VITOR PASSOS DA SILVA JÚNIOR

EFFICIENCY OF AUGMENTED AND DOUBLE-REPLICATED DESIGNS UNDER AN ARRAY OF GENETICAL AND SPATIAL SCENARIOS

LAVRAS - MG
2020

VITOR PASSOS DA SILVA JÚNIOR

EFFICIENCY OF AUGMENTED AND DOUBLE-REPLICATED DESIGNS UNDER AN ARRAY OF GENETICAL AND SPATIAL SCENARIOS

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração Genética e Melhoramento de Plantas para a obtenção do título de Doutor.

Prof. Dra. Flávia Maria Avelar Gonçalves
Orientadora

LAVRAS - MG
2020

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Silva Júnior, Vitor Passos da.
Efficiency of augmented and double-replicated designs under an array of genetical and spatial scenarios / Vitor Passos da Silva Júnior. - 2020.
45 p. : il.

Orientador(a): Flávia Maria Avelar Gonçalves.
Tese (doutorado) - Universidade Federal de Lavras, 2020.
Bibliografia.

Melhoramento genético. 2. Experimentos de campo. 3. Análises espaciaisII. Gezan, Salvador Alexandro. III. Título.

VITOR PASSOS DA SILVA JÚNIOR

EFFICIENCY OF AUGMENTED AND DOUBLE-REPLICATED DESIGNS UNDER AN ARRAY OF GENETICAL AND SPATIAL SCENARIOS

EFICIÊNCIA DE DELINEAMENTOS EM BLOCOS AUMENTADOS E "DOUBLE REPLICATED" CONSIDERANDO DIFERENTES CENÁRIOS DE PEDIGREE E ANÁLISES ESPACIAIS

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração Genética e Melhoramento de Plantas para a obtenção do título de Doutor.

APROVADA em 04/12/2020

Dr. Bráulio Fabiano Xavier Moraes	UFLA
Dr. Breno Alvarenga Rezende	UFLA
Dr. Bruno Marco de Lima	ESALQ
Dr. José Luiz Lima	UFLA

Prof. Dra. Flávia Maria Avelar Gonçalves
Orientadora

LAVRAS - MG
2020

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me ajudado a concluir mais essa etapa em minha vida, foi uma longa e enriquecedora jornada. A qual tive a oportunidade de conviver com grandes mentes que hoje posso chamar de amigos. Professores que me auxiliariam muito além da sala de aula, me deram conselhos para a vida, os quais carregarei comigo até meus últimos dias.

Gostaria de começar agradecendo a pessoa que, com extrema paciência e carinho me conduziu por essa jornada, Flávia Maria Avelar Gonçalves, minha orientadora. O meu muito obrigado pelas conversas que tivemos, por ter me ajudado a amadurecer e principalmente, por sempre procurar entender o momento em que estava e me ajudando a passar por eles, bons ou ruins, sempre podia contar com a sua ajuda, muito obrigado!

Meus mais sinceros agradecimentos aos professores Antônio Magno Pato Ramalho e Marcio Balestre (*in memoriam*), os quais foram meus mentores por um longo tempo, me auxiliando muito tanto nas análises quanto na parte de melhoramento. Foram pessoas fundamentais para a construção do conhecimento que possuo hoje, serei eternamente grato por seus ensinamentos.

Aproveito também para agradecer ao professor Salvador A. Gezan, que me recebeu na Flórida, para meu doutorado sanduiche, foram momentos magníficos, discussões ricas em conhecimento, tive a oportunidade de trabalhar com diversas culturas, graças à sua plasticidade como melhorista. Aprendi muito com suas habilidades em programação e coragem ao enfrentar desafios, sua determinação e força de vontade muito me ajudam a tomar decisões, foram momentos incríveis, que guardarei com muito carinho. Agradeço também à todos os amigos que fiz na Florida, em especial as tias do inglês, os amigos do Sunbay e também ao Andrew e Priscila, os quais tiveram grande participação nesse processo.

Nessa longa jornada, aos meus pais, foram um pilar essencial, sempre me dando forças e encorajando a ir mais longe, muito obrigado! Aos meus irmãos, por sempre acreditarem e sonharem junto a mim. Aproveito também para agradecer a minha companheira, Daniela, por estar sempre ao meu lado me encorajando e me ajudando sem medir esforços. Davi, meu maior orgulho, muito obrigado por fazer parte da minha vida, você é sem dúvidas, minha maior conquista e minha felicidade diária.

Por onde passei, tive a felicidade de fazer grandes amizades, pessoas que tenho orgulho imenso de chamar de amigos, em especial o Guilherme, Mario, Rafael, Iolanda, Rafael, Zina e Indalecio, os admiro muito e torço para o sucesso de cada um de vocês.

Não posso deixar de agradecer a todos os funcionários da UFLA e da UF, os quais me acolheram e trabalharam para que tivéssemos sempre o melhor ambiente possivel. Vocês também foram fundamentais para a realização desse momento.

Agradeço o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – pela bolsa de doutorado.

Agradeço também a todos os pesquisadores, pois esse trabalho foi realizado graças à trabalhos prévios, com todo o apoio e suporte que foi fornecido é que pude não apenas enxergar, mas alcançar essa conquista.

RESUMO

Os programas de melhoramento em estágios iniciais avaliam centenas ou milhares de genótipos em um ou mais locais, exigindo experimentos que geralmente apresentam genótipos com baixa ou nenhuma replicação. Alguns dos designs propostos são os delineamentos aumentados e duplamente replicados (em que todos os genótipos têm duas repetições). Neste trabalho, simulações em computador foram utilizadas para avaliar o desempenho de delineamentos sem replicação ou replicação limitada para obter estimativas de componentes de variância, herdabilidade, valores clonais e genéticos, e seus ganhos genéticos alcançados sob uma variedade de cenários genéticos e espaciais. Os resultados de todas as simulações avaliadas indicam que a análise espacial forneceu resultados superiores ao modelo não espacial. A incorporação do efeito pepita trouxe melhores precisões, mesmo quando não há pepita em campo. A incorporação de informações de pedigree na estimativa de valores clonais e genéticos resultou em ganhos de precisão com boas correlações entre valores clonais e genéticos verdadeiros e preditos. No entanto, os benefícios das análises espaciais foram menos relevantes uma vez que o pedigree foi incorporado. Delineamentos duplamente replicados apresentaram melhor desempenho do que os delineamentos aumentados; e esses delineamentos com parcelas de controle de 6,25% mostraram resultados semelhantes aos delineamentos aumentados com parcelas de controle de 25%. Em resumo, o acréscimo de uma repetição, conferindo duas parcelas por genótipo fornece estimativas muito boas de valores genéticos. No entanto, ensaios não replicados podem fornecer estimativas razoáveis desde que seja incorporado análises espaciais e/ou informações de pedigree.

Palavras-chave: Melhoramento genético. Experimentos de campo. Análises espaciais.

ABSTRACT

Early stage breeding trials evaluate hundreds or thousands of genotypes in one or more sites, requiring experiments which often present genotypes with low or no replication. Some of the proposed designs are augmented design (AD), and double replicated (DR, where all genotypes have two replications). In this work, computer simulations were used to evaluate the performance of designs with no or limited replication to be used for estimation of variance components, heritability, genotypic and breeding values (BV), and their achieved genetic gains under an array of genetical and spatial scenarios. Results from all simulation evaluated indicate that spatial analysis provided superior results than a no-spatial model. The incorporation of nugget effect brought better accuracies, even when there is no nugget on the field. Incorporation of pedigree information on the estimation of BV resulted on accuracy gains with good correlations between true and predicted BVs. However, the benefits of spatial analyses were less relevant once pedigree was incorporated. DR designs presented better performance than AD; and AD designs with 6.25% control plots showed similar results than AD with 25% control plots. In summary, the increase of only one replication, conferring two plots per genotype provided with very good estimations of genetic values. However, unreplicated trials provides with reasonable estimates when under spatial analyses and/or when they incorporate pedigree information.

Keywords: Breeding. Field trials. Spatial analysis.

LIST OF ABBREVIATIONS AND ACRONYMS

AD	augmented design
AR1	first order auto regressive;
BLUP	best linear unbiased predictor
DR	double replicated
GE	genotype by environment interaction
MET	multi-environmental trial
PEV	predict error variance
P-rep	partial replicated
RCBD	randomized complete block design
REML	restricted maximum likelihood

SUMMARY

1 INTRODUCTION

Most breeding trials require the evaluation of genetic materials in one or more environments with the objective of estimating breeding or genotypic values. For these evaluations, it is necessary to choose the experimental design that optimizes resources in function of the objective, achieving the best possible accuracy of the tested genotypes. The design of these field experiments should follow three key principles: randomization, control, and replication (COCHRAN; COX, 1957; STEEL; TORRIE, 1980; WELHAM et al., 2014). Replication is of special interest, as it allows to obtain an estimate of the experimental error and its associated variance by sampling several microenvironments in a field trial. Replication also plays an important role on the accuracy of the treatment (or genotype) mean estimates, where as the number of replications increases, the standard errors associated with this estimate decreases (COCHRAN; COX, 1957).

In most plant breeding programs, early stage trials often evaluate hundreds or thousands of genotypes in one or more sites, requiring challenging experiments which often present genotypes with low (<3) or no replication. Federer (1955) proposed a field layout capable of handling this sort of problem: the augmented design (AD). These designs evaluate a large set of unreplicated test genotypes together with a small set of replicated control genotypes used as checks. To generate these experiments, control genotypes are often arranged as a randomized complete block design (RCBD) and later, test genotypes are added to augment these blocks. One important aspect of these designs is the proportion of control plots to be considered, some recommendations suggest that 10 to 25% of the plots should be with checks (KEMPTON, 1984; KEMPTON; GLEESON, 1997; MÜLLER et al., 2010).

The advantage of the AD designs is the large number of genotypes that can be evaluated; however, one of its drawbacks is the low precision of the estimation of the genetic effects as a result of this lack of replication. In addition, AD designs assign to many plots to checks, limiting

the number of test genotypes, and leading to good estimates of control plots, which are not the main objective of these studies. An interesting variant of the AD that handles some of these drawbacks is the P-rep designs, where a proportion of the test genotypes are replicated and therefore, they are considered as checks in the statistical analyses. A property of these designs is that the increase of replications of a proportion of the test genotypes improves the efficiency of selection (CULLIS et al., 2006; WILLIAMS et al., 2011).

In contrast, fully replicated designs will not present these issues, but require considerable amounts of plant material and space to establish all replicated plots. A possible compromise is to establish field experiments, where all genotypes are replicated only twice, and experiment identified here as double replicated (DR). This limited replication will allow for reasonable precision on the estimate of the genetic effects, while limiting the use of resources.

Statistical analyses for most field experiment, not only AD, P-rep and DR, benefit from the incorporation of spatial correlations (GILMOUR et al., 1997) to efficiently control for environmental sources of variations that might be present in a trial, which are often not captured by the design factors (such as row, columns or incomplete blocks). In genetic analyses, modeling the error variance with spatial structure increases the precision of the estimates of the genotypic or breeding values (BURGUEÑO, 2018; GEZAN et al., 2006), leading to better selections, and therefore larger genetic gains. The use of the spatial matrices, such as the first order separable auto-regressive error variance-covariance (AR1) has proven to be effective (CULLIS et al., 1998). This error structure has been used broadly in agricultural (CULLIS; GLEESON, 1991; ZIMMERMAN; HARVILLE, 1991; GILMOUR et al., 1997; PIEPHO et al., 2015) and forestry trials (COSTA e SILVA et al., 2001; DUTKOWSKI et al., 2002). An important additional component of the AR1 error structure is the nugget (or microsite) variance, that is a parameter used to account for variability at short distances due to measurement errors or local random noise

(DALE; FORTIN, 2014). In field trials, nugget captures part of the heterogeneity of soil fertility and/or measurement error (PIEPHO et al., 2015).

Selecting the best spatial model to fit is a difficult task. There are many criteria to choose, such as: Akaike criteria of information (AIC) (AKAIKE, 1974), Bayesian criteria of information (BIC) (SCHWARZ, 1978) and also the REML log-likelihood value can be used as a criterion. For genetic analyses, it is common to use the H^2PEV (WELHAM et al., 2010), that relies on the predictor error variance. Data and analyses simulations present a good opportunity to evaluate these different criteria by comparing differences between true and predicted fixed or random effects.

The main objective of this study is to evaluate, through computer simulations, the performance of experimental designs with no or limited replication (i.e., AD and DR designs) to be used for estimation of variance components, heritability, genotypic and breeding values, and their achieved genetic gains under an array of genetical and spatial scenarios. The secondary objectives of this study are: 1) to contrast the benefits/drawbacks of the use of AD against DR designs, 2) to compare the effects of different levels of spatial correlation and nugget on estimation of genetic parameters, 3) to evaluate the goodness-of-fit statistics of an array of linear mixed models with and without spatial components, 4) to evaluate the accuracy of the estimation of breeding values under limited replication, and 5) to quantify the levels of genetic gains in terms of genotypic or breeding values under different scenarios for designs with no or limited replication. For this, an array of array of genetical and environmental scenarios were considered with different levels of narrow-sense heritability's and dominance, spatial correlations, presence or absence of nugget effects, and selection intensities were simulated and evaluated for a scenario with the aims of estimating genotypic values, and another with the focus on breeding values.

2 MATERIAL AND METHODS

2.1 Spatial Error Structure

For this study all simulations were based on a field trial with 1,024 plots on a contiguous rectangular grid of 64×16 positions, where each plot contained a single observation. The spatial surface was modelled by using an autoregressive of first order error structure with and without nugget (AR(x)⊗AR(y) + $\sigma_\eta^2$and AR(x)⊗AR(y), respectively). The AR(x)⊗AR(y) matrix considers two correlations, one on the x direction (ρ_x) and another on the y direction (ρ_y), which are perpendicular, and therefore this an anisotropic model. These parameters, ρ_x and ρ_y, correspond to spatial error correlation between the residuals defining a patch of similar residual values, and the expressions for this structure for two residuals e_i and $e_{i'}$ is: $cov(e_i, e_{i'}) = {\sigma^2}_e \rho_x^{hx} \rho_y^{hy}$ for off-diagonal elements, and $Var(e_i) = {\sigma^2}_e + {\sigma^2}_\eta$ for diagonal elements where $h_x = |x_i - x_{i'}|$, $h_y = |y_i - y_{i'}|$ are the absolute distance between plots in row and column positions, respectively; ${\sigma^2}_e$ is the random spatial error variance and η is the nugget or microsite error variance.

Based on the autoregressive error structure specified above, two sets of 500 sites with and without nugget were simulated considering ρ_x and ρ_y values that randomly varied between 0.02 and 0.98, and where their absolute difference was smaller than 0.85. Also, ${\sigma^2}_e + \eta$ was considered, for simplicity, equal to 1, where for the sites with nugget, the proportion of the variance for η ranged randomly between 30% and 70%. All the above site simulations were later used to superimpose field experiments with different statistical designs that were then analyzed to estimate genotypic or breeding values. Details of these experiments are detailed below for simulation of genotypic and breeding values.

2.2 Simulation for Genotipic Values

The experimental layouts considered to estimate genotypic values corresponded to three augmented designs (AD) all with eight blocks and four control genotypes or checks with different proportions of test plots out of the 1,024 total plots. The selected proportions corresponded to 6.25%, 12.5% and 25% (thereafter identified as AD6.25, AD12.5 and AD25, respectively). For simplicity, the distribution of the control plots was at random, as the arrangement does not provide no significative differences (MÜLLER *et al.*, 2010). These proportions of control plots were selected based on typical augmented experiments (PAYNE, 2006; MÜLLER *et al.*, 2010). In addition, a randomized complete block design based on two blocks with no control plots but with all genotypes replicated twice was considered, and therefore this is identified as a double-replication experiment (DR). Further details of the characteristics of these experiments are summarized in the supplementary material (TABLE S1).

Table S1. Details of experimental designs considered specifying number of plots and genotypes for control and test treatments. All designs form a grid of 64×16 positions for a total of 1,024 plots. AD: augmented design, DR: double replicated design.

Designs	Proportion Control Plots	# Plots		# Genotypes		# Blocks
		Control	Test	Control	Test	
AD6.25	6.25%	64	960	4	960	8
AD12.50	12.5%	128	896	4	896	8
AD25	25%	256	768	4	768	8
DR	0%	0	512	0	512	2

To simulate the phenotypic response the following model was used: $y_{ijk} = \mu + c_k + e_{(ijk)} + e_{\eta(ijk)}$; where y_{ijk} is the response variable of the k^{th} clone located in the i^{th} row and j^{th} column, μ is a population mean, which was arbitrarily set to 10 units; c_k is the random genetic genotypic effect simulated, with distribution $c_k \sim N(0, \sigma^2{}_c$; $e_{(ijk)}$ is the spatial error (or structured residual), with $e_{(ijk)} \sim N(0, \sigma^2{}_e$; and $e_{\eta(ijk)}$ corresponds to the microsite random error (or

unstructured residual) with distribution $e_{\eta(ijk)}$~N(0, ${\sigma^2}_{\eta}$. For all simulations, the genotypic variance, c_k was set to 0.5 and the total error variance, ${\sigma^2}_e + \eta$, was set to be 0.5, and therefore, broad-sense heritability corresponded to $H^2 = {\sigma^2}_c / \left({\sigma^2}_c + {\sigma^2}_e + {\sigma^2}_{\eta}\right) = 0.50$. This scheme provided with a total of 4,000 simulations (4 designs × 2 surfaces × 500 sites) generated and stored for further statistical analyses.

2.3 Simulation for Breeding Values

For this section, the designs selected to assess breeding values estimation corresponded to a subset of the experimental layouts previously described (i.e., AD6.25, AD25 and DR) for a total of 960, 768 and 512 clones, respectively. These belonged to a circular diallel crossing design based on 42 unrelated parents (21 males and 21 females) for a total of 64 families, where each parent was used in 3-4 crosses (see Figure S1). Here, a total of 15, 12 and 8 clones were considered for each of these 64 families for the designs AD6.25, AD25 and DR, respectively conforming to a total of 1,024 plots (TABLE S2). Hence, this testing population is formed by half- and full-sib individuals that allow for partition of additive, dominance and epistasis, thanks to the diallel crossing structure.

Table S2. Details of experimental designs considered specifying number of plots and families for control and test treatments. All designs form a grid of 64×16 positions for a total of 1,024 plots. AD: augmented design, DR: double replicated design.

Designs	# Plots		#Family		# Blocks
	Control	Test	Families	Clones per family	
AD6.25	64	960	8	15	8
AD25	256	768	8	12	8
DR	0	512	8	8	2

As before, to simulate the phenotypic response the following linear model was used: $y_{ijklm} = \mu + a_k + d_k + i_k + e_{(ijk)} + e_{\eta(ijk)}$; where a_k is the additive effect with distribution $a_k \sim$ N(0, ${\sigma^2}_a$) represents the dominance effect with $d_k \sim$ N(0, ${\sigma^2}_d$); i_k is the epistatic effect for the k^{th} individual with $i_k \sim$ N(0, ${\sigma^2}_i$); and all other terms were previously described. The variance-covariance of the a_k and d_k effects were modelled by considering the pedigree-based numerator relationship (**A**) and dominance (**D**) matrix, respectively (SCHAEFFER *et al.*, 1989; MRODE, 1996).

Two distinct genetic scenarios were simulated with different proportion of additive variance in relation to non-additive variance; one with high proportion (E1: $h^2 = 0.40$, $d^2 = 0.05$ and $i^2 = 0.05$) and another with low proportion (E2: $h^2 = 0.20$, $d^2 = 0.15$ and $i^2 = 0.15$). For these scenarios the total genetic variance ($h^2 + d^2 + i^2$) and error variance (${\sigma^2}_e + {\sigma^2}_\eta$) were both set to 0.50, corresponding to a total phenotypic variance of 1.0. For this subset of simulations only sites with nugget were evaluated; therefore, a total of 3,000 simulated datasets (3 designs × 1 surface × 500 sites x 2 genetic scenarios) were available.

2.4 Statistical Models

The linear mixed models to evaluate each of the simulated datasets for the three or four experimental designs considered the following options: non-spatial (M1), spatial without nugget (M2), and spatial with nugget (M3).

The fitted model for the simulations for genotypic values was:

$$y = 1\mu + Z_1 b + Z_2 g + e$$

where μ is the overall mean; $\boldsymbol{b}$ is a vector of random effects for blocks, with $\boldsymbol{b} \sim$ MVN($\boldsymbol{0}$, $\sigma_b^2 \boldsymbol{I_b}$); $\boldsymbol{g}$ is a vector of random effects for the genotypes, with $\boldsymbol{g} \sim$ MVN($\boldsymbol{0}$,$\sigma_g^2 I_g$; $\boldsymbol{e}$ is the vector of

random residuals, with $\boldsymbol{e}$ ~ MVN($\boldsymbol{0}$, $\boldsymbol{R}$). $\boldsymbol{1}$ is a vector of ones; $\boldsymbol{I}$ is an identity matrix of its corresponding size; $\boldsymbol{X}$, $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are incidence matrices for their respective factors. The residual variance-covariance matrix $\boldsymbol{R}$ was specified for each of the models with $\boldsymbol{R_1} = \sigma_e^2\boldsymbol{I_n}$, $\boldsymbol{R_2} = \boldsymbol{AR(x)}\otimes\boldsymbol{AR(y)}$, and $\boldsymbol{R_3} = \boldsymbol{AR(x)}\otimes\boldsymbol{AR(y)} + \sigma_\eta^2\boldsymbol{I_n}$, for models M1, M2 and M3, respectively. Here, $\boldsymbol{I_n}$ is an identity matrix of dimension *n*, where *n* is the total number of observations; $\boldsymbol{AR(\cdot)}$ are autoregressive order 1 correlation matrices for x- and y-directions; and $\sigma_\eta^2$is a nugget variance term.

Equivalently, the fitted generic model for the simulations for breeding values was:

$$y = 1\mu + X\beta + Z_1 b + Z_2 a + Z_3 f + Z_4 c + e$$

where $\boldsymbol{\beta}$ is the fixed effect of checks; $\boldsymbol{b}$ is a vector of random effects for blocks, with $\boldsymbol{b}$ ~ MVN($\boldsymbol{0}$, $\sigma_b^2\boldsymbol{I_b}$)*;* $\boldsymbol{a}$ is a vector of individual additive effects for each genotypes, with $\boldsymbol{a}$ ~ MVN($\boldsymbol{0}$,$\sigma_a^2 A$; $\boldsymbol{f}$ is a vector of random effects of family, with $\boldsymbol{f}$ ~ MVN($\boldsymbol{0}$,$\sigma_f^2 I_f$; $\boldsymbol{c}$ is a vector of random effects clone within a family, with $\boldsymbol{c}$ ~ MVN($\boldsymbol{0}$,$\sigma_c^2 I_c$; $\boldsymbol{e}$ is the vector of random residuals, with $\boldsymbol{e}$ ~ MVN($\boldsymbol{0}$, $\boldsymbol{R}$). $\boldsymbol{A}$ corresponds to the numerator relationship matrix obtained from pedigree, and $\boldsymbol{R}$ was identical to $\boldsymbol{R_1}$, $\boldsymbol{R_2}$ or $\boldsymbol{R_3}$, as described earlier for models M1, M2 and M3, respectively. All other terms were previously defined. The above model was fitted for all simulations for the DR datasets; however, for the unreplicated experiments AD6.25 and AD25 the same model was fitted but here the term $\boldsymbol{f}$ was dropped to facilitate model fitting, and therefore the variance associated with this term is contained (confounded) within the genotypic factor $\boldsymbol{c}$. In summary, there were 12,000 analyses for genotypic values (4 × 2 × 500 × 3) and 9,000 for breeding values (3 × 1 × 500 × 2 × 3) to be run.

2.5 Statistical Comparisons

Each of the analyses were used to obtain summary statistics for comparisons. First, for each of the estimated variance components means and standard deviations were calculated, together

with the REML log-likelihood for each analysis (logREML). For genotypic values plot-level broad-sense heritability was estimated as: $H_c^2 = \sigma_g^2/(\sigma_g^2 + \sigma_e^2)$. For breeding values analyses, the narrow-sense heritability (h^2), dominance ratio (d^2), epistatic ratio (i^2), and broad-sense heritability (H_c^2) were estimated using the following expressions: $h^2 = \sigma_a^2/(\sigma_a^2 + \sigma_f^2 + \sigma_c^2 + \sigma_b^2 + \sigma_e^2 + \sigma_\eta^2)$, $d^2 = 4\sigma_f^2/(\sigma_a^2 + \sigma_f^2 + \sigma_c^2 + \sigma_b^2 + \sigma_e^2 + \sigma_\eta^2)$, $i^2 = (\sigma_c^2 - 3\sigma_f^2)/(\sigma_a^2 + \sigma_f^2 + \sigma_c^2 + \sigma_b^2 + \sigma_e^2 + \sigma_\eta^2)$, and $H_c^2 = (\sigma_a^2 + \sigma_f^2 + \sigma_c^2)/(\sigma_a^2 + \sigma_f^2 + \sigma_c^2 + \sigma_b^2 + \sigma_e^2 + \sigma_\eta^2)$.

Recall that for the AD6.25 and AD25 experiments σ_f^2 is confounded within σ_c^2. In addition, a different expression of heritability, H^2_{PEV} and h^2_{PEV}, was used corresponding to: $H^2{}_{PEV} = 1 - \overline{PEV}/\sigma_g^2$ and $h^2{}_{PEV} = 1 - \overline{PEV}/\sigma_a^2$ (Welham et al. 2010), where $\overline{PEV}$ corresponds to the average predictor error variance from each genotypic (or additive) random effects, and σ_g^2 (or σ_a^2) is the variance component associated with the specific random effect. This expression is a generalized definition of heritability proposed by Cullis *et al.* (2006) and it represents the genotype mean broad-sense heritability (for genotypic values) and the mean narrow-sense heritability (for breeding values).

Finally, other statistics were calculated in order to assess the quality of the predictions for each of the fitted datasets and models in relation to the true genetic (genotypic) or breeding values. Pearson product-moment correlation was obtained between the predicted and the true values, for genotypic values ($CorP_c$) and breeding values ($CorP_a$). For each simulation, an estimation of the selection efficiency (SEF) was obtained by selecting the top 10, 40 and 100 clones based on their estimated BLUPs (i.e., using incomplete information), and a ratio was calculated based on the average true genetic value of these selection against the genetic values from the actual 10, 40 and 100 top individuals. Hence, for example, a value of 82% indicates that a given selection of genotypes provides with 82% of the maximum potential genetic gain. For the breeding values

evaluations, selection efficiency was calculated only the top 40 individuals (SEF_{40}) according to their estimated breeding values. Another statistic used corresponded to the genetic gain achieved by selecting the top 40 individuals based on their true genetic value (PM_{40}); hence, the larger these values the greater the realized genetic gains based on the selections.

All the statistical analyses, simulations and calculations were done using the package R version 3.4.4 (R CORE TEAM 2018). The fitting of the linear mixed models was done using ASReml-R (BUTLER *et al.*, 2009) as implanted in the package R. The package AGHmatrix (AMADEU *et al.*, 2016) was used to obtain the relationship matrix for simulations. Code is available upon request from the authors.

3 RESULTS

3.1 Genotypic Values

For the genotypic value simulations, in all scenarios, the genetic variance component (σ^2_g) estimates were unbiased, with their average value close to its true value of 0.5, with the exception of the analyses of the models M2 for the sites with nugget that presented some overestimation, with the highest average values for AD6.25, but this bias was still present for all designs for model M2 (TABLE 1). This genetic variance bias was followed by an important underestimation of the spatial correlations (ρ_x, ρ_y) for the same model M2 and across all designs resulting in an average spatial correlation estimate of 0.264 instead of 0.5. However, variance component estimations under the sites without nugget presented minimal bias in genetic variance and spatial correlation estimates for model M2, with a small overestimation of these correlations for model M3. For the replicate variance component (σ^2_b), bias values ranging from 0.004 to 0.043 were observed where its true value is 0; this bias was more relevant under M1 model for all designs, with the exception of DR design. This is likely to occur because the REML variance component estimates are constrained to be positive.

Table 1. Variance components, spatial correlations and heritability estimates for the four designs using the three statistic models analyzed for sites with and without nugget for genotypic value simulations.

		Sites with Nugget					
Designs	**Models**	σ^2_b	σ^2_g	σ^2_e	σ^2_η	ρ_x	ρ_y
AD6.25	**M1**	0.020	0.505	0.477	-	-	-
	M2	0.011	0.562	0.443	-	0.298	0.301
	M3	0.005	0.502	0.275	0.223	0.496	0.494
AD12.5	**M1**	0.020	0.504	0.480	-	-	-
	M2	0.012	0.544	0.463	-	0.271	0.270

	M3	0.004	0.506	0.272	0.228	0.499	0.499
	M1	0.021	0.496	0.484	-	-	-
AD25	**M2**	0.014	0.523	0.479	-	0.243	0.245
	M3	0.004	0.497	0.272	0.232	0.497	0.495
	M1	0.008	0.503	0.494	-	-	-
DR	**M2**	0.007	0.512	0.492	-	0.245	0.242
	M3	0.004	0.504	0.268	0.237	0.511	0.504
Sites without Nugget							
Designs	**Models**	σ^2_b	σ^2_g	σ^2_e	σ^2_η	ρ_x	ρ_y
	M1	0.043	0.497	0.465	-	-	-
AD6.25	**M2**	0.007	0.495	0.514	-	0.507	0.500
	M3	0.007	0.484	0.494	0.030	0.525	0.514
	M1	0.040	0.504	0.461	-	-	-
AD12.5	**M2**	0.006	0.503	0.506	-	0.511	0.503
	M3	0.006	0.497	0.489	0.022	0.528	0.515
	M1	0.040	0.500	0.465	-	-	-
AD25	**M2**	0.007	0.498	0.512	-	0.507	0.505
	M3	0.006	0.495	0.496	0.019	0.523	0.519
	M1	0.016	0.502	0.490	-	-	-
DR	**M2**	0.006	0.501	0.512	-	0.512	0.504
	M3	0.006	0.500	0.499	0.015	0.526	0.513

† AD6.25, AD12.5 and AD25 represents augmented designs with 6.25%, 12.5% and 25% of replicates respectively and DR is the double-replication design. M1 is the no-spatial model, M2 is the spatial model AR(x)⊗AR(y) and M3 is the spatial model AR(x)⊗AR(y) +σ^2_η; σ^2_b, σ^2_g, σ^2_e and σ^2_η represents the variance components for block, clone, residual and nugget, and ρ_x and ρ_y are the spatial correlations for row and column.

The fitting statistic logREML for the case of sites with nugget indicated that, as expected, the best model is the model that included this component, i.e., M3; and this model was clearly superior to M2 and M1 (TABLE 2). In contrast, for those sites without nugget, minimal differences were detected in logREML between M2 and M3, but interestingly, model M3 resulted marginally better. The estimation of the broad-sense heritability (H^2_c), as expected, presented a similar pattern

to the estimate of the genotype variance component, with some bias only for the M2 model for sites with nugget (TABLE 2). A boxplot generated with the H^2_c estimates (Figure 1a) indicates that under the no-spatial model (M1) for AD6.25 there was a large range of estimates; however, this range gets reduced once spatial analyses (M2) were performed. Across designs, for the no-spatial analyses, AD25 and DR present much smaller range of H^2_c estimates.

Table 2. Summary statistics for the three designs using the three statistic models analyzed for sites with and without nugget for genotypic value simulations. Numbers in bold corresponds to the best model for a given statistics within a design.

Sites with Nugget							
Designs	**Models**	**logREML**	**H^2_c**	**H^2_{PEV}**	**$CorP_c$**	**SEF_{40}**	**PM_{40}**
AD6.25	**M1**	-490.32	0.503	0.507	0.711	71.1%	11.52
	M2	-465.42	0.552	**0.612**	0.736	73.6%	11.57
	M3	**-460.15**	0.499	0.544	**0.740**	**74.0%**	**11.58**
AD12.5	**M1**	-470.13	0.501	0.504	0.713	71.1%	11.50
	M2	-443.35	0.533	**0.590**	0.737	73.4%	11.55
	M3	**-436.42**	0.501	0.548	**0.743**	**74.0%**	**11.56**
AD25	**M1**	-425.56	0.494	0.498	0.711	71.0%	11.45
	M2	-395.66	0.514	**0.567**	0.738	73.7%	11.51
	M3	**-386.28**	0.493	0.542	**0.744**	**74.3%**	**11.52**
DR	**M1**	-437.55	0.499	0.660	0.817	81.8%	11.76
	M2	-396.98	0.506	**0.713**	0.840	83.8%	11.81
	M3	**-384.62**	0.497	0.705	**0.845**	**84.6%**	**11.82**
Sites without Nugget							
Designs	**Models**	**logREML**	**H^2_c**	**H^2_{PEV}**	**$CorP_c$**	**SEF_{40}**	**PM_{40}**
AD6.25	**M1**	-481.43	0.493	0.503	0.719	72.4%	11.55
	M2	-381.78	0.488	**0.652**	**0.821**	**82.2%**	**11.76**
	M3	**-381.54**	0.478	0.637	**0.821**	**82.2%**	**11.76**
AD12.5	**M1**	-460.42	0.500	0.509	0.718	72.3%	11.52
	M2	-343.55	0.496	**0.672**	**0.826**	**82.9%**	**11.75**
	M3	**-343.31**	0.491	0.663	**0.826**	**82.8%**	**11.75**
AD25	**M1**	-415.93	0.496	0.505	0.719	71.9%	11.47
	M2	-261.13	0.490	**0.684**	**0.837**	**83.6%**	**11.71**
	M3	**-260.90**	0.487	0.680	**0.837**	**83.6%**	**11.71**
DR	**M1**	-434.64	0.497	0.654	0.818	81.7%	11.76
	M2	-204.60	0.493	**0.811**	**0.914**	**91.3%**	**11.97**
	M3	**-204.36**	0.491	0.808	**0.914**	**91.3%**	**11.97**

† AD6.25, AD12.5 and AD25 represents augmented designs with 6.25%, 12.5% and 25% of replicates respectively and DR is the double-replication design. M1 is the no-spatial model, M2 is the spatial model AR(x)⊗AR(y) and M3 is the spatial model AR(x)⊗AR(y) $+\sigma_{\eta}^{2}$; logREML is the log-likelihood REML value, H^2_c is the broad-sense heritability, H^2_{PEV} is the predictor-error-variance heritability, $CorP_c$ is the correlation between true and estimated genotypic values, SEF_{40} is the selection efficiency of selecting the top 40 individuals based on genotypic values, and PM_{40} is the genetic gain achieved by selecting the top 40 individuals.

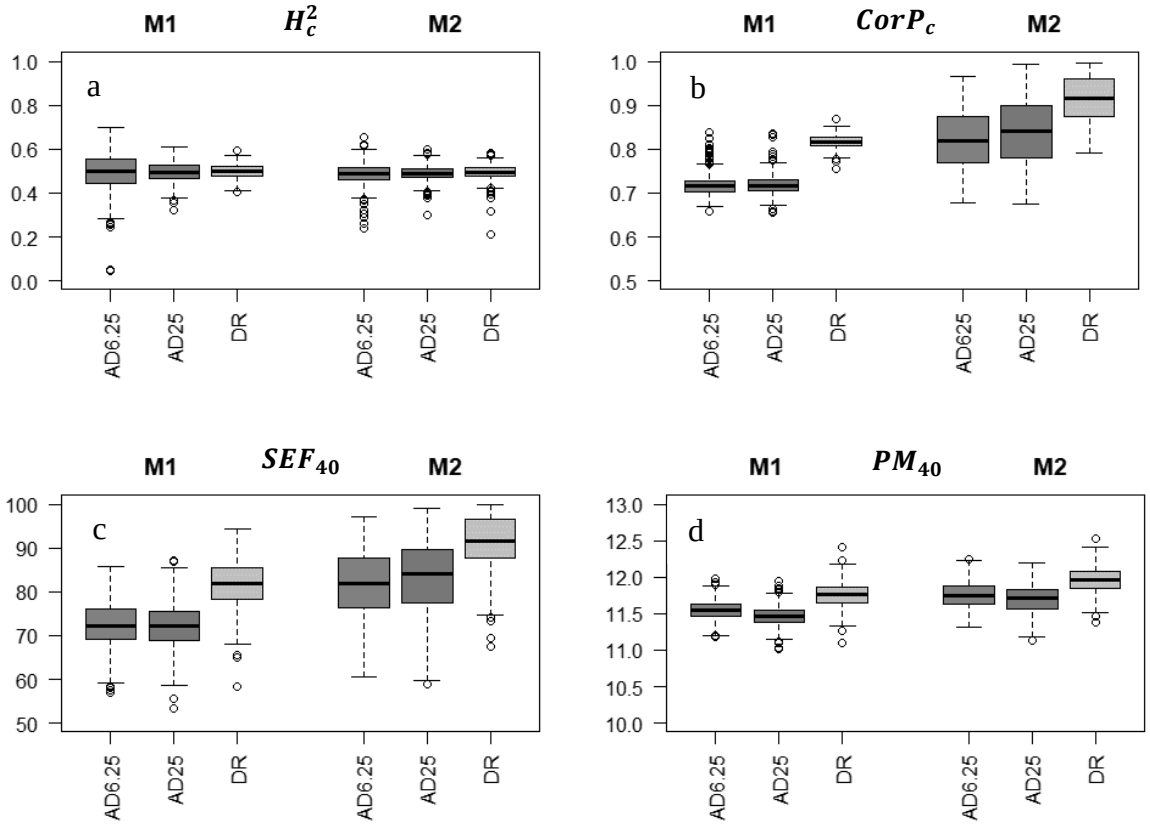

Figure 1. Box plots for the broad-sense heritability H^2_c (a), correlation between true and estimated genotypic values ($CorP_c$) (b), selection efficiency of selecting the top 40 individuals (SEF_{40}) (c)and genetic gain from the top 40 individuals (PM_{40}) (d) for no-spatial model (M1) and spatial model AR(x)⊗AR(y) (M2) for the sites without nugget for genotypic values simulation.

The average of the H^2_{PEV} values indicated that the best model, for both site surfaces, is M2; this is more noticeable for AD6.25 and with almost null differences for DR (TABLE 2). This appears to be a bias resulting by having overestimated genetic variance components as indicated earlier. Also, higher values of H^2_{PEV} were found for the sites without nugget with an average of

0.648 against 0.582. Interestingly, in the sites without nugget for any design, the fitting of a model that considers spatial analysis produced a relevant increase of 0.16 on H^2_{PEV}. As a result of having replication, an increase of 0.15 was noted between the use of a DR design instead of any augmented design, for all models.

Promising correlations between true and predicted genotypic values ($CorP_c$) were found in this study with values ranging from 0.711 to 0.914 (average of 0.785), with larger values for sites without nugget, models M2 and M3, and for the DR design (TABLE 2). The average $CorP_c$ values for sites without nugget across all models were 0.787, 0.790, 0.797, 0.882 for AD6.25, AD12.5, AD25 and DR, respectively. Overall, fitting spatial analyses models (M2 or M3) provided with an increase of the $CorP_c$ from 0.743 to 0.850 in relation to the no-spatial analyses (M1). However, these benefits vary considerably according to the level of spatial correlations (i.e., ρ_x and ρ_y).

In Figure 2, the $CorP_c$ for different spatial correlations values are presented as contour plots. Here, narrow ranges of $CorP_c$ values were observed for the no-spatial analyses indicating little or no sensitivity to spatial correlations. In contrast, for the M2 analyses a clear trend was noted with increasing values of $CorP_c$ as the spatial correlations get closer to 1. The highest $CorP_c$ for any of the designs were found when both spatial correlations ranged approximately between 0.75 and 0.98. A similar trend is noted for H^2_{PEV} with a clear increase of these values on the spatial correlations range of 0.67 to 0.98 (see Supplementary Figure S2). Interestingly, the results for $CorP_c$ and H^2_{PEV} (Figure 2 and Figure S2, respectively) show that high values of these statistics were achieved even when only one of the spatial correlations is high (> 0.75). In terms of $CorP_c$, the spatial analysis for the different designs, as indicated above, yielded higher average correlations between true and predicted; however, the spatial analyses present larger estimate ranges in comparison to the no-spatial analyses (Figure 2b). This increase in variability is likely to be due to the fact that spatial analyses provide with improvements on the genotypic estimation for some sites and no improvement for other sites.

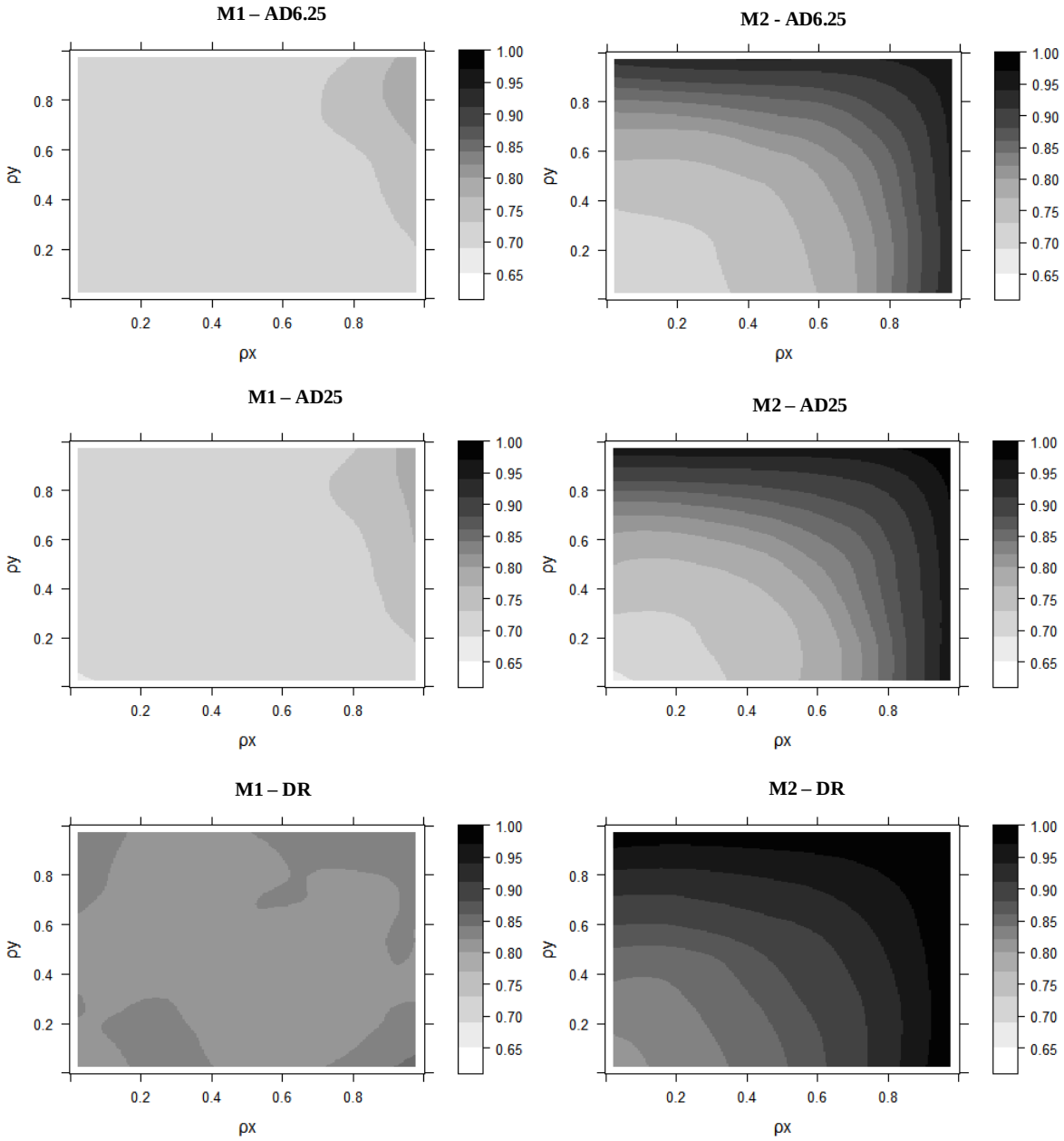

Figure 2. Contour plots for $CorP_c$ obtained by using the true row and column spatial correlations for the designs AD6.25, AD25 and DR fitted with no-spatial model (M1) and spatial model AR(x)⊗AR(y) (M2) for the sites without nugget for genotypic values simulation.

In relation to the statistic selection efficiency for AD6.25 the selection of the top 40 genotypes resulted in an average SEF_{40} of ~72% and ~83% for the no-spatial and spatial analyses, respectively. Also, as expected, the best results were obtained from the DR design with spatial analysis resulting in an SEF_{40} of ~91%. (Figure 2c) Hence, the use of spatial analyses results in ~11% increase in SEF_{40} for any design. Similar trends were observed for the top 10 and 100 individuals with SEF_{10} and SEF_{100} (TABLE S3). Under the spatial models, the consequences of low (≤ 0.5), or high (> 0.5) spatial correlations are shown in Figure 3 for SEF_{40}. Here, for AD6.25 and AD25 under model M2, larger values of spatial correlation yield to important increases on SEF_{40} changing from ~76% to ~89%; nevertheless, this effect is almost null under the DR design, where for the no-spatial and spatial analyses the SEF_{40} values are both approximately 92%. As expected, small differences were noted between low and high spatial correlations for model M1, regardless of the design.

Table S3. Summary statistics for the three designs using the three statistic models analyzed for sites with and without nugget for genotypic value simulations. Numbers in bold corresponds to the best model for a given statistics within a design.

Designs	**Models**	**WITH NUGGET**		**WITHOUT NUGGET**	
		SEF_{10}	**SEF_{100}**	**SEF_{10}**	**SEF_{100}**
AD6.25	**M1**	0.716	0.711	0.719	0.719
	M2	0.737	0.735	**0.820**	**0.822**
	M3	**0.742**	**0.739**	**0.820**	0.821
AD12.5	**M1**	0.711	0.712	0.717	0.721
	M2	0.733	0.738	**0.825**	**0.826**
	M3	**0.742**	**0.743**	**0.826**	**0.826**
AD25	**M1**	0.708	0.711	0.721	0.720
	M2	0.735	0.736	**0.838**	**0.835**
	M3	**0.743**	**0.743**	**0.838**	**0.835**
DR	**M1**	0.818	0.817	0.826	0.818
	M2	0.841	0.841	**0.911**	**0.914**
	M3	**0.843**	**0.846**	**0.911**	**0.914**

† AD6.25 and AD25 represents augmented designs with 6.25% and 25% of replicates respectively and DR is the double-replication design. M1 is the no-spatial model, M2 is the spatial model AR(x)⊗AR(y) and M3 is the spatial model AR(x)⊗AR(y) $+\sigma_{\eta}^{2}$; where SEF_{10} and SEF_{100} represents the efficiency selection for the top 10 and top 100 individuals, respectively.

In terms of genetic gain calculated based on the true genetic values of selecting the top 40 individuals, as measured by PM_{40}, spatial analyses of any of the designs yielded higher genetic gains (2% on average) with similar ranges (Figure 2d). Interestingly, the DR design presented, for no-spatial and spatial models, the highest PM_{40} values, even that this design has only 512 genotypes evaluated, in contrast with AD6.25 that has 960 test genotypes evaluated.

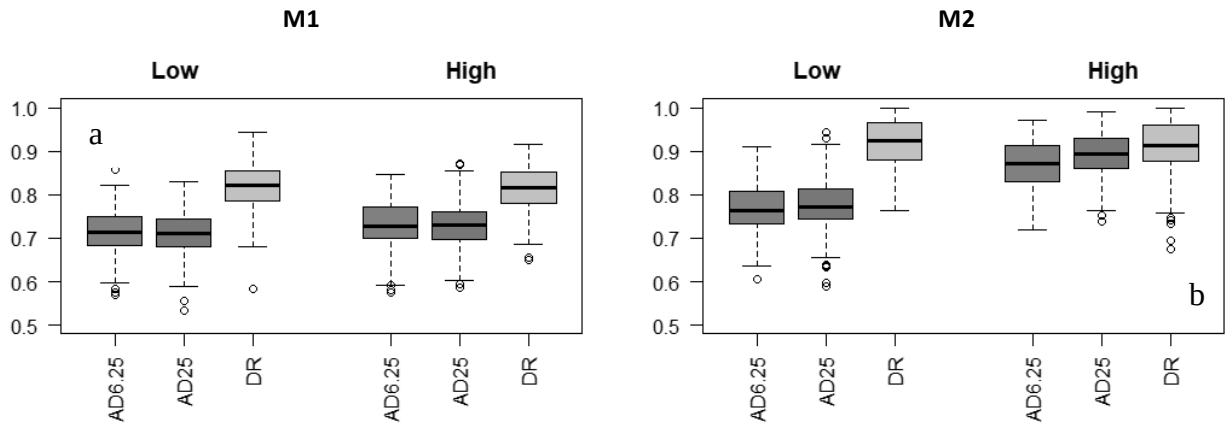

Figure 3. Box plots for SEF_{40} for no-spatial model (M1) (a) and spatial model AR(x)⊗AR(y) (M2) (b)for the sites without nugget for genotypic values simulation separated for low and high spatial correlations.

3.2 Breeding Values

For the breeding value simulations, the additive variance, σ^2_a, was closer to its true value for the scenario with high additive proportion (E1) in all designs (TABLE 3); however, an overestimation was noted for the scenario with low additive proportion (E2) in both the AD6.25 and AD25 designs with values close to 0.254 when its real value was 0.2; interestingly, no bias was noted for this component in the DR design. The analyses of the AD6.25 and AD25, with a model that does not separate dominance from epistasis, presented an overestimation for the σ^2_c that

is due to this confounding between the family and the genotypic effects (TABLE 3). The value for E1 was close to its expected value of $d^2 + i^2 = 0.10$, but for E2 this value was smaller (~0.25) than the true value of 0.30 (TABLE 4); hence, the difference appears to be contained as a bias in the additive variance. For the DR design, the partition of dominance and epistasis was possible, and their average estimated values were close to its true values for both E1 and E2 (TABLE 4). These results indicate that for higher proportions of non-additive variance with respect to additive variance it is difficult to properly partition the genetic components for the AD designs, but the DR design provides with reasonable and unbiased estimates. As with the genotypic value simulations, σ^2_b also presented values larger than its true value of 0, but these were no larger than 0.02 (TABLE 3). Unbiased estimates of spatial correlations (ρ_x, ρ_y) were found for the M3 model in all designs; however, an important underestimation of these correlations was reported by the model M2 on those sites with nugget (TABLE 3); a comparable finding that the one detected on genotypic value simulations (TABLE 1).

Table 3. Variance components, spatial correlations and heritability estimates for three designs using the three statistic models analyzed for sites with nugget for breeding value simulations for scenario E1 with high proportion of additive variance ($h^2 = 0.40$, $d^2 = 0.05$ and $i^2 = 0.05$) and scenario E2 with low proportion ($h^2 = 0.20$, $d^2 = 0.15$ and $i^2 = 0.15$).

Scenario E1									
Designs	**Models**	σ^2_b	σ^2_a	σ^2_f	σ^2_c	σ^2_e	σ^2_η	ρ_x	ρ_y
AD6.25	**M1**	0.020	0.407	-	0.101	0.474	-	-	-
	M2	0.010	0.412	-	0.166	0.428	-	0.306	0.301
	M3	0.004	0.408	-	0.097	0.265	0.231	0.501	0.495
AD25	**M1**	0.019	0.402	-	0.094	0.481	-	-	-
	M2	0.012	0.406	-	0.116	0.475	-	0.249	0.246
	M3	0.003	0.403	-	0.090	0.274	0.230	0.501	0.503
DR	**M1**	0.008	0.397	0.016	0.088	0.495	0.000	0.000	0.000
	M2	0.006	0.401	0.015	0.097	0.491	0.000	0.243	0.239
	M3	0.003	0.399	0.015	0.088	0.263	0.241	0.512	0.510
Scenario E2									
Designs	**Models**	σ^2_b	σ^2_a	σ^2_f	σ^2_c	σ^2_e	σ^2_η	ρ_x	ρ_y
AD6.25	**M1**	0.020	0.256	-	0.251	0.482	-	-	-
	M2	0.010	0.258	-	0.320	0.436	-	0.310	0.302

	M3	0.004	0.258	-	0.251	0.273	0.229	0.504	0.496
	M1	0.020	0.250		0.251	0.486	-	-	-
AD25	**M2**	0.013	0.250	-	0.279	0.478	-	0.247	0.252
	M3	0.004	0.250	-	0.253	0.269	0.237	0.504	0.503
	M1	0.009	0.209	0.037	0.258	0.495	0.000	0.000	0.000
DR	**M2**	0.007	0.208	0.037	0.268	0.492	0.000	0.241	0.241
	M3	0.004	0.208	0.037	0.259	0.277	0.230	0.501	0.495

† AD6.25 and AD25 represents augmented designs with 6.25% and 25% of replicates respectively and DR is the double-replication design. M1 is the no-spatial model, M2 is the spatial model AR(x)⊗AR(y) and M3 is the spatial model AR(x)⊗AR(y) $+\sigma_{\eta}^{2}$; σ^2_b, σ^2_a, σ^2_f, σ^2_c, σ^2_e and σ^2_η represents the variance components for block, individual (additive), family, clone, residual and nugget, and ρ_x and ρ_y are the spatial correlations for row and column.

In a similar way as with the genotypic value simulation, the broad-sense heritability (H_c^2) for the M2 model with the AD6.25 and AD25 was always higher than its true value, but for all other analyses its value was close to its true value of 0.5 (TABLE 4). As indicated earlier, overestimation of narrow-sense heritability was found for scenario E2 for both of the AD designs but not for the DR design.

Table 4. Summary statistics for the three designs using the three statistic models analyzed for sites with nugget for breeding value simulations for scenario E1 with high proportion of additive variance ($h^2 = 0.40$, $d^2 = 0.05$ and $i^2 = 0.05$) and scenario E2 with low proportion ($h^2 = 0.20$, $d^2 = 0.15$ and $i^2 = 0.15$). Numbers in bold corresponds to the best model for a given statistics within a design.

Scenario E1											
Designs	**Models**	**logREML**	**h^2**	**d^2**	**i^2**	**d^2+i^2**	**H^2_c**	**h^2_{PEV}**	**$CorP_a$**	**SEF_{40}**	**PM_{40}**
	M1	-413.477	0.403	-	-	0.102	0.505	0.569	0.748	73.9%	11.56
AD6.25	**M2**	-379.623	0.401	-	-	0.165	0.566	0.592	0.760	75.1%	**11.58**
	M3	**-372.732**	0.402	-	-	0.098	0.500	**0.593**	**0.762**	**75.3%**	**11.58**
	M1	-366.454	0.400	-	-	0.096	0.496	0.560	0.742	72.9%	11.47
AD25	**M2**	-327.722	0.399	-	-	0.116	0.515	0.583	0.754	74.1%	**11.49**
	M3	**-316.959**	0.400	-	-	0.092	0.491	**0.587**	**0.757**	**74.5%**	**11.49**
	M1	-386.214	0.392	0.062	0.04	0.105	0.497	0.607	0.788	77.7%	11.66
DR	**M2**	-341.244	0.392	0.061	0.05	0.114	0.506	0.630	0.801	79.1%	**11.69**
	M3	**-328.288**	0.392	0.060	0.04	0.105	0.496	**0.635**	**0.804**	**79.5%**	**11.69**

5

Scenario E2											
Designs	**Models**	**logREML**	**h^2**	**d^2**	**i^2**	**d^2+i^2**	**H^2_c**	**h^2_{PEV}**	**$CorP_c$**	**SEF_{40}**	**PM_{40}**
AD6.25	**M1**	-453.743	0.251	-	-	0.250	0.501	0.472	0.627	61.7%	11.31
	M2	-425.080	0.250	-	-	0.313	0.562	0.488	0.635	62.6%	11.32
	M3	**-418.964**	0.252	-	-	0.248	0.501	**0.491**	**0.636**	**62.8%**	**11.33**
AD25	**M1**	-399.221	0.246	-	-	0.250	0.496	0.455	0.616	60.6%	11.23
	M2	-364.351	0.243	-	-	0.274	0.517	0.470	0.626	61.5%	**11.25**
	M3	**-354.502**	0.245	-	-	0.250	0.495	**0.474**	**0.628**	**61.7%**	**11.25**
DR	**M1**	-413.458	0.205	0.148	0.14	0.294	0.499	0.405	0.648	63.5%	11.38
	M2	-370.898	0.203	0.147	0.15	0.302	0.506	0.416	0.656	64.3%	11.39
	M3	**-358.246**	0.203	0.147	0.14	0.293	0.497	**0.420**	**0.658**	**64.6%**	**11.40**

† AD6.25 and AD25 represents augmented designs with 6.25% and 25% of replicates respectively and DR is the double-replication design. M1 is the no-spatial model, M2 is the spatial model AR(x)⊗AR(y) and M3 is the spatial model AR(x)⊗AR(y) +σ^2_η; logREML is the log-likelihood REML value, h^2, d^2, i^2 and H^2_c represents the narrow-sense heritability, dominance ratio, epistatic ratio and broad-sense heritability. Also, h^2_{PEV} is the predictor-error-variance heritability, $CorP_a$ is the correlation between true and estimated breeding values, SEF_{40} is the selection efficiency of selecting the top 40 individuals based on breeding values, and PM_{40} is the genetic gain achieved by selecting the top 40 individuals.

For all the combinations of designs, models and scenarios, model M3 (spatial with nugget) always presented the best goodness-of-fit statistics of logREML, $CorP_a$ and SEF_{40}; however, differences were negligible with model M2, particularly in relation to $CorP_a$ and SEF_{40} (TABLE 4). It is important to note that $CorP_a$ values for model M1 (no spatial) are, on average across all combinations, only 1% lower than the values found for models M2 and M3. Hence, for DR design, even a no-spatial model can provide with reasonable estimates of breeding values. Also, note that the difference between the $CorP_a$ for the augmented designs AD6.25 and AD25 against the DR design are 2-4%, a small difference that contrasts with the 8-10% found for genotypic value estimations ($CorP_c$ from TABLE 2). In general, the use of DR designs results in an increase of ~3% in the correlation between true and predicted breeding values against the AD designs. Also,

for those simulations that include pedigree information, the incorporation of spatial correlations in the statistical analyses increases correlation marginally in ~2%. Again, a small difference compared to the ~10% that was found for the genotypic value estimations ($CorP_c$ from TABLE 2), reflecting the benefits of including pedigree information.

Similar trends of those found for $CorP_a$ were detected for SEF_{40} with values ranging between 72.9-79.5% and 60.6-64.6% for E1 and E2 scenarios, respectively. Hence, better results are obtained, as expected, for the case with larger additive variance. In addition, marginally better selection efficiency values were found for DR designs, where these benefits were more relevant in E2 than in E1. Finally, small differences were obtained with PM40, with similar trend of those found for genotypic values where, even with small number of genotypes tested, the DR design provided with marginally larger genetic gain achieved.

4 DISCUSSION

4.1 Reliability of Variance Component and Genetic Parameters

Genotypic value simulations showed that, as expected, model M2 provided with better variance component estimates for sites without nugget, and model M3 was better for sites with nugget, indicating that using the correct model for its corresponding site structure will lead to negligible bias on estimation of variance components. Model M2, in sites with nugget, for both, genotypic and breeding value simulations presented an important bias on the estimation of the variance components associated with the spatial components. Interestingly, model M3 was always adequate on sites with or without nugget, showing equivalent statistics with model M2 under the without nugget simulations.

The simulations for breeding value estimations (i.e., those that incorporate pedigree information) under both E1 and E2 scenarios provided with reasonable variance component estimates for models M2 and M3; however, important differences were noted between experiments. The analyses of the AD design showed that this design has some difficulty separating the non-additive from the additive components (TABLE 3), particularly when the additive component is relatively low (E2), producing upwardly biased heritabilities of ~0.25 instead of 0.20. Nevertheless, this bias was not almost completely absent when the additive component was relatively high (E1). These overestimations were not observed for the DR design analyses, where all genetic parameters, in both scenarios and for all models, were close to its true value, indicating that, for the conditions evaluated in this study, this is a better experimental design to separate additive, dominance and epistatic effects.

4.2 Incorporation of Spatial Correlations and Nugget Effect

For the analyses considered under sites without nugget, correlation between true and predicted genotypic values were all higher than 0.71, even for AD designs without spatial analyses. However, spatial analyses provided with an increase of this correlation to an average value of 0.85, with better results for models without nugget. The improvements on this correlation for the DR design were lower but still relevant with an increase from 0.82 to 0.88. These differences were much smaller in the case of breeding value simulations, with an increase of only ~0.02 for any scenario. This is probably the results of incorporating the pedigree information that allows for more accurate estimation of BLUP values. Interestingly, the effects of having large spatial components seems to be more relevant for AD designs, and once there is some replication, as with DR, these benefits are greatly diminished (Figure 1). Hence, this gain in precision on genetic value prediction is achieved by performing spatial analyses of field data, which agrees with the general recommendation of using spatial analysis for modelling the residual variance by Gilmour *et al*. (1997).

Interestingly, correlations between true and predicted values (and other goodness-of-fit statistics) reported almost null differences between models M2 and M3, on both sets of simulations with or without nugget. Therefore, it seems appropriate to always incorporate the nugget effect into the fitted linear model when spatial analyses are performed, but our study indicated that its beneficial effects are expected to be minimal. Similar results were indicated by GILMOUR *et al*. (1997) and MÜLLER *et al*. (2010) were the addition of the nugget effect on these studies resulted in better analyses in contrast to models without this effect.

4.3 Model Selection with Goodness-of-fit-Statistics

A goodness-of-fit statistics used to select a genetic linear model should identify the model that provides with the best correlation between true and predicted genotypic or breeding values. In this study, logREML resulted to be the most appropriate goodness-of-fit statistic to identify the adequate model for both genotypic and breeding value simulations under the array of conditions evaluated. This is reflected in the strong agreement between logREML and $CorP_c$ (or $CorP_a$) for sites with and without nugget (TABLES 2 and 4), justifying its use. The logREML as a goodness-of-fit statistics has been also widely used to select models in other genetic studies (COSTA e SILVA *et al.*, 2001; GEZAN *et al.*, 2006).

In contrast, H^2_{PEV} or h^2_{PEV}, a goodness-of-fit statistic often recommended to select the best models (CULLIS *et al.*, 2006; WELHAM *et al.*, 2014), appears to not completely agree with logREML, and therefore $CorP_c$ (or $CorP_a$) often selecting incorrectly model M2 for the genotypic simulations. However, for breeding value simulations this disagreement is not present, and all fitting statistics selected as best model M3. For the present study, it appears that this statistic was affected by the presence of the upward bias found on the estimation of the genotypic variance that is occurring in model M2, and therefore it should be used with caution, and further evaluations should be performed to determine the most adequate goodness-of-fit statistic for genetic analyses.

4.4 Augmented Design versus Double-Replicated Design

As expected, the use of DR designs, in comparison to any AD design, translated into better fittings. For example, producing an increase of the correlation between true and predicted values from 0.79 to 0.88 for genotypic values and from 0.69 to 0.73 for breeding values. An interesting result from this study was that the differences in goodness-of-fit statistics between DR and AD were lower in the simulations for breeding values, indicating that, for the conditions evaluated in

this study, AD designs were able to estimate breeding values almost as efficiently as replicated experiments. However, given the presence of confounding between family and genotypic effects in the AD design it might be useful to favor DR designs as these will be able to partition the additive, dominance and epistatic components for a given trait. This aspect is critical to assist breeders on implementing more efficient breeding strategies, where non-additive effects are relevant.

Under the same space and genetic material resources, one drawback of the DR design is the lower number of genotypes that can be evaluated. For example, the DR designs for genotypic experiments evaluated in our simulations a total of 512 test genotypes, in contrast to the AD6.25 that evaluated 960 genotypes (i.e., 87.5% more entries). This smaller set of genotypes translates into a reduction of selection intensity given that the pool of genotypes to make selection is smaller for DR than for AD designs. However, for the conditions evaluated in this study, the genetic gains achieved (in this case by selecting the top 40 individuals) based on their true genetic value (PM_{40}) still resulted in higher efficiencies for DR designs in genotypic or breeding values (TABLES 2 and 4) due to better identification of the top genotypes. This is a strong indication that the increase in replication at the cost of a reduced selection intensity, it is not greater than the benefits of increased precision of selection achieved by using double replication.

For all simulations evaluated in this study, the comparison of AD designs with 6.25% and 25% of control plots, according to the majority of the goodness-of-fit statistics, indicated that these two designs present minimal differences. Hence, AD designs with as little as 6.25% controls should be favored as these will require fewer control plots, and therefore increasing selection intensity (from 768 to 960 test genotypes). Burgueño *et al*. (2018) showed similar results for the proportion of control plots, where ~11% of check plots were adequate. Other authors suggest using between 10 and 25% of control plots (MARTIN *et al.*, 2006; MÜLLER *et al*., 2010).

4.5 Varying Design Parameters

The conditions evaluated in this study for genotypic simulations considered a broad-sense heritability of 0.5; however, this value will vary depending on the trait and field characteristics. It is possible to evaluate the effect of different levels of genetic control, replication and spatial correlation on the accuracy of genotypic value estimation mathematically. This was done by calculating the correlations between true and predicted genotypic values ($CorP_c$) as the square root of the genotypic mean heritability estimated with the expression: $H_c^2 = \sigma_g^2/\left(\sigma_g^2 + \sigma_e^2/[r(1+\rho/2)]\right)$, where r is the replication, ρ is the average spatial correlation for rows and columns, and the other are the variance components previously defined. Note that this expression translates the benefit of spatial correlation in terms of increased replication, and it can be written as $H_c^2 = H^2/(H^2 + (1-H^2)/[r(1+\rho/2)])$ when $\sigma_g^2 + \sigma_e^2 = 1$. Hence, for the model without spatial correlation (M1), ρ = 0, and therefore there are no benefits of the use of spatial correlation, and for a model with ρ = 1, there will be exactly doubling of the replication. Also, note that the nugget effect was not considered in this case, but it can be easily incorporated.

Based on the above expression different levels of broad-sense heritability (0.2 to 0.9) and replication (1, 2 and 3) were evaluated for models without and with spatial analyses (M1 and M2, respectively). These calculations are presented in Figure 4, and, as expected, the increase in the number of replications has a positive effect on the $CorP_c$, but this effect gets reduced once replications are greater than three, and its effect is more relevant under low heritability values. As this heritability increases, the differences between any of the conditions diminishes considerably, particularly on levels of heritability of 0.75 or higher. There is always an improvement on these correlations when model M2 is used instead of model M1. Again, a difference that gets smaller under higher heritability levels. A similar plot can be obtained for correlations for breeding value

estimations with similar trends, which is not presented here. In summary, this figure can be used to guide some decisions in terms of replications and desired levels of accuracy.

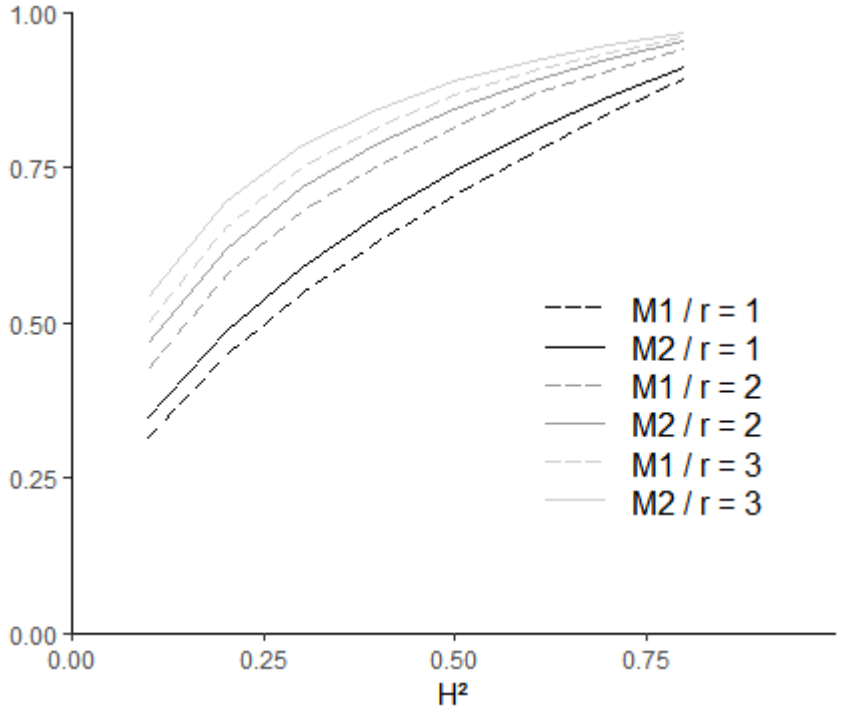

Figure 4. Correlation between true and predicted genotypic values ($CorP_c$) for different levels of broad-sense heritability (H^2) for no-spatial model (M1) and spatial model AR(x)⊗AR(y) (M2) on sites without nugget for varying replications (r).

4.6 Final Remarks

In this study, all spatial simulations and analyses were made considering the two-dimensional autoregressive error structure with and without nugget (AR(x)⊗AR(y) + σ_η^2 and AR(x)⊗AR(y), respectively) suggested by Cullis and Gleeson (1991). However, several other spatial error structures could have been considered (HU; SPILKE, 2009), and a different approach to model spatial variation, for examples based on cubic Splines, could also have been followed (VELAZCO *et. al*., 2017). Nevertheless, it is expected that similar trends will be found with any spatial approach, as long as it exploits information related to field correlations.

The interaction genotypes by environments (GE) was not the focus of this paper but some of the findings presented here can be easily extrapolated for multi-environmental trial (MET) analyses, particularly those with reference to accuracy achieved and effects of replication. For more details on this aspect we recommend revising literature on spatial analysis in MET (CULLIS *et al.*, 1998; SMITH *et al.*, 2006), and on effects of unreplicated designs in a MET (MOEHRING *et al.*, 2014).

The designs evaluated in the present study focused on two specific contrasting designs: AD and DR. However, the use of P-rep designs is also very common in many agronomical field experiments designs (CULLIS *et al.*, 2006). In these designs some test genotypes are replicated two or more times, and the rest are unreplicated. Here, replicated genotypes act as control plots helping to exploit and estimate spatial variability. Results from our current study can easily be extrapolated to P-rep designs as this design is intermediate between DR and AD designs, where for P-rep designs there will be two tiers of results: those that apply for the replicated genotypes and those for unreplicated genotypes.

Finally, it was shown in the simulations for breeding values presented here the tremendous benefit of considering the pedigree information to connect data, and therefore to improve the accuracy of breeding value estimation, as noted by the statistic $CorP_a$. This benefit was more relevant for the AD designs where, in some cases, results were almost equivalent to DR designs. Therefore, the use of a relationship matrix for estimation of additive effects is critical to maximize the information extracted from both of these designs. In a similar way, the use of molecular information (e.g., SNP markers) to obtain genomic relationship matrices (VANRADEN, 2008) could increase genetic gains due to improved accuracy on the determination of the relationship between individuals. Even further, the use of this information on relatedness can be incorporated not only for the analysis stage but also for the randomization of the field experiments, as illustrated by Mramba *et al.* (2018).

5 CONCLUSIONS

Moderate to large values of field spatial correlation, in at least one coordinate, increases the accuracy of the spatial analyses for any design. It was found that, in all simulation conditions evaluated, the use of spatial analysis provided with superior results than a no-spatial model. Even when there is no nugget on the experimental field, its incorporation on the model fit brings better, or at least equivalent, accuracies in most situations.

The incorporation of the pedigree information by obtaining a relationship matrix on the estimation of breeding values resulted on important gains in accuracy with good correlations between true and predicted breeding value. Interesting, the benefits of the spatial analyses were less relevant once pedigree information was incorporated into the model.

Based on the evaluated goodness-of-fit statistics it was determined, for the simulations performed, that the use of double replication (DR) designs presents better performance than augmented designs (AD). The use of AD with 6.25% control plots showed similar results than the AD designs with 25% of control plots, but the former has the benefit of fewer control plots; therefore, allowing for an increase number of test genotypes translating into better selection intensity.

In summary, it is always recommended the use replication, whenever possible. The simulations presented in this study indicated that a replication of only two plots per genotype provides with very good estimations of genotypic and breeding values. However, unreplicated trials, also provide with reasonable estimates of these genetic values, but these are particularly good under spatial analyses and when they incorporate pedigree information.

REFERENCES

AKAIKE, H. Maximum likelihood identification of Gaussian autoregressive moving average models. **Biometrika,** v. 60, n. 2, p. 255-265, 1974.

ALSPACH, P. A.; ORAGUZIE, N. C. Estimation of genetic parameters of apple (*Malus domestica*) fruit quality from open-pollinated families. **New Zealand Journal of Crop and Horticultural Science**, v. 30, n. 4, p. 219-228, 2002.

AMADEU, R. R. *et al.* AGHmatrix: R package to construct relationship matrices for autotetraploid and diploid species: A blueberry example. **The Plant Genome**. v. 9, n. 3, p. 1-10, 2016.

BURGUEÑO, J. 2018. Spatial analysis of field experiments. In: B. Glaz, K. M. Yeater, editors, **Applied Statistics in Agricultural, Biological, and Environmental Sciences**, ASA, CSSA, and SSSA, Madison, WI. p. 319-344.

BURGUEÑO, J. *et al*. 2018. Augmented designs-experimental designs in which all treatments are not replicated. In: B. Glaz, K. M. Yeater, editors, **Applied Statistics in Agricultural, Biological, and Environmental Sciences**, ASA, CSSA, and SSSA, Madison, WI. p. 345-370. doi:10.2134/appliedstatistics.2016.0005

BUTLER, D. G. *et al*. 2009. **ASReml-R Reference Manual**. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane, Australia.

COCHRAN, W. G.; COX, G. M. 1957. **Experimental Designs**. 2nd ed. John Wiley & Sons. Oxford, England

COSTA e SILVA, J. *et al*. Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual. **Canadian Journal of Forest Research**, v. 31, n. 11, p. 1887-1893, 2001.

CRESSIE, N. A.; NOEL, A. 1993. **Statistics for Spatial Data – Revised Edition**. John Wiley & Sons, Inc. New York, New York

CULLIS, B. R.; GLEESON, A. C. Spatial analysis of field experiments—An extension to two dimensions. **Biometrics**, v. 47, n. 4, p. 1449–1460, 1991.

CULLIS, B. *et al*. Spatial analysis of multi-environment early generation variety trials. **Biometrics**, v. 54, n. 1, p. 1-18, 1998.

CULLIS B. R. *et al.* On the design of early generation variety trials with correlated data. **Journal of Agricultural, Biological and Environmental Statistics**, v. 11, n. 4, p. 381–393, 2006.

DAETWYLER, H. D. *et al*. Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. **Animal Production Science**, v. 50, n. 12, p. 1004-1010, 2010.

DALE, M. R.; FORTIN, M. J. 2014. **Spatial Analysis: A Guide for Ecologists**. University Press, Cambridge, UK.

DENIS, M.; BOUVET, J. M. Efficiency of genomic selection with models including dominance effect in the context of Eucalyptus breeding. **Tree Genetics & Genomes**, v. 9, n. 1, p. 37-51, 2013.

DUARTE, J. B.; VENCOVSKY, R. Spatial statistical analysis and selection of genotypes in plant breeding. **Pesquisa Agropecuária Brasileira**, v. 40, n. 2, p. 107-114, 2005.

DUTKOWSKI, G. W. *et al.* Spatial analysis methods for forest genetic trials. **Canadian Journal of Forest Research**, v. 32, n. 12, p. 2201-2214, 2002.

FEDERER, W. T. 1955. **Experimental Design - Theory and Application**. Macmillan, New York

FEDERER, W. T.. Augmented (or hoonuiaku) designs. **Hawaiian Planters' Record**, v. 55, p. 191-208, 1956

FEDERER, W. *et al.* Some augmented row-column designs. **Biometrics**, v. 31, n. 2, p.361-373, 1975.

GEZAN, S. A. *et al.* Post hoc blocking to improve heritability and precision of best linear unbiased genetic predictions. **Canadian Journal of Forest Research**, v. 36, n. 9, p. 2141-2147, 2006.

GILMOUR, A. R. *et al.* Accounting for natural and extraneous variation in the analysis of field experiments. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 2, n. 3, p. 269-293, 1997.

KEMPTON, R. A. 1984. The design and analysis of unreplicated field trials. **Vortraege fuer Pflanzenzuechtung (Germany).** ISSN: 0723-7812

KEMPTON R.A; GLEESON, A. C. 1997. Unreplicated trials. In: Kempton R.A., Fox P.N., Cerezo M. (eds) **Statistical Methods for Plant Variety Evaluation**. Plant Breeding Series 3 Statistical Methods for Plant Variety Evaluation. Springer, Dordrecht.

HESLOT, N. *et al*. Genomic selection in plant breeding: a comparison of models. **Crop Science**, v. 52, n. 1, p. 146-160, 2012.

HU, X.; SPILKE, J. Comparison of various spatial models for the analysis of cultivar trials. **New Zealand Journal of Agricultural Research**, v. 52, n. 3, p. 277-287, 2009.

LITTELL, R. C. *et al.* 2006. **SAS® for Mixed Models**. SAS Institute Inc. Cary, North Carolina

LUND, M. S. *et al.* A common reference population from four European Holstein populations increases reliability of genomic predictions. **Genetics Selection Evolution**, v. 43, n. 1, p. 43, 2011.

MARTIN, R. J. *et al*. Chan. Some results on the design of field experiments for comparing unreplicated treatments. **Journal of agricultural, biological, and environmental statistics**, v. 11, n. 4, p. 394, 2006

MOEHRING, J. *et al*. Efficiency of augmented p-rep designs in multi-environmental trials. **Theoretical and Applied Genetics**, v. 127, n. 5, p. 1049-1060, 2014.

MRAMBA, L. K. *et al.* Generating improved experimental designs with spatially and genetically correlated observations using mixed models. **Agronomy**, v. 8, n. 4, p. 40, 2018.

MRODE R. A. 1996. **Linear Models for the Prediction of Animal Breeding Values**. 1st ed. CABI. Wallingford, Oxon, UK

MÜLLER, B. U. *et al*. Arrangement of check plots in augmented block designs when spatial analysis is used. **Plant Breeding**, v. 129, n. 6, p. 581-589, 2010

PAYNE, R. W. New and traditional methods for the analysis of unreplicated experiments. **Crop Science**, v. 46, n. 6, p. 2476-2481, 2006.

PIEPHO, H. *et al.* Problems in parameter estimation for power and AR(1) models of spatial correlation in designed field experiments. **Communications in Biometry & Crop Science**, v. 10, n. 1, p. 3-16, 2015.

R Core Team. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria. 2018.

SCHAEFFER L.R. *et al*. The inverse of the gametic relationship matrix. **Journal of Dairy Science**, v. 72, n. 5, p. 1266–1272, 1989.

SCHWARZ, G. Estimating the dimension of a model. **Annals of Statistics**, v. 6, n. 2, p. 461–464, 1978.

SMITH, A. B. *et al*. The design and analysis of multi-phase plant breeding experiments. **The Journal of Agricultural Science**, v. 144, n. 5, p. 393-409, 2006.

STEEL, R. G.; J. H. TORRIE. 1980. **Principles and Procedures of Statistics: A Biometrical Approach**. McGraw-Hill, New York, New York

VANRADEN, P. M. Efficient Methods to compute genomic predictions. **Journal of Dairy Science**, v. 91, n. 11, p. 4414-4423, 2008.

VELAZCO, J. G. *et al.* Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model. **Theoretical and Applied Genetics**, v. 130, n. 7, p. 1375–1392, 2017.

WELHAM, S. J. *et al*. A comparison of analysis methods for late-stage variety evaluation trials. **Australian & New Zealand Journal of Statistics**, v. 52, n. 2, p. 125-149, 2010.

WELHAM, S. *et al*. **Statistical Methods in Biology: Design and Analysis of Experiments and Regression.** New York: Chapman and Hall/CRC. 2014.

WILLIAMS E. R. *et al.* Augmented p-rep designs. **Biometrical Journal**, v. 53, n. 1, p. 19-27, 2011.

YOUNG, L. J.; YOUNG, J. H. **Statistical Ecology: A Population Perspective**. Kluwer Academic Publishers. Boston, Massachusetts. 1998.

ZIMMERMAN, D. L.; HARVILLE, D. A. A random field approach to the analysis of field-plot experiments and other spatial experiments. **Biometrics**, n. 47, n. 1, p. 223-239, 1991.

ATTACHMENTS

M \ F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21
M1	x	x	x																		
M2		x	x	x																	
M3			x	x	x																
M4				x	x	x															
M5					x	x	x														
M6						x	x	x													
M7							x	x	x												
M8								x	x	x											
M9									x	x	x										
M10										x	x	x									
M11											x	x	x								
M12												x	x	x							
M13													x	x	x						
M14														x	x	x					
M15															x	x	x				
M16																x	x	x			
M17																	x	x	x		
M18																		x	x	x	
M19																			x	x	x
M20	x																			x	x
M21	x	x	x																		x

Figure S1. Crossing arrangement for the breeding value simulations based on a circular diallel with 42 parents for a total of 64 families.

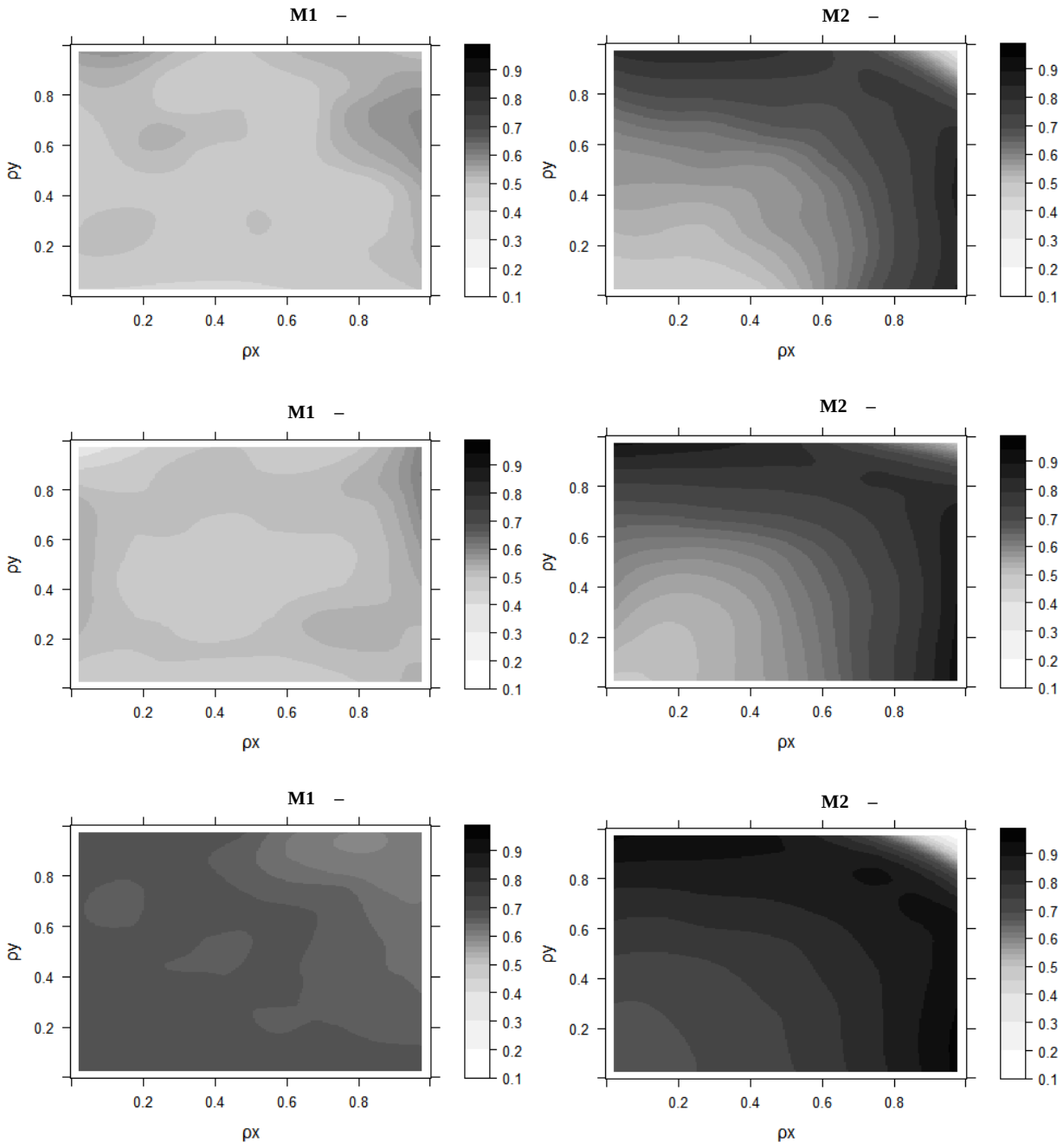

Figure S2. Contour plots for H^2_{PEV} obtained by using the true row and column spatial correlations for the designs AD6.25, AD25 and DR fitted with no-spatial model (M1) and spatial model AR(x)⊗AR(y) (M2) for the sites without nugget for genotypic values simulation.