special care under the standard approach. Conversely, the new spatial method based on 2D P-splines simplifies the problem of spatial model identification by

always modelling all types of field trend as a single continuous process. This unified modelling avoids the necessity of distinguishing between global and local trend. Both forms of continuous variation are simultaneously fitted by the flexible interaction surface with anisotropic smoothing. As a result, SpATS provides a straightforward representation of the spatial trend that is easy to interpret. Moreover, the ANOVA-type decomposition of the smooth surface facilitates the characterization of the spatial trend, providing additional insight into the structure of field variation.

Another difference between both spatial methods was evident regarding the estimation of the residual variance. In our data set, the standard spatial models exhibited a clear tendency to estimate smaller spatially independent components than the SpATS model (Fig. 4). The same discrepancy was reported by Rodríguez-Álvarez et al. (2016a) in a simulation study where they analysed data generated according to different autoregressive models with nugget. These authors showed that, when the autocorrelations are large ($\rho_r = \rho_c = 0.9$), the SpATS model provides relatively accurate estimates of the random error variance, whereas the autoregressive model tends to underestimate this term. The possibility of confounding the spatial component with the nugget variance when fitting autoregressive models in field trials was also reported by Cullis et al. (1998) and extensively discussed in Piepho et al. (2015). Given the large autocorrelations estimated in most of our sorghum trials, we may suggest that SpATS performed generally better in identifying the true spatially independent residuals, while the BSS models were actually modelling part of the random error as spatially correlated data. This potential confounding of parameters in autoregressive and other non-linear spatial models with nugget causes frequent convergence problems (e.g., Dutkowski et al. 2006; Müller et al. 2010; Liu et al. 2015; Rodríguez-Álvarez et al. 2016a). When convergence cannot be reached, one could fall back to alternative models without nugget effects (Müller et al. 2010; Leiser et al. 2012). This strategy is far from attractive given that the potential best fitting model would be deliberately ignored. Furthermore, our research (Table 3) and other studies demonstrated that a spatially independent component accounting for measurement error is frequently required (e.g., Cullis et al. 1998; Qiao et al. 2000; Liu et al. 2015). In contrast to the standard spatial modelling approach, the SpATS model always fits a random error variance on top of the spatial surface and, in our experience, it always converges readily; see also Rodríguez-Álvarez et al. (2016a). As a reviewer suggested, in addition to the identifiability issues mentioned above for the standard approach, it cannot be excluded that the difference in convergence performance between SpATS and the standard models may be related to the standard method using a covariance structure that is non-linear in the variance

parameters, while the covariance structure of SpATS is linear in the parameters. Further study is required here.

Comparison of SpATS and the standard method: performance

The comparison between SpATS and the best fitting standard spatial models revealed a similar performance for the evaluation criteria considered in this paper. Besides the differences discussed above, both methods caused similar reductions in the spatially independent residual variance compared with the error of the non-spatial model. This changes indicate the magnitude of spatial variation adjusted by the spatial models for both traits. The generally large decreases in the random error component (>30%) obtained for grain yield reflect that strong spatial trends affected this trait in most trials (Stroup et al. 1994; Yang et al. 2004). The lower reductions observed for plant height could be related to the dominant presence of random environmental variation (see Table 3). Interestingly, the same inferences can be drawn by considering the higher ED_s that were usually associated with grain yield trends (Fig. 1). The larger number of parameters effectively estimated by SpATS to better approximate the underlying spatial surface reflected the higher intensity of field trends for grain yield data. The ability of the ED_s to indicate the relative importance of spatial variation was evidenced by the strong positive association between the number of ED_s and the ratio of spatial to spatially independent variance from the standard models.

In general, the estimates of genetic variance from the SpATS model were comparable to those obtained by the BSS models. The inconsistencies between both models observed in some cases may result from the impossibility to clearly identify the genetic and the environmental variation in presence of spatial correlation. Several simulation studies have shown that unadjusted patchiness in the field may inflate the genetic variance (e.g., Loo-Dinkins et al. 1990; Magnussen 1993, 1994). This identification problem was apparent across candidate BSS models, where the autoregressive models ignoring the nugget estimated higher trial genetic variances than the better-fitting models using nugget (data not shown). The overestimation of genetic variation when adjusting autoregressive models without nugget was also reported by Dutkowski et al. (2002) in tree breeding trials and by Rodríguez-Álvarez et al. (2016a) using simulated data. In addition, the latter authors showed that SpATS produced more accurate estimates of genetic variance, which were highly consistent with those obtained from the best fitting standard model including the nugget. However, more extensive assessments of the SpATS model would be still necessary with respect to the validity of estimates when spatial variation is present.

Several studies considered the changes in heritability to measure the impact of alternative models on the efficiency of plant breeding evaluations (e.g., Smith et al. 2001; Welham et al. 2010; Sarker and Singh 2015). Following this approach, we used the generalized heritability to compare the performance of the SpATS model and the standard spatial models. The adjustment of spatial trends with the new spatial model led to levels of heritability equivalent to the standard models in all the sorghum trials. The increases in grain yield heritability compared to the randomization-based model were broadly consistent with the results from standard spatial analysis of sorghum breeding trials in West Africa (Leiser et al. 2012). The improvement in precision, measured as the increase in the correlation of genotype predictions between environments, was generally the same for both spatial methods. Similar magnitudes of improvements through standard spatial analysis were previously reported by Leiser et al. (2012) in sorghum, but smaller increases were achieved for wheat, sugar beet, and barley breeding trials (Qiao et al. 2004; Müller et al. 2010).

The analysis with SpATS affected the predictions of genotypic values, as the ranking of genotypes changed after modelling the spatial trends. A bigger impact on genotype ranks was usually observed in cases where the fit of a smooth surface produced larger increases in broad-sense heritability. Our results showed high consistency in the ranking of genotypes predicted by the SpATS model and the BSS models for all cases. This indicates that the use of the new spatial method would hardly produce changes in selection decisions compared to the more refined spatial models. The consistent but small changes in predicted rankings may be a consequence of the differences discussed above related to how both spatial methods accommodate global and local trends.

Comparison of SpATS and the standard method: modelling strategy

In this paper, we used a single-step modelling strategy to perform the definite spatial analysis in every trial. Furthermore, the same SpATS model was applied for individual-trial analysis across the whole data set. This approach differs from the common modelling procedure based on sequential fitting of alternative spatial models for each trial. The latter practice may be a limitation for efficient routine application given that several model selection steps are required to arrive at a final spatial model. To perform the standard spatial analysis in the present paper, we inspected alternative AR1 models. However, the number of potential candidate models increases if other spatial covariance structures are also considered. A strategy to simplify the model selection process may be to restrict the number of candidate spatial models, potentially reducing the

efficiency of analysis. A remarkable attempt to maximize efficiency of plant breeding trials through standard spatial analysis was reported by Leiser et al. (2012), who fitted 91 different models for each trial to identify the best models in 17 environments. Unfortunately, these efforts for further modelling usually result in modest benefits relative to simpler models. Alternative spatial methods based on kriging are also time-consuming and difficult to apply in practice (Zas 2006; de la Mata and Zas 2010).

Our approach using SpATS accounted for all types of spatial variation by fitting a single model rather than using a multi-step modelling procedure. Under this simplified strategy, model selection steps required to identify the appropriate spatial correlation and/or global trend terms are not needed; both local and global trends are automatically modelled in a single step by the smooth surface. The SpATS approach relies on the estimation procedure to effectively reduce the influence of the smooth surface components that are not needed. This implicit model selection is automatically tuned by specific smoothing parameters (or penalties) and is reflected in the ED_s . Accordingly, after convergence, the ED_s of unimportant components will tend to zero, meaning that these terms are not contributing to the complexity of the spatial model. The new method also simplifies the practice of using diagnostic graphics, such as variograms, to guide model selection. The reason is that the selections steps required to fit global and local trends under the standard method are reduced to one with SpATS, and thus, the diagnostic plots associated to those steps are essentially skipped. Random row and column effects were fitted by default in our SpATS model, as discontinuous spatial effects were also present in most cases (data not shown). The inclusion of these effects in a default spatial model is justified by the frequent existence of non-smooth effects caused by blocking factors or extraneous variation (e.g., Piepho and Williams 2010; Liu et al. 2015). We set the same number of equally spaced knots in each dimension of the 2D P-spline for every trial. These quantities were chosen to be so many as to ensure ample flexibility to the smoother. Other studies on spatial analysis with P-splines reported that models using different numbers of knots produced similar fits and results (Cappa and Cantet 2007; Cappa et al. 2011). Moreover, Eilers et al. (2015) demonstrated that, once a sufficient number of knots has been chosen, optimizing their quantity is not worthwhile, because the smoothing parameters will regulate the smoothness of the fit to optimize the bias-variance trade-off.

For the present research, we have used a general SpATS model considering the design and treatment factors of our data set. However, it is noteworthy that the mixed model formulation of SpATS enables more refined model building/selection according to specific situations. For instance,

having an ED of zero is equivalent to an associated variance component being zero. It implies that we could use any tests that evaluate the relative fit of a variance model (e.g., REML-LRT, AIC) to perform model selection.

The results from this study showed that the SpATS model performed comparably to more refined and site-specific spatial models. One advantage of the novel method is that all types of continuous spatial variation and genetic effects can be modelled simultaneously in a single modelling step. As Dutkowski et al. (2006) pointed out, this approach should be superior to fitting all terms in a multi-step process as it will avoid parameter identification problems derived from confounding spatial heterogeneity with genetic heterogeneity due to aggregation of related genotypes. An additional benefit is that the SpATS model may be useful to improve the efficiency of two-stage analysis of multi-environment trials (MET). The reason is that the same flexible model can be fitted in the first stage to account for the spatial surfaces of all the trials, obtaining adjusted genotype means to be used in the second stage. The gain in speed of analyses using the new method results from the fact that less computational steps would be needed to identify an appropriate spatial model for each trial.

Conclusion

The SpATS model provided a flexible and efficient alternative to account for spatial patterns in the sorghum breeding field trials. The performance of the new model was equivalent to the more elaborate standard spatial models when considering the improvement in precision and the predictions of genotypic values. The suitability of SpATS was consistent across trials and traits exhibiting different magnitudes of heritability and complexity of spatial variation. A major advantage of the new model over existing techniques is that global and local trends are jointly modelled by the smooth surface. Moreover, we used a general SpATS model to adequately fit all experiments, which avoids the examination of several candidate models for each trial. Given the results of this study, the use of the new method should be considered as a simple and effective strategy to optimize the practical application of spatial analysis in plant breeding trials.

Author contribution statement FvE conceived the research. JV, DJ, MM, and FvE designed the research. JV performed statistical analyses and wrote the manuscript. MXRA, MB, and PE supported the application and understanding of the SpATS methodology and corresponding R package. DJ coordinated the field trials and the data collection. MXRA, MB, MM, and FvE edited the manuscript. All authors read and approved the final manuscript.

Acknowledgments We thank two anonymous reviewers for insightful comments and suggestions on earlier versions of the manuscript. The trial data used in this research was generated with funding from the Grains Research and Development Corporation (GRDC) of Australia. J.G. Velazco acknowledges financial support from the National Institute of Agricultural Technology (INTA) of Argentina, Res. DN 1126/13. M.X. Rodríguez-Álvarez 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. M. Malosetti and F.A. van Eeuwijk worked on this paper as part of the Integrated Breeding Program.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bartlett MS (1978) Nearest neighbour models in the analysis of field experiments. *J R Stat Soc Ser B* 40:147–174
- Basford KE, Williams ER, Cullis BR, Gilmour A (1996) Experimental design and analysis for variety trials. p. 125–138. In: Cooper M, Hammer GL (eds) *Plant adaptation and crop improvement*. CAB Int., Wallingford
- Bernal-Vasquez AM, Möhring J, Schmidt M, Schönleben M, Schön CC, Piepho HP (2014) The importance of phenotypic data analysis for genomic prediction—a case study comparing different spatial models in rye. *BMC Genom* 15(1):646. doi:10.1186/1471-2164-15-646
- Brownie C, Gumpertz ML (1997) Validity of spatial analyses for large field trials. *J Agric Biol Environ Stat* 2(1):1–23
- Brownie C, Bowman DT, Burton JW (1993) Estimating spatial variation in analysis of data from yield trials: a comparison of methods. *Agron J* 85:1244–1253
- Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) *Mixed models for S language environments, ASReml-R reference manual*. Training and development series, No QE02001. QLD Department of Primary Industries and Fisheries, Brisbane
- Cappa EP, Cantet RJC (2007) Bayesian estimation of a surface to account for a spatial trend using penalized splines in an individual-tree mixed model. *Can J For Res* 37:2677–2688
- Cappa EP, Lstiburek M, Yanchuk AD, El-Kassaby YA (2011) Two-dimensional penalized splines via Gibbs sampling to account for spatial variability in forest genetic trials with small amount of information available. *Silvae Genet* 60:25–35
- Cappa EP, Muñoz F, Sanchez L, Cantet RJC (2015) A novel individual-tree mixed model to account for competition and environmental heterogeneity: a Bayesian approach. *Tree Genet Genomes*. doi:10.1007/s11295-015-0917-3
- Cullis BR, Gleeson AC (1991) Spatial analysis of field experiments—an extension to two dimensions. *Biometrics* 47:1449–1460
- Cullis BR, Gogel B, Verbyla AP, Thompson R (1998) Spatial analysis of multi-environment early generation variety trials. *Biometrics* 54:1–18
- Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. *J Agric Biol Environ Stat* 11:381–393
- Currie ID, Durbán M (2002) Flexible smoothing with p-splines: a unified approach. *Stat Model* 2:333–349
- Currie ID, Durbán M, Eilers PHC (2006) Generalized linear array models with applications to multidimensional smoothing. *J R Statist Soc Ser B* 68:259–280
- de la Mata R, Zas R (2010) Transferring Atlantic maritime pine improved material to a region with marked Mediterranean influence in inland NW Spain: a likelihood-based approach on spatially adjusted field data. *Eur J Forest Res* 129:645–658
- Durbán M, Currie ID, Kempton R (2001) Adjusting for fertility and competition in variety trials. *J Agric Sci (Camb.)* 136:129–140
- Durbán M, Hackett CA, McNicol JW, Newton AC, Thomas WTB, Currie ID (2003) The practical use of semiparametric models in field trials. *J Agric Biol Environ Stat* 8:48–66
- Dutkowski GW, Costa e Silva J, Gilmour AR, Lopez GA (2002) Spatial analysis methods for forest genetic trials. *Can J For Res* 32:2201–2214
- Dutkowski GW, Costa e Silva J, Gilmour AR, Wallendorf H, Aguiar A (2006) Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. *Can J For Res* 36:1851–1870
- Edmondson RN (1993) Systematic row-and-column designs balanced for low order polynomial interactions between rows and columns. *J R Statist Soc Ser B* 55:707–723
- Eilers PHC (1999) Discussion on: the analysis of designed experiments and longitudinal data by using smoothing splines (by Verbyla et al.). *J R Statist Soc Ser C* 48:307–308
- Eilers PHC, Marx BD (1996) Flexible smoothing with B-splines and penalties. *Stat Sci* 11:89–102
- Eilers PHC, Marx BD (2003) Multivariate calibration with temperature interaction using two-dimensional penalized signal regression. *Chemometr Intell Lab Syst* 66:159–174
- Eilers PHC, Marx BD, Durbán M (2015) Twenty years of P-splines. *SORT* 39(2):149–186
- Federer WT (1998) Recovery of interblock, intergradient, and inter-variety information in incomplete block and lattice rectangle design experiments. *Biometrics* 54:471–481
- Fu YB, Yanchuk AD, Namkoong G (1999) Spatial patterns of tree height variations in a series of Douglas-fir progeny trials: implications for genetic testing. *Can J For Res* 29:714–723
- Gilmour AR (2000) Post blocking gone too far! Recovery of information and spatial analysis in field experiments. *Biometrics* 56(3):944–945
- Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiments. *J Agric Biol Environ Stat* 2:269–293
- John JA, Williams ER (1995) *Cyclic and computer generated designs*. 2nd ed. Chapman & Hall, London
- Jordan DR, Mace ES, Cruickshank AW, Hunt CH, Henzell RG (2011) Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. *Crop Sci* 51:1444–1457
- Lado B, Matus I, Rodriguez A, Inostroza L, Poland J et al (2013) Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3* 3:2105–2114
- Lee D-J, Durbán M, Eilers PHC (2013) Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested bases. *Comput Stat Data Anal* 61:22–37
- Leiser WL, Rattunde HF, Piepho HP, Parzies HK (2012) Getting the most out of sorghum low-input field trials in West Africa using spatial adjustment. *J Agron Crop Sci* 198:349–359

- Liu SM, Constable GA, Cullis BR, Stiller WN, Reid PE (2015) Benefit of spatial analysis for furrow irrigated cotton breeding trials. *Euphytica* 201:253–264. doi:10.1007/s10681-014-1205-2
- Loo-Dinkins JA, Tauer CG, Lambeth CC (1990) Selection system efficiencies for computer simulated progeny test field designs in loblolly pine. *Theor Appl Genet* 79:89–96
- Mace ES, Hunt CH, Jordan DR (2013) Supermodels: sorghum and maize provide mutual insight into the genetics of flowering time. *Theor Appl Genet* 126:1377–1395
- Magnussen S (1993) Bias in genetic variance estimates due to spatial autocorrelation. *Theor Appl Genet* 86:349–355
- Magnussen S (1994) A method to adjust simultaneously for spatial microsite and competition effects. *Can J For Res* 24:985–995
- Müller BU, Kleinknecht K, Möhring J, Piepho HP (2010) Comparison of spatial models for sugar beet and barley trials. *Crop Sci* 50:794–802
- Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H (2006) Joint modelling of additive and non-additive genetic line effects in single field trials. *Theor Appl Genet* 113:809–819
- Patterson HD, Thompson R (1971) Recovery of interblock information when block sizes are unequal. *Biometrika* 31:100–109
- Patterson HD, Williams ER, Hunter EA (1978) Block designs for variety trials. *J Agric Sci (Camb)* 90:395–400
- Payne RW (2006) New and traditional methods for the analysis of unreplicated experiments. *Crop Sci* 46:2476–2481
- Piepho HP, Williams ER (2010) Linear variance models for plant breeding trials. *Plant Breed* 129:1–8
- Piepho HP, Richter C, Williams ER (2008) Nearest neighbour adjustment and linear variance models in plant breeding trials. *Biom J* 50(2):164–189
- Piepho HP, Möhring J, Pflugfelder M, Hermann W, Williams ER (2015) Problems in parameter estimation for power and AR(1) models of spatial correlation in designed field experiments. *Commun Biometr Crop Sci* 10(1):3–16
- Qiao CG, Basford KE, Delacy IH, Cooper M (2000) Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theor Appl Genet* 100:9–16
- Qiao CG, Basford KE, DeLacy IH, Cooper M (2004) Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. *Theor Appl Genet* 108:1256–1264
- R Development Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org>
- Rodríguez-Álvarez MX, Boer MP, van Eeuwijk FA, Eilers PHC (2016a) Spatial models for field trials. arXiv:1607.08255v1 [stat.ME]
- Rodríguez-Álvarez MX, Boer MP, Eilers PHC, van Eeuwijk FA (2016b) SpATS: spatial analysis of field trials with splines. R package version 1.0–4. <https://cran.r-project.org/package=SpATS>
- Ruppert D, Wand M, Carroll R (2003) Semiparametric regression. Cambridge University Press, Cambridge
- Sarker A, Singh M (2015) Improving breeding efficiency through application of appropriate experimental designs and analysis models: A case of lentil (*Lens culinaris* Medikus subsp. *culinaris*) yield trials. *Field Crops Res* 179:26–34
- Smith A, Cullis B, Thompson R (2001) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57:1138–1147
- Smith A, Cullis B, Luckett D, Hollamby G, Thompson R (2002) Exploring variety–environment data using random effects AMMI models with adjustments for spatial field trend: Part 2: Applications. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding. CABI, Wallingford, pp 337–351
- Stefanova KT, Smith AB, Cullis BR (2009) Enhanced diagnostics for the spatial analysis of field trials. *J Agric Biol Environ Stat* 14:392–410
- Stroup WW, Baenziger PS, Mulitze DK (1994) Removing spatial variation from wheat yield trials: a comparison of methods. *Crop Sci* 34:62–66
- Taye G, Njuho PM (2008) Smoothing fertility trends in agricultural field experiments. *Statistics* 42(3):275–289
- Verbyla AP, Cullis BR, Kenward MG, Welham SJ (1999) The analysis of designed experiments and longitudinal data using smoothing splines (with discussion). *J R Stat Soc Ser C* 48:269–312
- Welham SJ, Gogel BJ, Smith AB, Thompson R, Cullis BR (2010) A comparison of analysis methods for late-stage variety evaluation trials. *Aust N Z J Stat* 52:125–149
- Wilkinson GN, Eckert SR, Hancock TW, Mayo O (1983) Nearest neighbour (NN) analysis of field experiments. *J R Stat Soc Ser B* 45:151–211
- Williams ER, John JA, Whitaker D (2006) Construction of resolvable row-column designs. *Biometrics* 62:103–108
- Williams ER, John JA, Whitaker D (2014) Construction of more flexible and efficient p-rep designs. *Aust N Z J Stat* 56:89–96
- Wood SN (2006) Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics* 62(4):1025–1036
- Yang RC, Ye TZ, Blade SF, Bandara M (2004) Efficiency of spatial analysis of field pea variety trials. *Crop Sci* 44:49–55
- Yates F (1940) The recovery of inter-block information in balanced incomplete block designs. *Annals of Eugenics* 10:317–325
- Zas R (2006) Iterative kriging for removing spatial autocorrelation in analysis of forest genetic trials. *Tree Genet Genomes* 2:177–186
- Zimmerman DL, Harville DA (1991) A random field approach to the analysis of field-plot experiments and other spatial experiments. *Biometrics* 47:223–239