

Increasing Precision of Even Otherwise Well-Run Trials by Capturing Heterogeneity of Plot Error Variances

M. Singh · M. van Ginkel · A. Sarker · R. S. Malhotra · M. Imtiaz · S. Kumar

Received: 11 January 2012 / Accepted: 14 June 2012 / Published online: 6 July 2012
© NAAS (National Academy of Agricultural Sciences) 2012

Abstract Application of an experimental design based on blocking of homogeneous experimental units is an objective approach to reduce experimental error. Often, the experimental design and analysis efforts are considered to be executed satisfactory as long as the coefficient of variation is estimated to be below 10 %, say for yield, as a generally accepted guide for a well-run trial. Most of the statistical analyses of data are based on the assumption of a homogeneous variance of plot errors. In field plot experimentation, the question of heterogeneity of error variances has been addressed here. This study introspects a set of four such supposedly well-run trials in lentil and chickpea conducted in lattice designs. The presence of, in fact, heterogeneous error variances exhibits a distribution of coefficient of variation as opposed to a single index to measure heterogeneity of a field. This further suggests that it is more realistic to view the distribution of the heritability of a trait and its genetic advance due to selection, when assessing them, even in a single field. Objectives of the study were to examine the possibility of heterogeneous error variances, identify the associated experimental units and estimate the predicted means of the genotypes and gain due to selection. We show that accounting heterogeneous variances allows a significant increase in the efficiency of genotypic comparisons and the power of genotypic discrimination, higher heritability and genetic advance. Thus, it is likely to help shorten the breeding cycle for an expected genetic gain, and is in principle relevant to all experiments that use replications, involving crops or not.

Keywords Heterogeneity of experimental error variances · Experimental design · CV · Efficiency over RCB · Spatial analysis

Introduction

In field experimentation, the evaluation of treatment effects, for example, in crop variety trials, the effects of genotypes of a crop, is carried out against the experimental error variability, which arises due to a contribution of several factors such as local soil fertility, moisture profile, microbial activities, landscape variations, etc. Various standard approaches are followed to control the field variability by means of experimental design and statistical

analyses. Commonly-used experimental designs, often practiced by the crop improvement programmes, are based on anticipated, though often not confirmed, homogeneous blocks of suitable sizes (incomplete blocks) [2, 4, 6, 10]. Coupled with the incomplete block design, the possibility of accounting for any possible correlation between the plot errors, particularly auto-correlations along rows and columns, has been found useful in enhancing the efficiency of crop variety trials [3, 5, 12, 13]. These experimental designs and analysis led to higher efficiency for genotypic comparisons and genetic gain over analyses based on complete blocks. Similar advantages were found in chickpea trials [8, 12]. These analysis approaches were based on the assumption of homogeneous error variances. Singh et al. [14] argued that the probably more common situation is that field heterogeneity is present amongst plot

M. Singh (✉) · M. van Ginkel · A. Sarker · R. S. Malhotra · M. Imtiaz · S. Kumar
International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syria
e-mail: m.singh@cgiar.org

errors with different (heterogeneous) error variances. These errors need not have any obvious structured spatial pattern. Such skewed situations arise in experimental stations or on farms, where the residual effects of the crop genotypes may last for more than two crop seasons, making the inter-experimental cover-cropping less effective, or the presence of pests and diseases might create the heterogeneity.

The heterogeneity of a given field is often expressed in terms of the coefficient of variation (CV, in %), where the variation is due to residuals after accounting for any systematic factors such as block differences. Those field trials, which have CV less than 10 % are generally considered well-run trials, the analysis of which cannot be further improved. The findings are normally accepted, including in ISI peer-reviewed journal articles, as long as the CV is less than 10 % as it appears that no substantial precision can be gained by further exploring the plot residuals. Since there is no such theoretical limit on CV arising from the experimental plot errors, the aim of this study is to show that statistical analyses of such supposedly well-run trials can, in fact, be further improved and sharpened by identifying and partitioning such heterogeneities. This study also raises another basic question on the limitation to a single index in terms of CV for describing the heterogeneity of the entire field of multiple plots. The concept of a single homogeneous CV for all plots is based on the assumption of an equal error variance over all plots. It is argued that since plot error variances are most likely to be heterogeneous, field heterogeneity will be better described by calculation of CVs over individual plots. Its distribution summary could include an average of the CVs over the plots, with minimum, maximum, or various quantiles. We address the heterogeneity issue in trials, which are otherwise accepted as well-run and precise in practice, and illustrate the procedure, while discussing the case using legume crop trials. Crop variety trials are generally conducted as multi-environment trials where plot error variances are expected to vary with the environment as per the response scale in those environments. In these cases, a homogeneous error variance is still assumed over all experimental plots in a given environment. However, estimating heterogeneous variances in a single experiment is limited and is being addressed here.

Two legume trials each in lentil and chickpea, where the CV under the randomised complete block (RCB) analysis model was found to be less than 10 % for each trial (i.e. a 'well-run' experiment), were studied with the objectives to: (1) explore the un-assumed presence of heterogeneous error variances, (2) identify the groups of experimental units where error variances vary with the group, (3) estimate the standard error of differences of genotype effects and compare the efficiency of the design-analysis methods over the RCB analysis and (4) estimate gain due to

selection. Computational details are presented for one of these trials.

Materials and Methods

We consider a set of four trials selected to cover a range of genetic material, comprising F_7 and F_8 generations in lentil, genotypes suited for winter and spring planting in chickpea in three locations, Tel Hadya and Breda in Syria and Terbol in Lebanon, over the years 1998, 2002, 2004 and 2009. The statistical models for analysing data from RCB design and lattice design will be referred to as RCB model and lattice model, respectively. These trials had CV values of less than 10 % when analysed by means of the RCB model. The CV values were even lower under the lattice model.

Lentil Trials

Lentil-Yield Trial (YT)

Sixteen lentil genotypes were evaluated in a yield trial (F_7 generation) conducted in a triple lattice at Tel Hadya, Syria, in 2004 in a 3×16 rectangular layout. Tel Hadya is located at $36^{\circ}01'N$, $36^{\circ}56'E$ and at an altitude of 284 masl with a long-term average annual rainfall of 334 mm. The harvested plot size was 3 m^2 ($2 \text{ m} \times 5 \text{ rows} \times 0.3 \text{ m}$ inter-row distance), and the CV for seed yield was 6.4 % when analysed by means of either the lattice model or the RCB model. We refer to this experiment as Lentil-YT.

Lentil-Advanced Yield Trial (AYT)

An AYT was conducted in 2002 at Breda in Syria ($35^{\circ}56'N$, $37^{\circ}10'E$, at an elevation of 300 masl, long-term average annual rainfall of 266 mm) to evaluate 25 lentil genotypes of an F_8 generation in a triple lattice in a 9×10 rectangular layout with buffer plots. The harvested plot size was 4.8 m^2 ($2 \text{ m} \times 8 \text{ rows} \times 0.3 \text{ m}$ inter-row distance). The CV was 9.4 and 9.9 % under lattice and RCB models, respectively. We refer to this trial as Lentil-AYT.

Chickpea Trials

Chickpea-Winter Planting (WP)

An experiment was conducted to evaluate 49 genotypes of winter habit chickpeas in a simple lattice on a 4×25 (rows \times column) layout in 1998 in the winter season at Terbol ($33^{\circ}49'N$, $35^{\circ}59'E$), in the Beqaa valley in Lebanon, which has an elevation of 950 masl and a long-term average annual rainfall of 515 mm. The harvested plot size

was 2.45 m² (4 m × 4 rows × 0.35 m inter-row distance: only middle 2 rows of 3.5 m harvested). The CV for grain yield was 5.8 and 6.5 % under lattice and RCB models, respectively. We refer to this trial as Chickpea—WP.

Chickpea-Spring Planting (SP)

A final experiment was conducted at Tel Hadya during the 2009 winter season to evaluate 36 spring habit genotypes of chickpea in a simple lattice. The harvested plot size was 2.45 m² (4 m × 4 rows × 0.35 m inter-row distance). The CV for yield was 7.8 % under the lattice model and 9.3 % under the RCB model. We refer to this trial as Chickpea—SP.

Statistical Method

Here, we first describe the method used by Singh et al. [14], and then illustrate its application with an example of a Lentil-YT with 16 genotypes.

Consider an experimental design used for generating the response (data) and a model for the data analysis. The data analysis model might have been selected out of a set of candidate models using a certain criterion e.g. an Akaike information criterion (AIC; [1]). Then, plot residuals are obtained by fitting the selected model.

Step 1 (K-Means Clustering)

Based on the data as the squared residuals, the experimental plots are clustered by means of K-means clustering and the criterion which maximizes between group sum of squares. Since variance components are based on the squares of residuals, the clustering is based on squared residuals. The K-means clustering method requires the number of groups or clusters to be set a priori. We varied the number of groups from $k = 2, \dots, 10$. Then, the models, in terms of block effects and any other spatial correlation structure for plot errors, were fitted, while allowing different error variances varying with the cluster/group for each of $k = 2, \dots, 10$. For a given k , the error variance for the plots of group i and the number of plots in it are denoted by σ_i^2 and n_i , respectively, where $i = 1, 2, \dots, k$. For example, if the experimental units were grouped into 3 clusters ($k = 3$), then the fitted model will produce estimates of the 3 error variances: σ_1^2 on n_1 plots, σ_2^2 on n_2 plots and σ_3^2 on n_3 plots. The model fitting is assessed in terms of maximum likelihood of the data for the fitted model. Genstat software [11] was used for computation in this study and expresses the maximum likelihood as restricted maximum likelihood (REML) [9]. The REML value is displayed as a quantity called *deviance*, which is defined as ‘minus twice the logarithm of REML value’, ignoring a constant, which depends on the fixed terms in the model. It

also produces a residual degree of freedom, which varies with the number of error variance components associated with the number of clusters. By means of a clustering method on squared residuals, clusters/groups of plots can be determined. Furthermore, the plots with similar squared residuals would be within a cluster; the error variances arising from different groups of plots could be expected to be different. Two questions arise: how many (heterogeneous) groups are present within plot errors, and whether the error variances are heterogeneous for a selected group of the clusters, that is, for a pre-specified value of k .

Suppose that we want to compare two groups of clusters of the experimental plots, say one group has J clusters (Group J), i.e. with J heterogeneous variances, and another $J'(>J)$ clusters (Group J'). If the Group J' is nested within Group J , then the change in the *deviance* can be used as a Chi-square test with $J'-J$ degrees of freedom (d. f.) to test the hypothesis that the error variances in the nested group (Group J') of units are the same as those of Group J . However, in general, one group of clusters with J' variances is not necessarily nested within the other group of J variances, in which situation, the model selection can be carried out using a criterion such as AIC (Akaike Information Criterion). AIC has been written in terms of the deviance, $AICD = deviance + 2 \times \text{number of variance-covariance parameters}$ [13, 17].

Then, based on the AICD values for the model fitted using a constant variance ($k = 1$, say) and heterogeneous variances for $k = 2, \dots, 10$, select that value of k , say k^* , for which the AICD is lowest. Suppose that based on the selected group of k^* clusters of plots the estimated variance components are denoted by $\hat{\sigma}_i^2 (i = 1, 2, \dots, k^*)$. This Step 1 does not indicate that the variances are all different or whether there is no scope to merge the clusters with variances close by. Then follows Step 2, checking whether the two clusters with the closest variance estimates could be merged and reduced to a new group of clusters, resulting in a reduction down to just one cluster.

Step 2

Let D_k be the deviance for the model fitted with k^* groups selected by the AICD criterion. Then, merge the two clusters, which have the nearest error variance estimates. In this case, the group with larger number of clusters is nested within the group with the merged smaller number of clusters. The model is fitted again, and error variances and deviance, D_{k^*-1} , are computed. If $D_{k^*-1} - D_{k^*}$ is greater than the critical values for Chi-square at 1 d. f. at a chosen significance level, then the merging is not accepted and the k^* heterogeneous groups are considered final and the process is stopped. Otherwise, the two clusters are merged into one, and the model is fitted with the newly formed group of

$k^* - 1$ clusters and the associated AICD value is computed. With this group of $k^* - 1$ clusters, repeat the above step. If merging continues, one needs to proceed until all the units are in one group, which is the (relatively rare) case of homogeneous error variances.

Step 3

Having decided on the clusters of heterogeneous variances, the model is then used to compare genotypes effects and the efficiency of the design. With multiple error variances arising in a single experiment, there will be several CVs based on these variances. The field heterogeneity can then be expressed in terms of the distribution of the CVs, or at least in terms of minimum, maximum, and mean on a plot basis. The mean will be the weighted mean of the CVs, with weights equal to the number of plots within the clusters.

Furthermore, the heritability and genetic advance (gain) due to selection were estimated by fitting the model where the genotypic effects are treated as independent random variables. We obtain estimates of the genotypic variance component and of the heterogeneous error variances, which are likely to differ from that of the model when genotype effects are assumed fixed. Heritability estimates vary with the environmental variance. Thus, for a given trait, if σ_g^2 stands for genotypic variance component and σ_i^2 the experimental error (environmental variance for a single trial), for the i -th cluster of plots, the heritability on a plot basis is given by $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_i^2)$ ($i = 1, 2, \dots, k$). Considering a 20 % selection intensity, the genetic gain, GA (20 %), will be given by

$$\%GA(20\%) = 100C(\sigma_g^2/\bar{Y}) / (\sigma_g^2 + \sigma_i^2/r)^{1/2}$$

where for a general $p =$ intensity of selection ($0 < p < 1$), $C = \frac{1}{p\sqrt{2\pi}} e^{-z_p^2/2}$, the truncation point z_p in the standard normal distribution is given by the equation $\int_{z_p}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - p$, \bar{Y} is the trial or location mean and r is the number of replications. For $p = 20 \%$, $C = 1.4$ [7].

An Illustration

Consider the dataset for Lentil-YT. In an earlier analysis, the best model identified, under the assumption of homogeneous error variances, was ‘randomized complete blocks with a separable first order autoregressive structure along rows and along columns’ (RCBArAr). Further details on such models’ descriptions and notations are given in Singh et al. [13]. We attempted fitting heterogeneous error structures based on the clusters obtained from the residuals under RCBArAr, but convergence of the restricted maximum likelihood procedure did not take place. Therefore,

we used the residuals under the lattice design and analysis model with constant variance. The K-means clustering approach, which maximizes the between group sum of squares, was applied on the squared residuals. Clusters of the experimental units were obtained for the groups of $k = 2, \dots, 10$ clusters. In Table 1, row $k = 1$ stands for no grouping having occurred, taking all the units as a single group. For each of these groups of clusters, the mixed linear model was fitted to the data by means of the REML procedure of Genstat. The model was described in terms of random effects for replications, blocks within replications and plot errors having variances varying with the clusters of a given group $k = 1, \dots, 10$. Some of the key directives in fitting the mixed linear models in this case were:

```
VCOMPONENTS[Fixed = Geno]Rep + Rep.Blk
+ f.Rows.Cols; CONSTRAINTS = POSITIVE
VSTRUCTURE[TERM = f.Rows.Cols] MODEL = diag;
Factor = fREML[PRINT = m, c, w, mean, d] Yield
```

where Rep, Blk, Geno, f, Rows, Cols and Yield stand for the factors/variante representing plot-wise assignment to replication, blocks within replications, genotypes, grouping factor from the cluster analysis, row and column position on the layout, and yield response, respectively. After fitting these models, we obtained the ‘deviance’, reflecting the departure of the model from the data (Table 1).

We note that in Table 1, using a $k = 3$ cluster group results in the lowest AICD value. Thus, there could be three heterogeneous groups. The estimates of the variances for the three groups were obtained as follows: $\hat{\sigma}_1^2 = 5443$ on 28 plots/experimental units, $\hat{\sigma}_2^2 = 8018$ on 11 plots and $\hat{\sigma}_3^2 = 23247$ on 9 plots. Of these, estimates of the first two variances σ_1^2 and σ_2^2 are closest as their difference is the

Table 1 The residual degrees of freedom, number of random terms, deviance and Akaike Information Criterion expressed as deviance (AICD) when variances were allowed to vary over various clusters

No. of clusters (k)	Residual d.f.	Random terms	Deviance	AICD ^a
1	29	3	341.8	345.8
2	27	5	327.6	335.6
3	26	6	316.1	326.1
4	25	7	315.8	327.8
5	24	8	312.6	326.6
6	23	9	313.7	329.7
7	22	10	312.8	330.8
8	21	11	312.5	332.5
9	20	12	312.2	334.2
10	19	13	312.6	336.6

^a AICD = deviance + 2 × number of variance–covariance parameters where the deviance = $-2 \ln(\text{REML value})$, ignoring a constant, which depends on the fixed terms in the model

smallest of all the paired differences. They were, therefore, merged to form a total of just two groups with $28 + 11 = 39$ and 9 plots. Here, the three cluster group is nested within the two cluster group. The mixed model with these two groups was fitted, and the deviance was calculated as $D_2 = 327.6$ (d.f. = 27).

When compared with deviance D_3 (316.1, d.f. = 26, Table 1) for the group with three clusters, we find a change in the deviance, $D_2 - D_3 = 327.6 - 316.1 = 11.5$, which, compared against the Chi-square on $27 - 26 = 1$ d.f., has a p value of 6.81×10^{-04} , which is less than the 5 % level of statistical significance. Thus, the two clusters of the three cluster group should not be merged. Therefore, we conclude that the plot error variances fall into three heterogeneous groups, with variance estimates as above. The positions of the plots with the three cluster group are shown in the layout, as in Table 2.

Using this heterogeneous structure, with factor ‘f’ standing for the three cluster groups in the above directives, one can obtain the best linear unbiased estimates (BLUEs) of the genotype means and an estimated average standard error of the differences of the genotypes effects. In order to compute the heritability and genetic advance due to selection, the first directive in the above codes can be modified as:

```
VCOMPONENTS Rep + Rep.Blk + Geno
+ f.Rows.Cols; CONSTRAINTS
= POSITIVE
```

The other two lines of codes that follow remain unchanged. After running Genstat with these directives, one would arrive at estimates of the variance components due to genotypes (σ_g^2) and due to heterogeneous error variances σ_k^2 ($k = 1, 2, 3$) that are likely to be different from those when genotype effects were assumed fixed. The Genstat codes required to carry out the analysis presented here are available from the authors on request.

Results

CV, Efficiency and Genotypic Significance

Following the method described above, clusters of plots associated with heterogeneous variances were identified, and estimates of their variances, mean CVs per plot, average SED (standard error of differences of genotype effects) and efficiencies of the genotypic comparison were computed. The statistical test of significance of genotypic effect (based the hypothesis of equality of genotype means) is given in terms of a Chi-square distribution based Wald statistics in Genstat [11] and the associated p values for the

Table 2 Field layout of the Lentil-PYT and the plot positions of the three (1–3) heterogeneous error variances (1:28 plots with $\hat{\sigma}_1^2 = 5443$; 2:11 plots with $\hat{\sigma}_2^2 = 8018$; and 3:9 plots with $\hat{\sigma}_3^2 = 23247$)

Field layout (genotypes in cells)			
Rows:	1	2	3
Columns			
1	1	10	13
2	16	13	15
3	2	14	9
4	10	8	2
5	5	16	14
6	7	7	7
7	4	15	1
8	13	3	12
9	9	1	8
10	3	5	16
11	12	6	5
12	8	9	11
13	6	2	10
14	15	4	3
15	11	11	6
16	14	12	4

Field layout (variance groups in cells)			
Rows:	1	2	3
Columns			
1	1	1	3
2	1	3	1
3	2	3	1
4	1	2	1
5	1	1	3
6	2	1	2
7	1	2	3
8	1	2	1
9	1	3	3
10	1	1	1
11	1	2	1
12	1	1	1
13	1	2	1
14	2	2	3
15	1	1	2
16	1	1	3

test of significance are given in Table 3. Table 3 also shows results for RCB and lattice models for comparison.

Lentil-YT

In this trial, the lattice design did not show any improvement over RCB. However, a spatial model with first order

Table 3 Summary statistics on significance of genotypic differences, CV % and efficiency of design–analysis model duo

Design/model	Group number	No. of plots	σ_e^2	CV%	SED (standard error of difference)	Efficiency (%)	Wald statistic	<i>p</i> -value (Chi-square)
Lentil-YT: Tel Hadya, 2004; 16 genotypes, 3 replications. Trial mean = 1427 kg/ha								
RCB		48	8335	6.40	74.55		172	<0.001
Lattice		48	8335	6.40	74.55	100	172	<0.001
RCBArAr ^a model		48	4765	4.80	49.65	225	912.5	<0.001
Lattice design and heterogeneous plot errors					32.95	512	1595	<0.001
	1	28	544.3	1.63				
	2	11	8018	6.27				
	3	9	23247	10.68				
			Mean (per plot)	4.40				
Lentil-AYT: Breda, 2002; 25 genotypes, 3 replications. Trial mean = 1771 kg/ha								
RCB		75	30745	9.90	145.9		69.48	<0.001
Lattice		75	27935	9.44	144.1	103	73.81	<0.001
RCBAr ^b		75	32139	10.12	130.3	125	87.39	0.006
RCBAr ^b and heterogeneous plot errors					91.92	252	169.7	<0.001
	1	8	124614	19.93				
	2	67	16823	7.32				
			Mean (per plot)	8.67				
Chickpea-WP: Terbol, 1998; 49 genotypes, 2 replications. Trial mean = 2764 kg/ha								
RCB			32659	6.54	180.7	100	282.5	<0.001
Lattice			25258	5.75	169.1	114	334	<0.001
Lattice design and heterogeneous plot errors					142.8	160	1368	<0.001
	1	28	1225	1.27				
	2	8	89593	10.83				
	3	28	9807	3.58				
	4	20	40676	7.30				
	5	14	24444	5.66				
			Mean (per plot)	4.57				
Chickpea-SP: Tel Hadya, 2009; 36 genotypes, 2 replications. Trial mean = 1069 kg/ha								
RCB			9817	9.27	99.08		364.32	<0.001
Lattice			7023	7.84	90.4	120	461	<0.001
Lattice design and heterogeneous plot errors					77.16	165	3354	<0.001
	1	34	349.1	1.75				
	2	6	39310	18.55				
	3	22	10842	9.74				
	4	10	4540	6.30				
			Mean (per plot)	6.22				

^a RCBArAr: randomized complete blocks with a separable first order autoregressive structure along rows and columns

^b RCBAr: randomized complete blocks with first order autoregressive structure along rows

autocorrelation structure along rows and columns was found to be superior using the AICD model selection criterion [13]. This model had an efficiency of 225 % compared with RCB. Further information on modelling the heterogeneity of variances is given in the illustration above. When modelled with lattice blocks, the plot errors showed three heterogeneous variances. Improved efficiency in this case was over

500 %. Further, the average SED was much lower when the variance heterogeneity was accounted for, which is also reflected in the high value of the Wald test statistic compared to the other three cases of constant error variances. The per-plot CV (4.4 %) was much reduced from 6.4 % in RCB/lattice to 4.8 % under the RCBArAr model (an almost one-third reduction).

Table 4 Estimates of variance components, heritability and genetic advance due to 20 % selection intensity

Design/model	Group number (<i>i</i>)	No. of plots	σ_g^2	σ_e^2 or σ_i^2	h^2	Genetic advance (%)
Lentil-YT: Tel Hadya, 2004; 16 genotypes. Trial mean = 1427 kg/ha						
RCB		48	29076	8335	0.78	15.98
Lattice		48	29076	8335	0.78	15.98
RCBAr ^a		48	35352	11461	0.76	17.11
Lattice design and heterogeneous plot errors			33527			
	1	28		574	0.98	17.89
	2	11		8027	0.81	16.98
	3	9		22939	0.59	15.51
				Mean (per plot)	0.87	17.23
Lentil- AYT: Breda, 2002; 25 genotypes. Trial mean = 1771 kg/ha						
RCB			20238	30938	0.40	8.47
Lattice			20923	28961	0.42	8.79
RCBAr ^b			20909	32000	0.40	8.60
RCBAr ^b and heterogeneous plot errors			22502			
	1	8		128658	0.15	6.04
	2	67		16592	0.58	10.14
				Mean (per plot)	0.53	9.70
Chickpea-WP. Terbol, 1998; 49 genotypes. Trial mean = 2764 kg/ha						
RCB		98	79775	32659	0.71	13.03
Lattice		98	82786	26795	0.76	13.52
Lattice design and heterogeneous plot errors			87867			
	1	28		1260	0.99	14.96
	2	8		73706	0.54	12.60
	3	28		9878	0.90	14.61
	4	20		48696	0.64	13.29
	5	14		23002	0.79	14.12
				Mean (per plot)	0.83	14.21
Chickpea-SP: Tel Hadya, 2009; 36 genotypes. Trial mean = 1069 kg/ha						
RCB		72	46183	9817	0.82	26.75
Lattice		72	48485	7558	0.87	27.77
Lattice design and heterogeneous plot errors			50318			
	1	34		362.2	0.99	29.32
	2	6		32571	0.61	25.53
	3	22		11403	0.82	27.84
	4	10		4521	0.92	28.74
				Mean (per plot)	0.90	28.47

^a RCBArAr: randomized complete blocks with a separable first order autoregressive structure along rows and columns

^b RCBAr: randomized complete blocks with first order autoregressive structure along rows

Lentil-AYT

In this trial, the CV for the RCB model was 9.9 %. The lattice block design was not very effective, as the efficiency for the lattice design was only 103 % over RCB. The best model using the AICD criterion, the best model with constant variance was ‘randomised complete blocks with plot errors along following first order auto-regression along rows’ (RCBAr), and the efficiency of this model was

125 % compared to RCB. Although the efficiency was higher than the lattice design, the autocorrelation along rows increased the CV only slightly (from 9.9 to 10.1 %). In this case, residuals from the RCBAr models were used for identifying heterogeneous groups for error variances. Except for two clusters, in groups of $k = 3, \dots, 10$ clusters, at least one cluster was found of size 1, i.e. one plot stood out as a plot with a different variance. With poor support to estimate a variance based on just one observation, we

selected the RCBAr model with two heterogeneous variances. In this case, the efficiency of genotypes comparison was 252 % compared to RCB. There was also an increased statistical significance in terms of the Wald test statistic of 169.7 for heterogeneous variances vs. 87.39 for constant variance (the RCBAr structure being common). The CV varied from 7.3 to 19.9 % with still a lower average of 8.7 % per plot.

Chickpea-WP

For the Chickpea-WP trial with 49 genotypes and two replications, there were five heterogeneous error variances, with the number of plots per group varying from 8 to 28. The lattice design had an efficiency of 114 % compared with RCB (100 %). As can be expected, the lattice model had a higher power of discriminating between the genotypes as the Wald statistic value was higher than that for RCB. When we incorporated the heterogeneous variances, the efficiency for pairwise genotypes comparisons increased to 160 % compared to RCB. The CV for RCB was 6.5 % and 5.8 % for the lattice model. In the case of heterogeneous variances, the CV varied in the range of 1.3 % (28 plots) to 10.8 % (8 plots), with a mean CV per plot of 4.6 %, which is less than those for the RCB and the lattice models. The Wald statistic was 1368, which showed an enhanced power to discriminate the genotype effects (Table 3).

Chickpea-SP

Four heterogeneous groups of plots with different variances were statistically detected. Accounting for the heterogeneity, in addition to having used a lattice model, increased the efficiency to 165 %, and the significance level in terms of the Wald statistic to 3354 from 461 for the lattice model. The CV per plot was much lower 6.2 % (range: 1.8 % for 34 plots to 18.6 % for 6 plots).

Effect on Genotypic Variance Component, Heritability, and Genetic Advance

Table 4 gives the estimates of variance components due to genotypes and experimental errors, estimates of broad sense per-plot heritability and genetic advance/gain due to selection at 20 % intensity and their mean per-plot basis.

Lentil-YT

There was an increase in genotypic variance from the RCB or lattice model to the spatial model, and the models with heterogeneous plot errors. The spatial model gave the highest genotypic variance component value of the four

models. The heterogeneous model involved a lattice model, but not the spatial error autocorrelation structure. Heritability was nearly the same for all the models, except in the case of heterogeneous variances where it varied from 59 to 98 %. A high heritability of over 95 % for seed yield in lentil was also reported by Sarker et al. [12] in a lattice model. The genetic advance values were slightly higher on average by 1 % (from 16 to 17.2 %) over the RCB/lattice models.

Lentil-AYT

There was an increase in genotypic variance under the heterogeneous error variance model over the other three models: RCB, lattice and RCBAr. The average heritability per plot of 53 % (range: 15–58 %) showed a considerable increase over the other three models. The genetic advance increased on average by nearly 1 % (from 8.5 to 9.7 %).

Chickpea-WP and Chickpea-SP

In the case of the two chickpea trials, winter and spring planting, a similar trend as in the case of the two lentil trials was observed. Accounting for the heterogeneity of plot error variance led to a net increase in genotypic variance of 10 % in the Chickpea-WP and 9 % in the Chickpea-SP trials when compared with RCB, while in the case of lattice models, these values were 3.8 % for the Chickpea-WP and 5 % for the Chickpea-SP. There was on average a 10 % increase in heritability over RCB, and a net increase in genetic advance of more than 1 % (14.2 % for the heterogeneous variance model vs. 13 % in RCB for the Chickpea-WP; 28.5 % under the heterogeneous variance model vs. 26.8 % under RCB for the Chickpea-SP).

Discussion and Conclusions

In field trials, crop scientists focus on separating the genotypic differences or variation as precisely as possible from the factors inherent in the field, even after accounting for such known factors as replication, blocks within replications, and any other factors accounting for local fertility trends, or structural parameters such as autocorrelation between plot errors. The magnitude of the CV of plot errors provides a guide to the researchers on how far one should intensify the experimental design and analysis of data aspects to achieve the highest precision level. In orthogonal designs such as complete blocks, the CV and the number of replications determine the precision for comparisons amongst genotypes. In incomplete block (balanced or unbalanced) designs, the autocorrelated plot errors also determine the precision of genotype differences. In the

usual context of data analysis from field experiments, the common assumption is that the plot errors have a homogeneous variance. However, due to the presence of many factors creating different degrees of interference, effects on a plot due to residual effects over years or due to within-plot interference due to diseases, pests or competition between and within crop rows [15] such an assumption may not be tenable. Errors in the methodology applied in data collection and entry may also be reflected in extreme residual values, which can be examined for the presence of outliers, or else such errors might be confounded with plot heterogeneity. Often, plot scores such as percentage plant stand (after emergence and at harvest) are taken to explain the heterogeneity, but such variables actually depend on the genotype and should not be used as covariates for plot heterogeneity. Even after accounting for various systematic factors, the pattern of the residuals was found to have a non-constant variance; for example, its spatial structure was found to reflect an exponential variogram model in lentil [12] and a spherical model in herbage plant trials [16].

In the four ‘well-run’ (i.e. CV values of less than 10 %) trials studied, we noticed considerable improvement in efficiency for genotype comparisons and the standard error of differences when lattice blocks and spatial autocorrelations were compared with an RCB model (Table 3). An examination of the graphs (or the table of deviance) established that in all four cases, there was a presence of at least two heterogeneous variances of plot error, as detected by a statistically significant difference in the deviances, when compared with the homogeneous variance model (Table 5). Singh et al. [14] provided a systematic approach for exploring the number of heterogeneous variances. The effect of accounting for the heterogeneity of variances has demonstrated a substantial gain in efficiency of the experimental design, and the enhanced analysis method used in all four trials led to reduced SED values, and hence increased the power of genotypic discrimination. Singh et al. [14] listed examples of high CV values under RCB models where genotypes’ effects were found to be non-significant under a homogeneous error variance model. But, by bringing heterogeneous variances into the model, the genotypic difference became apparent. Therefore, some genotypes which were not significantly different from a best check in a homogeneous error variance analysis might become significant in the heterogeneous model or vice versa. While one expects a gain by such an attempt in highly heterogeneous fields, this study showed that improvement in precision can be brought about even in trials considered ‘well-run’, that is to say, with CV values of less than 10 %.

There were cases where various groups of heterogeneous variances could not be fitted, while retaining a

spatially autocorrelated error structure along rows as well as along columns. This happened in a specific case because of convergence failure in the model fitting. In the case of Lentil-YT, even when sacrificing the autocorrelation structure and introducing heterogeneity of error variances, this led to a substantially higher efficiency (SED = 49.65 for the RCBArAr model vs. SED = 32.95 for the lattice model with three heterogeneous error variances).

The groups of plots with heterogeneous variances identified as mentioned in the Materials and Methods section were further used to estimate the genotype variance component and genetic advance. This process also resulted in a general increase in genotypic variance. This implies that the estimates of heritability and genetic advance in reality could be higher. The presence of multiple error variances, however, does pose a controversial/inconvenient issue of having to deal with multiple indices for heritability and genetic advance. But, the multiple indices offer a more satisfactory explanation of reality than having to compromise on a single value of heritability or genetic advance. In the present four trials, the number of heterogeneous variances varied from 2 to 5. Therefore, it is more realistic to report the distribution of heritability and genetic advance rather than mean values. The mean of genetic advance over the plots gave an additional 1 % gain in each of these cases. Realizing the fact that we are analysing trials with relatively low CV values, this proposed procedure appears to still capture heterogeneous variances, which could help shorten the breeding cycle further by achieving the expected genetic gain sooner and thus increasing breeding efficiency. Modelling has shown that homozygosity can be achieved quicker if we know allele status, which is as expected. In fact, a fully homozygous individual can be identified as early as in the F₂, if population size is large enough to accommodate all the loci segregating, which is pure Mendelian genetics. An extension of this study could be to allow the plot error variance to vary with the genotype and evaluate the precision of the predicted means of

Table 5 Comparing the case of at least two heterogeneous variance groups

Trial	Number of groups (<i>k</i>)	Residual d.f.	Deviance
Lentil-YT	1	29	341.8
	2	27	327.6
Lentil-AYT	1	44	572.1
	2	43	544.4
Chickpea-winter planting	1	46	589.9
	2	44	574.4
Chickpea-spring planting	1	33	390.5
	2	31	374.7

the genotypes, and will be addressed separately. Furthermore, alternative approaches for formation of groups of homogeneous variances clusters may be explored and compared.

This study was conducted to explore the possibility of the role of heterogeneous variances of plot errors in field experiments, and the study supports that heterogeneous error variances are a reality even in ‘well-run’ trials. The method of analysis proposed by Singh et al. [14] has been described and illustrated here for such ‘well-run’ trials. Modelling of heterogeneous variances, in addition to accommodating the effects of systematic factors and autocorrelations between plot errors, has added substantial value to the trials in terms of a significantly higher efficiency of the experimental design and analysis model, a higher power of genotypic discrimination, a lower average CV and a higher genetic advance. This new analysis approach is recommended for adapting the analysis of field trials, but is also relevant to many other experiments (including outside of agricultural research) permitting more precise examination of the experimental residuals and improved precision of conclusions drawn.

Acknowledgments Authors are thankful to Drs Ram Sharma and Wuletaw Tadesse, ICARDA and an anonymous reviewer for their constructive comments and suggestions on earlier versions of the manuscript.

References

1. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* AC-19:716–723
2. Cochran WG, Cox GM (1957) *Experimental designs*, 2nd edn. Wiley, New York
3. Cullis BR, Gleeson AC (1991) Spatial analysis of field experiments—an extension to two dimensions. *Biometrics* 47:1449–1460
4. Fisher RA (1990) In: Bennett JH (ed) *Statistical methods, experimental design, and scientific inference: a re-issue of statistical methods for research workers, the design of experiments, and statistical methods and scientific inference*. Oxford University Press, Oxford
5. Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiments. *J Agric Biol Environ Stat* 2:269–293
6. Hinkelmann K, Kempthorne O (2005) *Design and analysis of experiments, vol 2: advanced experimental design*. Wiley-Interscience, New York, xii + 780 pp
7. Kempthorne O (1983) *The design and analysis of experiments*. R. E. Krieger Publishing Co., Malabar
8. Malhotra RS, Singh M, Erskine W (2004) Application of spatial variability models in enhancing precision and efficiency of selection in chickpea trials. *J Indian Soc Agric Stat* 57:71–83
9. Patterson HD, Thomson R (1971) Recovery of interblock information when block sizes are unequal. *Biometrika* 58:545–554
10. Patterson HD, Williams ER (1976) A new class of resolvable incomplete block designs. *Biometrika* 63:83–92
11. Payne RW (ed) (2011) *The guide to GenStat® release 14. Part 2: statistics*. VSN International, Hemel Hempstead
12. Sarker A, Singh M, Erskine W (2001) Efficiency of spatial methods in yield trials in lentil (*Lens culinaris* ssp. *culinaris*). *J Agric Sci (Cambridge)* 137:427–438
13. Singh M, Malhotra RS, Ceccarelli S, Sarker A, Grando S, Erskine W (2003) Spatial variability models to improve dryland field trials. *J Exp Agric* 39:151–160
14. Singh M, Chaubey YP, Sarker A, Sen D (2010) Modelling unstructured heterogeneity along with spatially correlated errors in field trials. *J Indian Soc Agric Stat* 64:313–321
15. Stringer JK, Cullis BR, Thompson R (2011) Joint modeling of spatial variability and within-row interplot competition to increase the efficiency of plant improvement. *J Agric Biol Environ Stat* 16:269–281
16. Watson S (2000) Spatial dependence and block designs in spaced plant herbage trials. *J Agric Sci (Cambridge)* 134:243–258
17. Wolfinger RD (1996) Heterogeneous variance-covariance structures for repeated measures. *J Agric Biol Environ Stat* 1:205–230