MISCELLANEOUS

Getting the Most Out of Sorghum Low-Input Field Trials in West Africa Using Spatial Adjustment

W. L. Leiser[1,2], H. F. Rattunde[2], H.-P. Piepho[3] & H. K. Parzies[1]

1 Department of Applied Genetics and Plant Breeding, University of Hohenheim Stuttgart, Germany
2 International Crops Research Institute for the Semi-Arid Tropics Bamako, Mali
3 Department of Bioinformatics, University of Hohenheim Stuttgart, Germany

Abstract

Breeding sorghum for low-input conditions is hindered by soil heterogeneity. Spatial adjustment using mixed models can help account for this variation and increase precision of low-input field trials. Large small-scale spatial variation (CV 39.4 %) for plant available phosphorus was mapped in an intensely sampled low-input field. Spatial adjustments were shown to account for residual yield differences because of this and other growth factors. To investigate the potential of such models to increase the efficiency of low- and high-input field trials, 17 experiments with 70 sorghum genotypes conducted in Mali, West Africa, were analysed for grain yield using different mixed models including models with autoregressive spatial correlation terms. Spatial models (AR1, AR2) improved broad sense heritability estimates for grain yield, averaging gains of 10 and 6 % points relative to randomized complete block (RCB) and lattice models, respectively. The heritability estimate gains were even higher under low phosphorus conditions and in two-replicate analyses. No specific model was best for all environments. A single spatial model, AR1 × AR1, captured most of the gains for heritability and relative efficiency provided by the best model identified for each environment using Akaike's Information Criterion. Spatial modelling resulted in important changes in genotype ranking for grain yield. Thus, the use of spatial models was shown to have potentially important consequences for aiding effective sorghum selection in West Africa, particularly under low-input conditions and for trials with fewer replications. Thus, using spatial models can improve the resource allocation of a breeding program. Furthermore, our results show that good experimental design with optimal placement and orientation of blocks is essential for efficient statistical analysis with or without spatial adjustment.

Introduction

Sorghum (*Sorghum bicolor* L. *Moench*) is the world's fifth and Africa's second most widely grown cereal crop (FAO 2010). Farmers in the Savannah zone of West Africa depend on sorghum as a staple crop that is grown primarily under low-input conditions. Although sorghum can achieve grain yields of several tons ha^{-1} in West Africa, average yields are only about 1 t ha^{-1} (FAO 2010), due in part to low soil fertility and low-input production systems (Vom Brocke et al. 2010). Soil-phosphorus deficiency is a major factor reducing sorghum and pearl millet growth and productivity across the range of rainfall zones in West Africa (Buerkert et al. 2001). However, little selection work specifically targeting low P and low-input production conditions has been carried out due, in part, to problems of high environmental error levels encountered when selecting in low productivity conditions and possibly lower heritability values (Ceccarelli 1994). Genotypic selection aims at selecting the best genotypes based on the genetic differences. Spatial variation in fertility, moisture and other environmental factors can bias variation and selection and

increase residual variation (Grondona et al. 1996). Therefore, environmental effects need to be controlled by design and analysis for effective selection. Many different field designs and analysis methods have been developed in the last century (Edmondson 2005). Environmental effects can be accounted for by using control plots, replications and blocks. These techniques have limitations where spatial variation cannot be well captured by blocks. More advanced spatial adjustments may improve the analysis in such cases. Various spatial adjustment techniques have been developed (e.g. autoregressive models) and have been shown to significantly increase heritability and thus make selection more efficient especially in abiotic stress environments (Gilmour et al. 1997, Singh et al. 2003). As part of a 5-year multi-location experiment for phosphorus efficiency selection in sorghum, field trials were analysed with spatial models. The objectives of this study were (i) to detect soil small-scale heterogeneity for plant available phosphorus and assess its relationship with grain yield and spatial adjustments for sorghum performance in one low P trial, (ii) to detect types of spatial models that are superior, (iii) to estimate the impact of spatial models on heritability and relative efficiency (RE) estimates for grain yield in a wide range of trials differing for levels of fertilization and replication and (iv) to evaluate the impact of spatial models on genotype ranking and genotypic selection.

Material and Methods

Experiments

A set of 70 sorghum varieties considered to be well adapted to the Sudanian zone of West Africa was established for this study. The varieties were sampled from sorghum breeding programs of the Malian Institut d'Economie Rural (IER) and the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Four very late maturing varieties proved to be unadapted, frequently suffering severe yield reductions as a result of sorghum midge (*Stenodiplosis sorghicola*) attacks and were therefore excluded from the analyses. The remaining varieties generally exhibited phenotypic uniformity and were considered to be homozygous lines. Yield trials were conducted in Mali, West Africa, at the IER-Kolombada (12° 40′N, 7°0′W) and ICRISAT-Samanko (12°31′N, 8°4′W) station in the years 2006–2010 under rain-fed conditions. Two separate trials were conducted, one with phosphorous fertilization (denoted 'high P') and one without (denoted 'low P') in each location and year. The high P and low P trials were conducted in adjacent fields at Samanko (Sko), whereas they were located in a single field divided into high P and low P sections at Kolombada (Kba). The high P fields were fertilized with diammonium phosphate at

rates of 100–200 kg ha^{-1} as basal fertilizer and urea (50 kg ha^{-1}) as top dressing. The low P fields were fertilized only with urea at rates that gave equivalent units of N as received by the high P fields. Each trial consisted of 70 genotypes sown in an α-design with four complete replicates and block sizes of five plots. Field layouts differed in environments, with column numbers ranging from 15 to 35 and row numbers ranging from 8 to 20. Plots consisted of two rows three meter in length and had a 0.75 m distance between rows, with 30 cm between hills within the row. Hills were thinned to two plants, resulting in a total of 44 plants per plot. A single guard row separated each test plot to minimize neighbour effects. The guard rows were left unsown in 2006–2008, whereas they were sown with a common locally adapted variety in 2009 and 2010. Soil from the entire low P field in Samanko 2009 was sampled in a 2.2 m × 3.6 m grid and grain yields of the corresponding guard rows were measured. Soil samples were analysed for plant available phosphorus using the Bray-1 method (Bray and Kurtz 1945).

Analysis

Description of models

Data for grain yield were analysed in each environment with 91 different models including 82 models with autoregressive spatial correlation terms. Column and row factors were assigned to spatially reference each trial plot's position. A conservative approach to spatial modelling is to form a baseline model with no spatial trend, add a spatial error term to the model and check whether the extra spatial term can improve the fit of the model (Williams et al. 2006, Piepho et al. 2008). The baseline model accounts for some global trend, for example, by terms for replicates, whereas the spatial error term models the local trend. A randomized complete block (RCB) model was used as the baseline model for each trial with fixed genotype and replication effects and can be expressed as:

$$Y_{ij} = \mu + \alpha_j + \gamma_i + e_{ij},$$

where μ is the general mean, α_j the effect of j-th replicate, γ_i the effect of i-th genotype, and e_{ij} is the error term of Y_{ij}. Although our trials were randomized as α-designs and thus would normally be analysed by a model with fixed effects for replicates and random or fixed effects for incomplete blocks, we have selected a RCB model as baseline model because many scientists in West Africa are still using RCBD as standard field design. While a lattice model is typically more efficient, analysis by an RCB model can be justified from randomization theory (Speed et al. 1985). A comparison of spatial and lattice models can always be obtained via

the comparison of both of these models with the RCB model.

Nine of the 91 models, inclusive the RCB model, did not have a specific autoregressive spatial correlation term. Instead, these nine models had one of the following effects added compared to the baseline model: (i) incomplete random blocks (referred to as 'lattice' in Tables 2–4), (ii) incomplete random blocks and random row term, (iii) incomplete random blocks and random column term, (iv) incomplete random blocks and random row and column term, (v) linear trend across rows, (vi) linear trend across rows with nugget effect, (vii) linear trend across columns and (viii) linear trend across columns with nugget effect.

For all other models, an extra term τ_{ij} was added to the baseline model that accounted for the local spatial trend (autoregressive spatial correlation). The term τ_{ij} assumes that neighbouring plots have more similar environmental influences than plots that lie further apart. In addition to the local spatial trend, all spatial models were analysed with and without a nugget effect ε_{ij} to account for the extra error of each observational unit (e.g. plot) in a trial. The model with nugget and local spatial trend can be stated as:

$$Y_{ij} = \mu + \alpha_j + \gamma_i + \tau_{ij} + \eta_{ij}$$

Global linear trends across columns or rows were accounted for by fitting a linear trend across all columns or rows by using the column or row number as fixed regression variate. By fitting an extra random factor for rows or columns, further trends along columns or rows were modelled. For modelling the influence of neighbouring plots along both rows and columns, autoregressive (AR) models of order 1 (AR1) and 2 (AR2) were applied (Cullis and Gleeson 1991). Those models assume a decreasing covariance of neighbouring plots with increasing distance (Schabenberger and Pierce 2001). Even though there are different models for modelling spatial trends (Grondona et al. 1996, Schabenberger and Pierce 2001, Piepho et al. 2008), we only focused on AR1 and AR2 models, because those models are easily fitted in GENSTAT 12.1 and have shown superiority over Gaussian and Spherical models in other studies (Müller et al. 2010). Furthermore, adding these models would have increased the number of tested models drastically. Both one-dimensional and two-dimensional AR1 models and two-dimensional AR2 models were used to model the spatial influence of neighbouring plots along columns and rows. While AR1 models consider only the correlation of adjacent plots, AR2 models consider additionally the correlation between plots being two plots apart. The formula for the correlation between plots under the AR2 model can be stated as:

$$C_{ii} = 1,$$

$$C_{i+1,i} = \varphi_1/(1 - \varphi_2),$$

$$C_{ij} = \varphi_1 C_{i-1,j} + \varphi_2 C_{i-2,j}, i > j + 1,$$

$$|\varphi_1| < (1 - \varphi_2), |\varphi_2| < 1,$$

while $\varphi_2 = 0$ for AR1, where C_{ij} is the correlation between plots and φ the correlation term (Gilmour et al. 2009). Table 1 shows the used autoregressive models, when autocorrelation between neighbouring plots was modelled across the whole field. For ten models, columns and rows were considered as nested within replicate (models not shown). To each model stated in Table 1, a linear trend across columns (LC) and across rows (LR) was added to test for global trends in addition to the local trend. Each trial was analysed using data from all four replicates and from only the first two replicates, simulating a two-replicate trial.

Model selection

Akaike's Information Criterion (AIC) was used to select the best fitting model of each trial (referred to as 'best spatial model'). This approach has been shown to be efficient for spatial model selection (Kehel et al. 2010). AIC is defined as: $AIC = 2k - 2\log(L)$, where k is the number of used variance parameters and L refers to the REML likelihood (Wolfinger 1996). The smaller the AIC value, the better the model fit. All models of Table 1 had the same fixed effects (replication and genotype) and were compared

Table 1 Used autoregressive spatial models as random effects in one and two dimensions across whole experimental field

Spatial models	
One-dimensional	Two-dimensional
Row AR1	AR1 × AR1
Row AR1 with nugget-effect	AR1 × AR1 with '*units*'
Row AR1 + random row	AR1 × AR1 + random row
Row AR1 + random row with '*units*'	AR1 × AR1 + random row with '*units*'
Row AR1 + random column	AR1 × AR1 + random column
Row AR1 + random column with '*units*'	AR1 × AR1 + random column with '*units*'
Column AR1	AR2 × AR2
Column AR1 with '*units*'	AR2 × AR2 with '*units*'
Column AR1 + random row	AR2 × AR2 + random row
Column AR1 + random row with '*units*'	AR2 × AR2 + random row with '*units*'
Column AR1 + random column	AR2 × AR2 + random column
Column AR1 + random column with '*units*'	AR2 × AR2 + random column with '*units*'

AR1, autoregressive order 1; AR2, autoregressive order 2; '*units*, nugget effect.

by AIC in each environment (1. step). Model selection for fixed effects was not conducted because the statistical package GENSTAT 12.1 uses REML and therefore model selection based on AIC could be only carried out regarding the random effects. Further, the best model based on the AIC was tested for any significant global trends (LC, LR) (2. step). If a significant global trend existed, this model inclusive of the global trend was selected as best spatial model (see Table 2). In most of the environments, the best model of the first step selected for its covariance structure was also selected in the second step. Only in a few environments, did adding the significant global trend change the superiority of the previously selected covariance structure thus a different covariance structure model was selected compared to step one. Finally, a variogram of the selected best spatial model was evaluated to see whether it fit the expected variogram of the selected model.

Precision assessment

The relative precision and efficiency obtained with contrasting models were assessed for individual trials by computing RE and broad sense heritability (h^2) estimated with the respective adjusted data. For RE estimates, genotypes were considered as fixed effects, while for heritability estimates, genotypes were considered random. This approach was necessary (Möhring and Piepho 2009), because the adjusted means of genotypes from each single environment are to be used in further analysis in a weighted combined

two-stage analysis with genotypes as random (results not shown). RE was calculated based on SED as suggested by Qiao et al. (2000) and can be described as:

$$\text{RE in } \% = \frac{\text{SED}_{\text{RCB}}}{\text{SED}_{\text{model}}} * 100,$$

where SED is the REML-based average standard error of the mean differences between genotypes for the baseline model (SED_{RCB}) and the different alternative models ($\text{SED}_{\text{model}}$). The higher the RE estimate, the more efficient the field evaluation of genotypes.

The broad sense heritability within one environment was calculated with an adjusted formula for unbalanced experiments based on Piepho and Möhring (2007)

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\text{VD}}{2}},$$

where σ_g^2 is the genotypic variance component and VD the average variance of differences between genotype means obtained from the default procedure in GENSTAT 12.1.

Results

Pattern of environmental variability and spatial models

Considering the Samanko 2009 low P field as an example, the plant available P (Bray-1 P) soil content showed

Table 2 Spatial correlated error terms of selected best spatial models for eight high P and nine low P environments

| | Model | | | | |
| | Four replications | | Two replications | | |
Environment	Global trend	Local trend	Global trend	Local trend
Kba06L	LC	AR1 × AR1	LC	AR1 × AR1
Kba07L	LC	AR2 × AR2	LC	AR1 × AR1
Kba08L	LC	AR1 × AR1 + random row	LC	AR1 × AR1
Kba09L		Rep. column AR1 for column		AR1 × AR1
Sko06L		AR1 × AR1 + '*units*'		AR1 × AR1 + random column
Sko07L	LC	AR1 × AR1 + '*units*'		Random column
Sko08L	LR	Column AR1		Column AR1
Sko09L	LC	AR1 × AR1	LR	AR1 × AR1 + '*units*'
Sko10L	LR	AR1 × AR1 + random column	LR	AR1 × AR1
Kba06H	LC	AR1 × AR1	LC	AR1 × AR1
Kba07H	LC	Column AR1 + random column	LC	Column AR1 + random column
Kba08H	LC	Column AR1	LC	AR1 × AR1
Kba09H	LR	AR2 × AR2		AR1 × AR1
Sko06H		Rep. block + column		Random column
Sko08H	LR	AR2 × AR2		Column AR1
Sko09H		AR2 × AR2 + random row		AR2 × AR2 + random row
Sko10H		Row AR1 + random row		AR2 × AR2 + random row

Kba, Kolombada; Sko, Samanko; 06,07,08,09,10, years 2006–2010; L, low P; H, high P; LC, linear trend across columns, LR, linear trend across rows; AR1, autoregressive order 1; AR2, autoregressive order 2; Rep. column, column nested within replication as random term; '*units*', nugget effect.

considerable spatial variation in a small-scale grid (Fig. 1), with a coefficient of variation for Bray-1 P of 39.4 %. Although there is visually some correspondence between the blocking structure of replicates or incomplete blocks (5 plots per block) with the spatial variation for measured Bray-1 P values, it is certainly difficult to capture all variation with incomplete blocks (Fig. 1). In addition, Bray-1P plot values were used as covariate in a model accounting only for fixed genotype effects, therefore eliminating genotypic differences, Bray-1P was significant by a Wald-test (P < 0.001), as is visually confirmed by comparing Bray-1P and plot residuals in Figure 1. In contrast, Bray-1P used as covariate in the best spatial model did not yield a significant Wald-test (P = 0.63). A subset of 55 soil samples from the total sample set was analysed for P_{total}, pH, Mg^{2+}, Na^+, C_{org}, Al-Saturation, Bray-1P, Ca^{2+} and K^+. The soil parameters were used in a forward step-wise regression to explain the genotype-model residuals of these 55 plots. A linear model stated as: Residual = constant + P_{total} + pH + Mg^{2+} + Bray-1 P + Na^+ was selected. It explained 49 % of the variation, while Bray-1P showed a significant F-statistics and third largest sum of squares (data not shown). Using

the best spatial model residuals as response variable resulted in only 20 % explained variation by the model and Bray-1 P no longer showed a significant Wald-statistics (data not shown).

Best spatial model composition

The best spatial models as identified by AIC differed considerably over environments (Table 2). A significant linear trend either across rows (LR) or columns (LC) occurred in 70 % of the four-replicate and 47 % of the two-replicate environments analyses. The two-dimensional (AR1 × AR1; AR2 × AR2) models were higher ranked based on AIC than one-dimensional models in 64 % of four-replicate and 70 % of two-replicate environments analyses. Adding the nugget effect ('*units*') improved the model fit in only a few cases and often resulted in convergence problems. To exemplify our approach to graphically check spatial models, Figure 2 shows as example residual variograms for one environment of the RCB model, the AR1 × AR1 model and the LC + AR1 × AR1 model. It shows that the fit to an exponential curve, corresponding to an AR-model, is
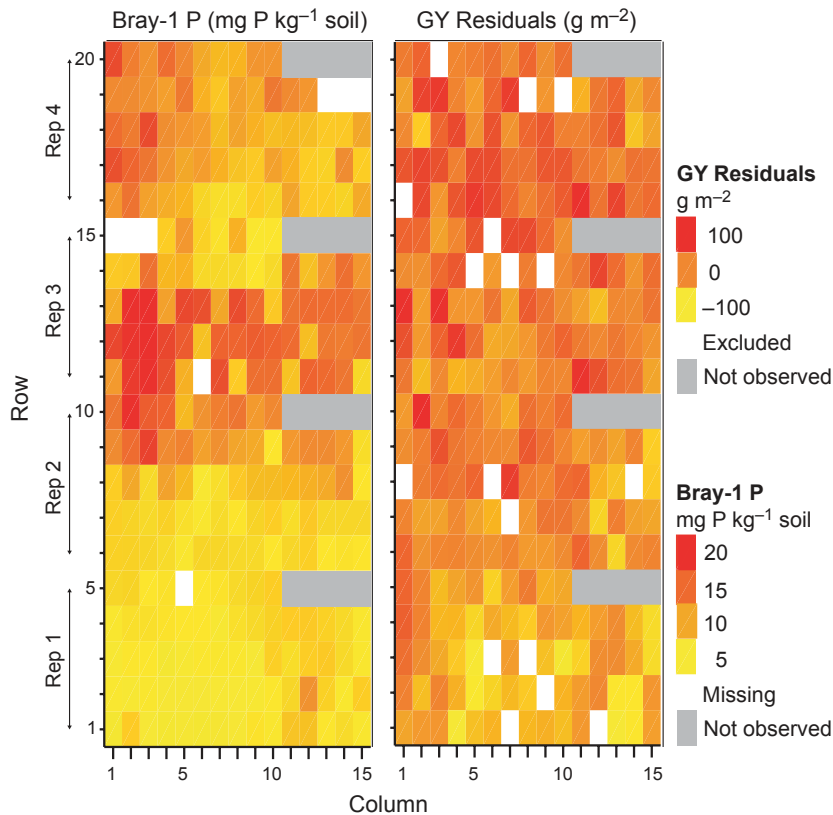

Fig. 1 Heat map of plant available phosphorus (Bray-1P) in mg P kg^{-1} soil and residuals of a model accounting only for genotype effects for grain yield (gm^{-2}) of one unfertilized sorghum field (Sko09L). Field (2408 m^2) consisted of 20 rows (3.5 m/row) and 16 columns (2.15 m/column). Each square represents one sample taken in one plot. Replications indicate usually applied field trial design.

improved by the spatial models and by including the global trend (LC) the fit could be improved even more.

Relative efficiency and broad-sense heritability

Broad-sense heritability (h^2) estimates for grain yield within individual environments, also referred to as repeatability, were significantly raised by using the best spatial models (based on AIC, Table 2) compared to RCB (Table 3). The highest observed h^2 increase relative to the RCB model was 49 % points while the lowest was 1.5 % points (data not shown). Mean h^2 increase of the best spatial model compared to the lattice model was only significant in the two-replicate analyses. The h^2 differences between lattice and best spatial model were rather small (1–10 % points) in the four-replicate analyses whereas in the two-replicate analyses a maximum h^2 increase of 33 % points was observed (data not shown). Using an AR1 × AR1 model with any significant linear trend (e.g. LC, LR) resulted in h^2 values that were generally close to results obtained by the best spatial model. However, in some environments, those models were inferior to lattice models if either row or column terms showed no spatial correlation. Heritability estimates were generally lower in low P environments, although lattice, best spatial and AR1 × AR1 models were more effective in increasing h^2 relative to the high P environments (Table 3). Similarly, although analyses of only two replicates resulted in lower heritabilities, the best spatial model resulted in major increases in h^2, especially for the low P environments and when compared to lattice.

Estimates of the Relative Efficiencies (RE) of best spatial models (Table 4) followed the same pattern as observed for heritability estimates (Table 3). The lattice model increased significantly the efficiency relative to RCB in the four-replicate but not in the two-replicate analyses across all environments. Also, the best spatial models increased significantly the efficiencies relative to RCB and lattice models when averaged across all environments, and particularly across low P environments and analyses based on only two replications. The standard error of genotype differences (SED) of the RCB model was significantly larger than those for the best spatial model, averaging 27 % and 37 % larger over all low P environments with four- and two-replicate data, respectively.

Consequences of spatial adjustment for selection decisions

Estimated entry mean grain yields obtained with RCB, lattice and best spatial model were positively correlated in each environment, ranging from 0.59 to 0.99, while lattice and best spatial model had generally a stronger correlation (r = 0.76–0.99). Correlations were weaker in environments where spatial modelling had a strong impact on h^2 and efficiency (data not shown).

Plotting spatially adjusted (best spatial model) estimated genotype grain yields of entries against the corresponding RCB estimates enables visualization of the absolute consequences of adjustments, with genotypes below the 1 : 1 line showing larger estimates based on the RCB model and those above showing smaller ones (Fig. 3). The differences between RCB and spatially adjusted values were quite small for the high P environments for both the four- and two-replicate data. In contrast, there were considerable differences in the low P environments. In some environments, certain genotypes appear to have been overestimated by 100 % with the RCB model (Fig. 3b,d). Based on Figure 3, it can be shown that genotype ranking for grain yield in each environment can be drastically different if data are analysed as RCB or best spatial model.
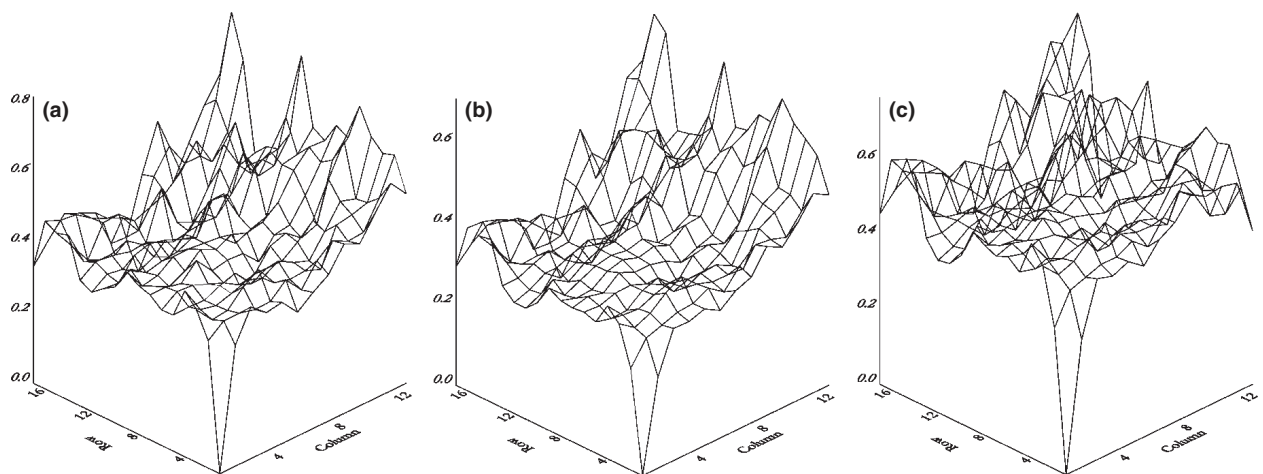

Fig. 2 Variograms of RCB model (a), AR1 × AR1 (b) and LC + AR1 × AR1 (c) model of environment Sko09L.

Table 3 Broad sense heritability (%) averaged over 17 environments (eight high P and nine low P) analysed with four and two replications as randomized complete block design (RCB), lattice, AR1 × AR1 with any significant global trend and best spatial model

	Nr Rep	RCB	Lattice	AR1 × AR1	Best spatial	LSD 5 %
Min	4	34.2	44.3	60.3	61.3	
Max	4	84.3	85.3	86.8	87.2	
Mean	4	68.0	72.1	77.6	78.2	7.7
High P Mean	4	74.0	77.5	80.6	81.4	8.3
Low P Mean	4	62.6	67.3	74.9	75.3	12.8
Min	2	17.0	30.8	53.0	53.0	
Max	2	78.3	77.5	78.4	88.4	
Mean	2	56.1	60.3	68.9	70.8	8.2
High P Mean	2	60.8	65.5	72.0	73.9	9.1
Low P Mean	2	51.9	55.6	66.2	68.0	13.5

LSD5 % values for within row comparison.

Table 4 Relative efficiency (%) averaged over 17 environments (eight high P and nine low P) analysed with four and two replications as randomized complete block design (RCB), lattice, AR1 × AR1 with any significant global trend and best spatial model

	Nr. Rep	RCB	Lattice	AR1 × AR1	Best spatial	LSD 5 %
Min	4	100	100.0	104.9	105.6	
Max	4	100	139.3	180.8	184.4	
Mean	4	100	109.7	122.5	123.8	8.2
High P Mean	4	100	108.4	117.2	119.3	10.7
Low P Mean	4	100	110.9	127.2	127.9	19.9
Min	2	100	100.0	102.3	103.3	
Max	2	100	137.5	195.6	195.6	
Mean	2	100	109.6	127.5	131.8	13.5
High P Mean	2	100	108.7	121.9	125.8	11.9
Low P Mean	2	100	110.4	132.4	137.1	23.6

LSD5 % values for within row comparison.

The actual consequences on selection decisions by using spatial adjustment were examined in two environments contrasting for the efficiency gains provided by spatial adjustments as indicated by heritability increases. Spatial adjustment in the first environment, Kba06L, had a strong impact on h^2, with an increase of 38 % points (4-reps) and 49 % points (2-reps), and relatively low correlation between RCB and spatially adjusted genotype mean yields (r = 0.71 (4-reps), r = 0.59 (2-reps)). The second environment, Sko06H, exhibited only a slight increase in h^2 with spatial adjustment (1.2 % points (4-reps) and 1.5 % points (2-reps)) and close correlation between the RCB and spatially adjusted genotype yield means (r = 0.99). Ranking genotypes in Kba06L using RCB entry means showed that the ten top-yielding genotypes (15 % selection intensity) with this model agreed only with 60 % and 40 % of the best-ranked genotypes using the lattice and best spatial models, respectively. The lattice model top ten genotypes overlapped with 60 % of the best-ranked genotypes analysed with the best spatial model. Further, using the genotype ranks of four-replicate adjusted mean yields as the standard, the two-replicate analyses with the best spatial, lattice and RCB models would have missed 20 %, 50 % and 50 % of the highest ranked genotypes, respectively. In contrast, genotype ranking in Sko06H identified the same genotypes using the RCB, lattice and best spatial models with four-replicate data and all but three genotypes using two-replicate data.

Discussion

Analysis of spatial variation of plant nutrients recently gained more importance in industrialized countries because of precision farming (Raun et al. 1997). In West Africa, high spatial variability for nutrients and plant growth has been reported by several authors (Brouwer et al. 1993, Lamers and Feil 1995, Voortman et al. 2004). Our sampled low P field showed a very high variation for Bray-1P (CV > 36 %) (Wilding 1985) and agrees with observed high small-scale variation for P in soils of Niger and USA (Raun et al. 1997, Voortman 2010). Although Bray-1P could explain some phenotypic variation, as seen in the regression, several soil factors influenced grain yield. Soil water availability was visually observed to be an important factor but was not measured. The findings based on the step-wise regression for explaining residual variation showed that the applied best spatial model captured the Bray-1 P variation together with other environmental factors and is therefore suitable for genotype analysis. Intensive soil sampling accounting for the variability in a field is mostly not affordable and leads often only to a limited proportion of explained variation by the soil properties (Voortman and Brouwer 2003). Thus, spatial adjustment methods accounting for any environmental and soil variation are a suitable inexpensive way of correcting for microvariability and should be considered especially in trials conducted in highly weathered soils with high spatial variation. The high heterogeneity of soils in West Africa is owing to many factors such as abiotic (e.g. wind, water), biotic (e.g. termites, trees) and human (e.g. farm practices) influences (Brouwer et al. 1993) and are therefore rather difficult to be well recorded and reduced to a limited amount of influencing factors. Bänziger and Cooper (2001) suggested the use of spatial models to increase efficiency of on-station and on-farm trials for low-input breeding.

Out of the 91 applied models, in 16 of 17 environments, spatial AR models were superior to lattice and RCB analysis in terms of AIC. Cullis and Gleeson (1989) and Müller et al. (2010) showed that one-dimensional AR1 models

Table 5 Entry numbers of top ten ranked genotypes in an environment with a high heritability (h^2) increase and in an environment with a low h^2 increase through spatial adjustments. Ranks are based on the adjusted means for grain yield (g m^{-2}) analysed with four replications and two replications using a randomized complete block (RCB), lattice and best spatial models

	High h^2 increase (Kba06L)						Low h^2 increase (Sko06H)					
	RCB		Lattice		Best spatial		RCB		Lattice		Best spatial	
Rank	4 Reps	2 Reps	4 Reps	2 Reps	4 Reps	2 Reps	4 Reps	2 Reps	4 Reps	2 Reps	4 Reps	2 Reps
1	55	55	34	34	34	34	6	60	6	60	60	60
2	47	26	38	14	47	9	60	58	60	58	6	64
3	38	14	14	55	49	67	65	64	65	64	65	58
4	36	70	49	70	41	49	42	42	58	42	58	6
5	41	36	13	36	38	14	58	6	42	6	42	42
6	30	42	64	67	14	4	64	65	64	65	64	65
7	49	34	2	38	67	41	45	13	14	13	14	67
8	26	67	55	49	4	21	14	67	45	67	45	13
9	60	2	41	9	13	3	50	51	50	51	50	51
10	2	33	60	26	9	47	67	38	67	38	67	38

were superior compared to RCB and lattice analysis. Superiority of two-dimensional AR1 and ARIMA (autoregressive integrated moving average) models have been reported in wheat and barley yield trials by several authors (Cullis and Gleeson 1991, Grondona et al. 1996, Gilmour et al. 1997, Qiao et al. 2000, Smith et al. 2001, Singh et al. 2003). In our data, two-dimensional AR (1&2) models were superior to one-dimensional AR models in 64 % (4 reps) and 70 % (2 reps) of the environments, respectively. In 70 % (4 reps) and 47 % (2 reps) of the environments significant global linear trends across rows or columns existed. If a linear trend was detected, a spatial model including this linear trend was selected. Linear trends can be also included in lattice models. In many of our environments, lattice models could be improved by adding a global linear trend across columns (data not shown). The two-replicate analyses showed less global trends as the field size decreased and thus field trends, for example, linear column trends, were less dominant. Fewer global trends were also reported by Qiao et al. (2000) in shorter fields compared to longer fields. The higher frequency of significant column trends corresponded with visually observable patterns of field heterogeneity in these trials and is most likely related to the fact that many field operations (like ridging, fertilizer topdressing) are performed down the length of the field, and not across the ridges. The lattice design poorly accounted for this spatial trend in our experiments, because the orientation of blocks was not optimal, and the layout (orientation) of incomplete blocks of large lattice design experiments was typically across the ridges to facilitate observations by walking down the alley between bands rather than walking through plots down the field in 'column' direction. An appropriate design layout with incomplete blocks along the ridges probably would have captured more field variation and thus the efficiency and h^2 differences between lattice and spatial models would have been much less, because blocks would then have better captured spatial field trends. For future experiment designs with one-dimensional blocking structure, an appropriate orientation of blocks is very important to cover most of the field variation. In case strong trends are expected on both, rows and columns, an efficient row–column design (John and Williams 1995) is especially useful for covering two-dimensional spatial variation. In our trials, mostly two-dimensional spatial models were selected, thus row–column designs would have been an appropriate experimental design (Piepho and Williams 2010). An additional nugget effect mostly did not improve the fit of a model as also stated by Qiao et al. (2000). This could be due to predominantly existing convergence problems in AR models with nugget effects and hence an inappropriate selection process. Müller et al. (2010) experienced the same convergence problems and therefore excluded nonlinear models with nugget effects. With linear variance models, these problems are much less prominent (Piepho et al. 2008, Piepho and Williams 2010). We included nonlinear models with nugget effect because a spatially non-correlated extra error term on a plot basis could be useful in stress-prone low P fields, because non-adapted plants could die off in their early development. In many environments, a few individual plots showed extreme residuals even after spatial adjustment (see Figs 1 and 2). These outliers could be due to micro-variability because of termites or reduced plant stand at harvest and could mostly not be accounted for by an extra error term. Thus, the spatial models corrected for most of the field variation but could not correct for extreme outliers.

Heritability and RE are important measures for plant breeding trials. Even though a high heritability and a low SED are intended for plant breeding trials, they were not
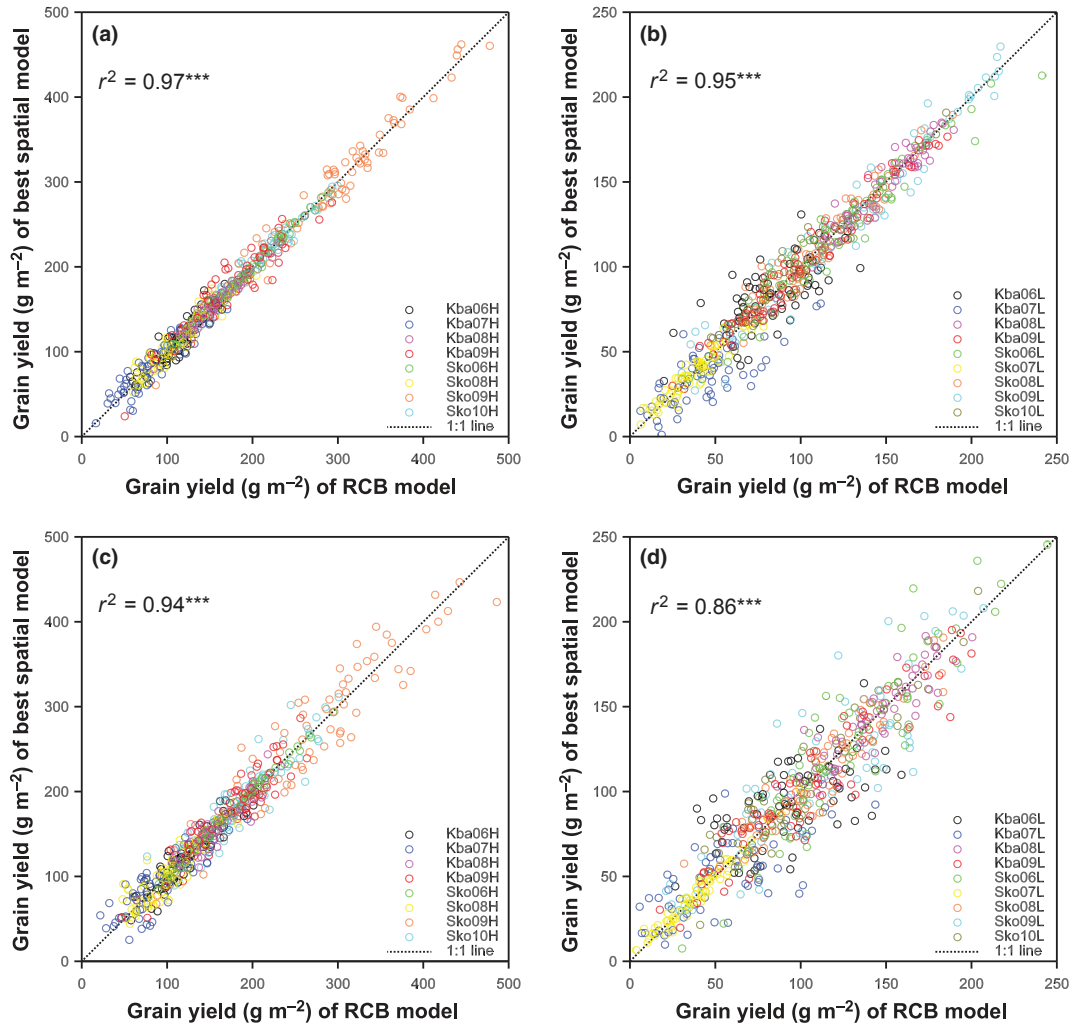
Fig. 3 Estimated adjusted mean grain yields (gm^{-2}) of 66 Sorghum varieties analysed with Best Spatial Model and Randomized Complete Block Design (RCB) model in eight high P (a, c) and nine low P (b, d) environments, respectively. (a) and (b) represent data analyzed including four replications while (b) and (c) show data from two replications.

used as criteria for model selection, because this would lead to selection of over-fitted models. A similar extent of increase in h^2 compared to RCB as in our study could be found in Australian wheat trials (Smith et al. 2001). In our study, heritability was generally higher in high P environments than in low P trials, indicating a stronger environmental error in low P trials and corresponding to the expectation of higher broad sense heritability values in high input trials (Ceccarelli 1989). A mean RE of spatial models of 123.87 % (4-reps) and 131.85 % (2-reps) could be considered rather low compared to results by Cullis and Gleeson (1991), Gilmour et al. (1997) and Singh et al. (2003) and are comparable to Qiao et al. (2000), Müller et al. (2010). Increase in RE and h^2 by the best spatial model was more pronounced in low P environments and in two-replicate analyses. This trend could not be found for the lattice

model, showing the particular advantages of spatial modelling relative to our lattice designs in low P fields and trials with few replications. In high-input trials, soil fertility differences are likely diminished by fertilizer application. Additionally, plant responses to small soil fertility differences between medium to high input are smaller than those from low to medium input (Marschner 1995, pp. 184–186; Voortman and Brouwer 2003), hence spatial variation for grain yield is lower in high-input trials.

The h^2 levels in low P trials and two-replicate trials approached those in high P and four-replicate trials when using spatial adjustments. Thus, using spatial adjustments can increase the effectiveness of direct selection in low P and less replicated trials. Also, reallocation of testing resources to increase genotypes tested at lower replication could be considered, because for example, mean RCB

model heritability from four-replicate analyses was actually lower than best spatial model heritability from two-replicate analyses. Especially studies aiming at improving stress tolerance (e.g. nutrient efficiency, salt tolerance, drought tolerance) are mostly confronted with rather high environmental errors. In these cases, spatial adjustment and adequate experimental design are crucial tools for improving phenotypic data and thus being able to find better associations between genotypic and phenotypic data such as in locating QTLs for different stress responses.

The heritability increases with spatial adjustments showed no relation to productivity level (mean yield) across the low P trials ($r^2 = 0.01$). However, in the high P trials, the heritability increases declined at higher productivity levels ($r^2 = 0.38$). Thus, use of spatial modelling is less important in higher productivity environments with P fertilizer application.

Even though environmental mean grain yields did not change through spatial models (data not shown), genotype means and thus genotype ranking changed. Especially in environments (e.g. Kba06L) with a big impact of spatial modelling, rank changes were frequently observed. Different genotype ranking after spatial modelling have been reported for chickpea, lentil, barley and wheat (Qiao et al. 2000, Singh et al. 2003). Genotype effect estimates including a rather high environmental error (like RCB estimates) can lead to over or underestimation of certain genotypes and therefore inappropriate variety selection. Genotype estimates of all environments analysed with RCB, lattice and best spatial model were correlated based on an approach used by Müller et al. (2010). The best spatial models estimates showed a significantly higher correlation ($r = 0.49$ (4-rep), $r = 0.41$(2-rep)) across all environments than the RCB ($r = 0.43$ (4-rep), $r = 0.36$ (2-rep)), but not significantly higher than the lattice model ($r = 0.47$ (4-rep), $r = 0.39$ (2-rep)) indicating that genotype estimates based on spatial modelling provided more precision for genotype selection compared to RCB estimates, but not compared to lattice estimates. Therefore, lattice models with efficiently laid out incomplete blocks should give high precision and be suitable for most environments. However, spatial adjustment would be preferred where it can be modelled, especially in low-input and less-replicated trials, because major increases in h^2 and RE were found compared to lattice models.

Spatial models have not been used for sorghum yield testing in West Africa up to now, but it appears that benefits from more widespread use may be likely, as the 17 trials analysed in this study sample the range of edaphic-, climatic- and productivity-conditions generally encountered in on-station testing. The use of spatial models would appear to be even more justified for the extensive on-farm testing conducted in the region, where strong spatial trends and lower replication are encountered. Widespread application of spatial models is feasible but should be always combined with an appropriate experimental design and field layout. In particular, spatial analysis should not be regarded as a substitute for good blocking. Simple spatial models (like AR1 × AR1) can provide most of efficiency gains. It is not necessary to do extensive modelling, as in this study with 91 models, to identify the best model, as the best spatial model provided only modest gains over simpler spatial models. Our experimental designs were α-designs which did not consider any row–column or error-correlation structure and were mostly laid out in an inappropriate direction. An appropriate experimental breeding design that corrects for two dimensions (row–column) or blocking along ridges would have been a feasible approach for capturing more field variation even without spatial models. Furthermore, as shown by Qiao et al. (2000) and Williams et al. (2006), the experimental designs can have a big impact on efficiency increase, and classical recovery of inter-row and inter-column information can always be combined with spatial analyses. In the design generation packages CycDesigN (VSN International Ltd., Hempstead, UK) and the R package DiGGer (NSW DPI, Wagga, Australia) different experimental row–column designs can be generated and, if needed, error-correlation structures for autoregressive spatial models can be considered for design generation. Although the error-correlation structure (decay factor) is often not known for trials in new fields, Williams et al. (2006) point out that spatial designs based on the linear variance models are quite robust to the choice of decay parameter included for design generation. For future plant breeding experiments, we recommend to use resolvable row–column experimental designs, facilitating correction for spatial trends in two dimensions and use of row–column models for the analysis. Additionally these models can be coupled with spatial add-on components to correct for any remaining global and local field trends.

Acknowledgements

References

Bänziger, M., and M. Cooper, 2001: Breeding for low input conditions and consequences for participatory plant breeding examples from tropical maize and wheat. Euphytica 122, 503–519.

Bray, R., and L. T. Kurtz, 1945: Determination of total, organic, and available forms of phosphorus in soils. Soil Sci. 59, 39–46.

Brouwer, J., L. K. Fussell, and L. Herrmann, 1993: Soil and crop growth micro-variability in the West African semi-arid tropics: a possible risk-reducing factor for subsistence farmers. Agric. Ecosyst. Environ. 45, 229–238.

Buerkert, A., A. Bationo, and H.-P. Piepho, 2001: Efficient phosphorus application strategies for increased crop production in sub-Saharan West Africa. Field Crops Res. 72, 1–15.

Ceccarelli, S., 1989: Wide adaptation: how wide? Euphytica 40, 197–205.

Ceccarelli, S., 1994: Specific adaptation and breeding for marginal conditions. Euphytica 77, 205–219.

Cullis, B. R., and A. C. Gleeson, 1989: The efficiency of neighbour analysis for replicated variety trials in Australia. J. Agric. Sci. 113, 233–239.

Cullis, B. R., and A. C. Gleeson, 1991: Spatial analysis of field experiments-an extension to two dimensions. Biometrics 47, 1449–1460.

Edmondson, R. N., 2005: Past developments and future opportunities in the design and analysis of crop experiments. J. Agric. Sci. 143, 27–33.

FAO, 2010: FAO Stat: Sorghum. http://faostat.fao.org/site/567/default.aspx [last accessed 5 February 2011].

Gilmour, A. R., B. R. Cullis, and A. P. Verbyla, 1997: Accounting for natural and extraneous variation in the analysis of field experiments. J. Agric., Biol. Environ. Stat. 2, 269–293.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson, 2009: ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, UK, www.vsni.co.uk [last accessed 29 May 2012].

Grondona, M. O., J. Crossa, P. N. Fox, and W. H. Pfeiffer, 1996: Analysis of variety yield trials using two-dimensional separable ARIMA processes. Biometrics 52, 763–770.

John, J. A., and E. R. Williams, 1995: Cyclic and Computer Generated Designs, 2nd edn. Chapman and Hall/CRC, London, UK.

Kehel, Z., D. Z. Habash, S. A. Gezan, S. J. Welham, and M. M. Nachit, 2010: Estimation of spatial trend and automatic model selection in augmented designs. Agron. J. 102, 1542.

Lamers, J. P. A., and P. R. Feil, 1995: Farmer's knowledge and management of spatial soil and crop growth variability in Niger, West Africa. Neth. J. Agric. Sci. 43, 375–389.

Marschner, H., 1995: Mineral Nutrition of Higher Plants. Academic Press, London, UK.

Möhring, J., and H.-P. Piepho, 2009: Comparison of weighting in two-stage analysis of plant breeding trials. Crop Sci. 49, 1977–1988.

Müller, B. U., K. Kleinknecht, J. Möhring, and H. P. Piepho, 2010: Comparison of spatial models for sugar beet and barley trials. Crop Sci. 50, 794–802.

Piepho, H.-P., and J. Möhring, 2007: Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177, 1881–1888.

Piepho, H. P., and E. R. Williams, 2010: Linear variance models for plant breeding trials. Plant Breeding 129, 1–8.

Piepho, H. P., C. Richter, and E. Williams, 2008: Nearest neighbour adjustment and linear variance models in plant breeding trials. Biom. J. 50, 164–189.

Qiao, C. G., K. E. Basford, I. H. DeLacy, and M. Cooper, 2000: Evaluation of experimental designs and spatial analyses in wheat breeding trials. Theor. Appl. Genet. 100, 9–16.

Raun, W. R., J. B. Solie, G. V. Johnson, M. L. Stone, R. W. Whitney, H. L. Lees, H. Sembiring, and S. L. Taylor, 1997: Small-scale variation in soil test phosphorus and Bermudagrass yield. Better Crops 81, 14–16.

Schabenberger, O., and F. J. Pierce, 2001: Contemporary Statistical Models for the Plant and Soil Sciences, 1st edn. CRC Press, Boca Raton, FL, USA.

Singh, M., R. S. Malhotra, S. Ceccarelli, A. Sarker, S. Grando, and W. Erskine, 2003: Spatial variability models to improve dryland field trials. Exp. Agric. 39, 151–160.

Smith, A., B. Cullis, and R. Thompson, 2001: Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57, 1138–1147.

Speed, T. P., E. R. Williams, and H. D. Patterson, 1985: A note on the analysis of resolvable block designs. J. R. Stat. Soc. Series B (Methodol.) 47, 357–361.

Vom Brocke, K., G. Trouche, E. Weltzien, C. P. Barro-Kondombo, E. Gozé, and J. Chantereau, 2010: Participatory variety development for sorghum in Burkina Faso: farmers' selection and farmers' criteria. Field Crops Res. 119, 183–194.

Voortman, R. L., 2010: Explorations into African Land Resource Ecology. On the chemistry between soils, plants and fertilizers. Phd Thesis, Vrije Universiteit, Amsterdam, the Netherlands.

Voortman, R. L., and J. Brouwer, 2003: An empirical analysis of the simultaneous effects of nitrogen, phosphorus and potassium in millet production on spatially variable fields in SW Niger. Nutr. Cycl. Agroecosyst. 66, 143–164.

Voortman, R. L., J. Brouwer, and P. J. Albersen, 2004: Characterization of spatial soil variability and its effect on millet yield on Sudano-Sahelian coversands in SW Niger. Geoderma 121, 65–82.

Wilding, L. P., 1985: Spatial variability: its documentation, accomodation and implication to soil surveys. In: D. R. Nielsen and J. Bouma, eds. Soil Spatial Variability, pp. 166–189. PUDOC, Wageningen, The Netherlands.

Williams, E. R., J. A. John, and D. Whitaker, 2006: Construction of resolvable spatial row-column designs. Biometrics 62, 103–108.

Wolfinger, R. D., 1996: Heterogeneous variance: covariance structures for repeated measures. J. Agric. Biol. Environ. Stat. 1, 205–230.