Available online at www.sciencedirect.com
ScienceDirect
Geoderma 140 (2007) 346 – 352
GEODERMA
www.elsevier.com/locate/geoderma

Spatially-Balanced Complete Block designs for field experiments

H.M. van Es [a,*], C.P. Gomes [b,c], M. Sellmann [b], C.L. van Es [c]

[a] *Department of Crop and Soil Sciences, Cornell University, Ithaca, NY 14882, United States*
[b] *Department of Computing and Information Sciences, Cornell University, Ithaca, NY 14882, United States*
[c] *Department of Applied Economics and Management, Cornell University, Ithaca, NY 14882, United States*

Available online 9 May 2007

Abstract

Spatial heterogeneity in fields may affect the outcome of experiments. The conventional randomized allocation of treatments to plots may cause bias and variable precision in the presence of trends (including periodicity) and spatial autocorrelation. Agricultural scientists appear to mostly use conventional experimental designs that are susceptible to adverse affects from field variability. The objectives of this research were to (i) quantify the use of different experimental designs in agronomic field experiments, and (ii) develop spatially-balanced designs that are insensitive to the effects of both trends and spatial autocorrelation. A review was performed of all research efforts reported in Volumes 93–95 of the Agronomy Journal and the frequency of various experimental designs was determined. It showed that the vast majority (96.7%) of agronomic field experiments are implemented through Randomized Complete Block (RCB) designs. The method of simulated annealing was used to develop Spatially-Balanced Complete Block (SBCB) designs based on two objective functions: promoting spatial balance among treatment contrasts, and disallowing treatments to occur in the same position in different blocks, when possible. SBCB designs were successfully developed for designs up to 15 treatments and 15 replications. Square SBCB designs were realized as Latin Squares, and perfect spatial balance was obtained when feasible. SBCB designs are simple to implement, are analyzed through conventional ANOVAs, and provide protection against the adverse effects of spatial heterogeneity, while randomized allocation of treatments still ensures against user bias.

Keywords: Experimental design; Geostatistics; Autocorrelation; Trend; Periodicity; Agronomy

Field experiments in agronomy and related disciplines have traditionally been affected by soil heterogeneity. This is especially of concern when treatment effects are small and soil variability is high, as this inflates the error term. Intrinsic soil variability is the result of the geological, hydrological, and biological factors that affect pedogenesis. The fact that soils are routinely mapped suggests that areas can be identified that are relatively uniform, but more recent research suggests that soils generally constitute a continuum with variability at different scales (van Es, 2002).

The structure of soil variability has important implications for the design of experiments. Most agronomic field experiments are based on the concepts of replication, local control (blocking) and randomization (Atkinson and Bailey, 2001). Replication allows for estimation of the experimental error by applying treatments to different plots under the same experimental conditions. Sufficient replication is needed to distinguish treatment effects from background variability. Blocking is used in field experiments to control the adverse effects of soil heterogeneity. Yates (1936) extended this concept by proposing incomplete blocks where the smaller units are assumed to adhere better to the assumption of uniformity.

The use of randomization has been justified in many ways. Its basic purpose is to remove bias from the estimation of treatment effects (Atkinson and Bailey, 2001), and to equalize the error over all treatment differences (Yates, 1939; Fagroud and van Meirvenne, 2002). Randomization is often considered the best protection and assurance against malicious manipulation of plot layout. Randomization is also believed to better justify the assumption of normal errors. A concern with randomization is the possibility of undesirable outcomes such as treatments being repeatedly located in the same location in different blocks, and treatment pairs being repeatedly located in adjacent positions. This poses no concern when variability is truly random and stationary, but agricultural scientists often admit to minor adjustments to randomized designs when treatment allocations appear undesirable.

* Corresponding author.
E-mail address: hmv1@cornell.edu (H.M. van Es).

0016-7061/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2007.04.017

1. Accounting for nonstationarity

The common assumption in experimental design is that observations y_i are realizations of a random variable Y_i which is independently distributed with the expectation of Y_i being constant (stationary) in the experimental domain:

$$E(Y_i) = \mu \text{ for all } i, \tag{1}$$

and the variance, σ^2, being constant and estimable:

$$E[(Y_i - \mu)]^2 = \sigma^2 \text{ for all } i \tag{2}$$

μ (mean) and σ are often assumed to be parameters of a normal (Gaussian) probability distribution function, thereby allowing for a series of powerful statistical testing procedures. Past research demonstrated that these assumptions are generally erroneous for agricultural fields, and common deviations from the above model are:

- *Nonuniformity of the mean (first-order nonstationarity)*: Within the experimental domain, the land property cannot be assumed to have the same expected value (i.e., Eq. (1) is invalid), but shows structural variation through a trend or discontinuity: The presence and significance of a simple field trend can be identified (David, 1977; Davidoff et al., 1986). A special case of first-order stationarity is the presence of periodicity or cyclical trends, which tend to be associated with cultural practices such as ridge and furrow patterns, wheel traffic, etc., and may be detected by spectral analysis (McBratney and Webster, 1981).
- *Spatial autocorrelation*: This implies that the assumption of independence among observations is incorrect (Nielsen et al., 1973; Vieira et al., 1981; Russo and Bresler, 1981). In such cases, Y_i is considered to be a *regionalized* variable and the variance is expressed in terms of the relative spatial location (h):

$$E(Y_i - Y_{i+h})^2 = 2\gamma_\iota(h) \text{ for all } i \tag{3}$$

or

$$E[(Y_i - \mu_\iota)(Y_{i+h} - \mu_\iota)] = C_i(h) \text{ for all } i \tag{4}$$

where $\gamma_\iota(h)$ and $C_i(h)$ are the semivariogram and autocovariance function, respectively, which can be estimated to verify the presence of autocorrelation. The use of blocking is an implicit recognition of the common presence of spatial autocorrelation and the fact that variance generally increases with scale, i.e., smaller experimental areas have lower variability than larger ones.

Student (1938), as also cited by Atkinson and Bailey, 2001) recognized that field trends can affect the outcome of experiments and argued that plot allocations are “balanced” rather than randomized to reduce bias and the variance of the estimators of treatment differences. Jeffreys (1939) concluded that ‘one should balance or eliminate the larger systematic effects first, and then randomize the rest’, as is done in randomized block designs. Standard analyses (ANOVA) generally are considered to yield valid estimates of treatment effects in the presence of trends and spatial autocorrelation (Brownie and Gumperts, 1997), but detrending methods (Kirk et al., 1980; Tamura et al., 1988) and nearest neighbor analysis and related techniques (e.g., Papadakis, 1937; Wilkinson et al., 1983; Gill and Sukla, 1985) have been successfully employed to improve the precision of estimators of treatment effects.

2. Spatial autocorrelation and design

van Es and van Es (1993) evaluated the spatial nature of randomized arrangement of plots in RCB designs, and determined its effect on the outcome of experiments. Under the common condition of spatial autocorrelation, the distance between plots affects the error variance, efficiency and the outcome of tests (Martin, 1986). If the distance between plots (h_p) equals unity when they are adjacent, the mean distance ($\mu_{h\mathrm{p}}$) associated with any two treatment contrasts increases with the number of treatments (t) in an experiment (van Es and van Es, 1993):

$$\mu_{h\mathrm{p}} = (t+1)/3 \tag{5}$$

This implies that experiments with larger numbers of treatments in (complete) blocks have higher experimental errors, assuming spatial autocorrelation, than those involving lower number of treatments. Also, the spatial nature of randomization is such that the mean distance for any two treatment contrasts has higher variance ($\sigma^2_{h\mathrm{p}}$) with increasing number of treatments, but decreases with the number of replications, r (van Es and van Es, 1993):

$$\sigma^2_{h\mathrm{p}} = (t-2)(t+1)/18r \tag{6}$$

This implies that, when randomized plot allocation is used within blocks, high discrepancy will exist in the spatial distance associated with treatment contrasts when the blocks are large and the number of replications low. It was concluded from probability distributions and a simulation study involving wheat yield uniformity trial data that commonly-used randomization and replication in RCB designs may result in unequal precision in treatment comparisons and erroneous assumptions about test confidence levels in the presence of spatial autocorrelation. Similarly, it can be argued that the presence of field trends or periodicity may generate false treatment effects under certain randomization realizations if some treatments are disproportionally represented in areas of high or low fertility. Incomplete block designs provide some protection against spatial imbalance and improve efficiency (van Es et al., 1989; Lopez and Arrue, 1995; Watson, 2000). Others (e.g., Cheng and Steinberg, 1991; Watson, 2000; Fagroud and van Meirvenne, 2002; Martin et al., 2004) have addressed this concern by considering spatial autocorrelation or trend structures, in some cases from prior soil or crop information, to optimize field designs. Concerns with such approaches are that the design process becomes more costly and cumbersome, and that the autocorrelation structure is

Table 1
Types and frequency of experiments discussed in Agronomy Journal volumes 93 through 95

Type of research	Frequency
Field-based experiment	414
Greenhouse experiment	37
Laboratory experiment	22
Modeling/simulations	20
Review/symposium	27
Other	17
Total	537

difficult to define as variability patterns often change among response variables and may not be temporally stable (Katsvairo et al., 2003; Magri et al., 2005).

Problems associated with trends and spatial autocorrelation can be addressed through improved design and analysis. It was hypothesized that few of these methods are applied by agricultural scientists, because they require considerable additional effort and cost. We set out to quantify the fact that most field scientists prefer simple designs that can be easily implemented and analyzed. Yet, many are also concerned about undesirable realizations of conventional randomized designs that may result in artificial treatment effects due to trends. This research therefore also addressed a need for experimental designs that are robust to both spatial autocorrelation and trends, as suggested by van Es and van Es (1993), and that can be readily implemented by a wide range of agricultural scientists and professionals.

The objectives of this research were to:

1. Through a journal review, quantify the adoption rate of advanced design and analysis methods for dealing with spatial heterogeneity in agronomic field experiments, and
2. Develop a set of spatially-balanced designs that are insensitive to the negative effects of both trends and spatial autocorrelation using the method of simulated annealing, and can be readily adopted by agricultural scientists and professionals.

3. Journal review

3.1. Methods

Volumes 93, 94 and 95 (2001 through 2003) of the Agronomy Journal were reviewed to assess the distribution of experimental design types used by current agricultural scientists. This journal is considered to be a leading scientific publication in the discipline of agronomy with six issues per year. The ISI Journal Citation Report® listed 5753 total citations to the journal in 2003, an impact factor of 1.243, and an immediacy index of 0.148. The number of articles published in Volumes 93, 94, and 95 were 183, 163, and 183, respectively.

For each paper, the research environment (field, laboratory, greenhouse, or other), experimental design, and number of treatments and replications were determined. In cases where multiple experiments were reported in the same article, each was considered separately. In cases where the number of treatments and replicates in the experiment changed over multiple years, average values were used. When experiments involved splits, the main-plot arrangement was used to classify the design type, if known.

3.2. Results and discussion

Volumes 93 through 95 of the Agronomy Journal reported 537 research efforts, some papers including more than one experiment (Table 1). Of those, 414 (77%) were reported to be field experiments, 37 (7%) were greenhouse trials, and 22 (4%) laboratory efforts. The remainder of the papers involved reviews or symposium reports, or others (methodology, notes, survey, etc.). The applied nature of the journal is therefore reflected in the large fraction of field experiments that are discussed in these volumes. Since the concerns with trends and autocorrelation are mainly associated with field experiments, we analyzed the types of designs used for those (Table 2). Of the 414 field experiments, the majority (300, 72%) were implemented as RCB designs. Completely Randomized, Randomized Incomplete Block, Split Block and Latin Square designs were rarely used (4, 3, 2, and 1 occurrences, respectively; Table 2). In addition, 9 experiments involved non-randomized field strips, typically involving on-farm research efforts, and 53 involved other field sampling efforts (surveys, etc.). The journal volumes discussed 42 field experiments that were conducted as split plot without any indication of the main-plot design, which is a notable omission by both authors and editors. Some other experiments were also conducted as split plot, but were classified under the main-plot arrangement.

The review of these three volumes of the Agronomy Journal shows that 96.7% (300/310) of the field experiments with known main-plot design were implemented using randomized complete blocks. Clearly, agronomists favor this design and rarely see compelling reasons to use more advanced designs that more explicitly address spatial variability concerns (i.e., incomplete blocks, Latin Squares, etc.). Also, no experiment was analyzed using trend or nearest neighbor analysis. It is presumed that most agronomists prefer RCB designs for their simplicity and intuitive layout.

Table 2
Characterization of designs used in field-based experiments reported in Agronomy Journal volumes 93 through 95

Design type	Frequency	Mean	
		# of treatments	# of replicates
Randomized Complete Block	300	8.0	3.8
Completely Randomized	4	17.3	14.7
Randomized Incomplete Block	3	43.6	3.7
Split Block	2	4.0	4.5
Latin Square	1	4	4
Field strips — unknown design	9	4.6	3.2
Split plot — with unknown main-plot arrangement	42	NA	NA
Other	53	NA	NA
Total	414		

It is notable that Completely Randomized designs had higher mean number of treatments and replications (Table 2), and were often associated with studies involving many species or varieties (e.g., trees). The three incomplete block experiments included a variety trial with 120 entries, and two studies with only 5 or 6 treatments, of which one explicitly mentioned the potential gains from incomplete blocks, even in smaller experiments.

An analysis of the greenhouse experiments revealed that out of a total of 37, 16 (43%) were implemented as Completely Randomized designs, 17 (46%) as RCB designs, and 4 did not involve an experimental design. This suggests that agronomists are less concerned about spatial variability in greenhouse experiments and more frequently use unblocked designs, despite the fact that greenhouses are known to have spatial trends in environmental indicators such as air temperature, humidity, and solar radiation.

4. Development of spatially-balanced designs

4.1. Methods

The previous section documented the strong preference by field agronomists for complete block designs and their presumed reluctance to deal with more complex plot layouts and analysis

Table 3
Spatially-Balanced Complete Block designs for experiments with up to ten treatments (a–j) and eight blocks

# of treatments	# of blocks						
	2	3	4	5	6	7	8
2	ba ab	ab ba ba	ba ab ab ba	ab ba ab ba ab	ba ab ab ba ba ab	ab ba ab ab ba ab ba	ba ab ab ba ba ab ba ab
3	bca cab	cba bac acb	bac acb abc cab	bac acb cba bca cba	acb cba bac acb cba bac	bac abc bca cab acb cba cab	cab abc acb cab cba bac bca abc
4	dcab cbda	cdab dbca dabc	abcd cdab dabc bcda	cbda dcab acbd badc bdca	dcba bcda bdac abcd cadb dabc	cdba dabc bacd adcb dbac bcad acdb	cadb dbca bacd adbc dcab adbc bcad cbda
5	debac bdcea	cebad eadcb dceba	edacb dbcea acdbe baedc	cdebf dabce aecdb ebdac bcaed	caedb dbaec bcdae becad aebcd edcba	cdeba dabce aecdb beacd bcade adbec ebdac	baecd ecadb dbcea cbdae adebc becda cabed edbac
6	cabfed becdaf	facedb cbadef efbacd	bcedfa cafedb efbacd fdcbae	beacdf afbdec fceabd cbdfae dacefb	feadbc baecdf edcbfa cfbead acdfeb dbface	fbecda cfdabe bdaefc decbaf eafdcb acbfed ecbafd	deabcf cbefda bdfcea cabdfe afcebd fdcaeb ecdfab febadc
7	gbcdeaf dagfbec	bcgadef gedcbfa caefgbd	fdcbage bgfcead afbecdg cagfdeb	cdgfeab fbdagce gabcdef deagfbc acfebdg	eafdbcg fcgaebd gbaecdf cebfdga dgecfab adcbgfe	fbdgcae afcegdb cdefbga dagbecf gcbafed egfdabc beacdfg	fgdceba cabfdge bgcdfae fbegacd afgcedb gceabfd ecfdabg dagebfc
8	cfehgdba hdfacbge	cefdbhag hfgcdaeb facbgehd	dbfhegac bcheadfg hdgcbaef egbdcfha	fhegcabd cfdegbha ebchafdg aefdbhgc hdecagfb	gbfceahd efdbhgac hgefacdb aebgdhcf gdafcebh fabhgdec	hcdabefg debafhcg bfcheagd fehdcgba caefgdhb ahfgdbec ehgbacdf	beachgdf gahdcefb dbcgafhe afdeghbc hdebfcga chfabdeg fgbheacd ecgfdbah
9	fbicagdeh cdbghfeai	caefdhibg ehicgfdab dicbeaghf	dgeafhibc aighcefdb ehabdigcf iecdafbgh	bacghfdei ahdceibgf ebhfagidc geadfbcih fceaihgbd	fedigbcah dabgecfhi hdcebgifa ibehafdcg bcfdiahge gfhbdiaec	ifachdgeb fbcgdaeih gahfebicd hcbeaidgf edfhcgabi dhgbifcae cgeifhbda	icedabhgf aifhdgcbe dfhebigac fbgcidaeh geafchdib hgibeafcd badghcefi chbafeidg
10	ciadbgfhje gdhiecjafb	dgbhfjceia bcfadhiegj fidjbagech	gcbejdfaih edacighbfj bieagfjchd agjdbheicf	gcfijdhbae dfaghibjec bgjdefcaih idbcagefhj fbchdegija	chbigdfeja diabhegjfc ijhedfcagb egicahdfbj gdjhcaebif hafgijbcde	fecgaibjdh ahfbegidcj gihfjedbac bgdcfhjeia edgahjcfbi hciedbfajg jbehgcaifd	adjheifgcb eghfjbacdi jcgiahedbf baechjifgd hibdgecafj chfadgbjie ifaecdgbjh decjbfhiag

Table 4
Spatially-Balanced Complete Block designs for experiments with 11 to 15 treatments (a–o) and up to eight blocks

# of treatments	# of blocks						
	2	3	4	5	6	7	8
11	bcijakgehdf kjhcgfbdiea	acjgiehkbfd bghafjdicek ifgkcbadehj	eikfacdgjbh fdihjabkgec jkfcdbehiag dakhejgfcib	jfkdagcheib abckfhejigd gahjbefdcki ekgbjcdaihf cjeaifgbkdh	fibacekdhjg ahikdjbcgfe dbeighajfkc ijcehdfbagk eajdfkigbch dcagifhekbj	bfgcjhadkie gaejdfbkihc fkjaicgehbd ecjbkadhfgi djigcbkaefh kgcdfehjiab jhkgbeifdca	fhebkjagdic ejcdbgfkiha iekgjachfbd gbaehidcjfk acbkefidhjg kdhjaebicgf bijhckgfade jafidbkegch
12	fjcahlgibked liakjdfecbhg	cjeglhikabfd gkfhecdbijla hblcfgajdkei	cjeglhikabfd gkfhecdbijla hblcfgajdkei	dfgichebakjl ailfhedjkgcb fkhgabldecij hcadjfblgike bdihklcfjaeg	iadkhgfcbjel bkijgcaehldf gfkejadilbhc hjafbkldgcie aglbifehkdcj dbghejifcalk	eijfgcblhkda bafiklecdhgj flhjaedikcbg hbedfgkajilc dfcbjihklgae kjbledfgcahi ldigbajhfeck	kdilfagcebjh djlcgbeikhaf acjikhlbdfge gabkdjhflcei ehdaclbkgifj cfkbedajighl blafjeidhkcg lkejagfhcdib
13	fkihgabcdmjel adhjklbfeimgc	fgdbjcmklhaie kdacbeglhfijm chgmadifekjbl	kjfbaglmidche hamjifckdbleg bcijlhkeagmdf ikmgbhfejcdal	mfdcgkejibhal hgeimcajfldbk jidglmakehfbc kichdaflgebjm agfibkmhjdcle	kmedhblafigjc ajdeifckbhgml kigdlabcemjhf leihjkcbamgdf cdhimalfkjgeb bimjdlfkcegha	jdfgkbeilahcm hjkblagmdcfie flhedkmgaibjc gilkcdhjemafb lejmgcbkfhida mkifjladcbehg kaecfgjhilmbd	jgdkcmelabfih bjmekihdfgacl debchgijlfmak hmldgbakeijfc eaghmfjcbdlki mcigbakfdhelj gbflekcmhjida khcbajlgiedmf
14	afegjhlmdikncb ldjkecabfngihm	dnbjhcagfkelim cengljmbkdfiah kjalcdinehmfbg	mdfeagcknjilhb jgdhicmelnfbka eingbmahkdjflc nchadbejfilmgk	lmnbjhgkafedci jcfmhabdgelink alecbmkihgfnjd gbcnafiljkdmeh fkbhldcnejiamg	enkldamjficbhg kjfhndgimlbeca dmhcfelkgbjani jclngmhakeifbd gkmiejcdbnhlaf mbnfkcagdjehil	jdbhaeilkcfmng bncdlfhaikgjem ifebgdcmahknjl dkinhmeblfjgac eanfdlkgbjmhci lmdaigbnjeckhf kbamfjnicdelgh	dnhgmljacebkfi lkedjbainhcgmf mabkdngceflihj khclagdfbinjem fdjacehnkmilgb gjkniaflmcdebh egaihdkmfljbcn alnfehbgjkmdic
15	nbamcdefgkjlohi ekclahnifbomjdg	becdfhjialnkgom ihgelncodmbjakf ndaimefgbkhocjl	ikobhlafndjcmge minfljekbocghad fhjgkmodiealncb ldkcfgibmhneajo	jgoebhnmcflkdia imekhcjlagbdfno clbmgiokndjaehf hdmoljifbekgacn legdinhcoamfjbk	ielngfmbchkajdo bgkldjimfahnoce dfncjghikebolam gachlidobfeknmj flhjaebdngomcik hmdgeofklcjianb	aiodcjblhkgenmf lbfogamiknjhdec gnaclkjofdeimhb jfighclmdabneok dgbjkmafceiohnl ikcbfndhogmalje hakfdlegibnjocm	mfokailjcbnhegd cidmnbjfahogkel jgmlcnhafeikbdo khcfjdelmobnaig ijhoemnkdgfaclb nkjadlocgiemhbf lnfihodgkcmbjae ocnefjgbikldmha

methods. Designs were therefore developed that are inherently insensitive to non-random field variability within the framework of the generally-favored complete block layout. Spatially-Balanced Complete Block (SBCB) designs are based on standard, spatially optimized experimental layouts to which treatments can be randomly assigned. Such random assignment of treatments ensures against user bias and allows for a large number of possible design outcomes. SBCB designs are a subset of all possible design realizations for RCB designs, with those designs excluded that may cause bias and imprecision when implemented on trended or autocorrelated experimental domains.

SBCB designs were developed using the computational method of simulated annealing, based on the successful approach for the Traveling Tournament Problem (Anagnostopoulos et al., 2003). Two simultaneous objectives were applied (Gomes et al., 2004): (i) promoting spatial balance among treatment contrasts, and (ii) disallowing treatments to occur in the same position in different blocks (unless the number of replications are greater than the number of treatments). Spatial balance among treatment contrasts was evaluated based on the average distance between plots for each of the $\binom{t}{2}$ possible treatment comparisons, for which $\sigma^2_{\bar{h}p}$ (Eq. (6)) was minimized. Designs were therefore balanced based on distances of all treatment contrasts, not based on first-order autoregressive assumption as done by Cheng and Steinberg (1991) and Martin et al. (2004). Our designs therefore are optimized based on a linear variogram model with a range greater than the width of a complete block. Simulations were performed for up to 15 treatments and 15 replications, which covers the majority of experiments conducted by agronomists.

Simulated annealing is a heuristic computational optimization method where in this case designs were generated and progressively improved through a local search approach with a pre-selected search neighborhood. Simple neighborhoods appeared to result in better performance than more complicated neighborhoods. In this study, one simple move was used for each subsequent design improvement, i.e., swapping a random pair of treatment indicators in a random block (Gomes et al., 2004). Five variations were used on the strength of the two objective functions. One design was selected from the five obtained for each treatment–replication combination based on the degree to which objectives were met, and whether treatment allocations in blocks were unique within a given design, as much as possible. The simulated annealing approach was implemented in C++ compiled with the GNU G++ compiler version 3.2.2, and executed on the Cornell University Department of Computing and Information Science computer cluster.

4.2. Results and discussion

The computational requirements for the simulated annealing effort went up exponentially with the increasing size of the design, as discussed by Gomes et al. (2004). For example, the optimum solution for square designs was found with 0.01, 0.36, 153, and 883 CPU seconds (mean values for ten runs) for designs of order 3, 6, 9, and 12. Obtaining 225 optimum SBCB designs (up to 15 treatments and 15 replicates) therefore required several weeks of simulations on a 25-unit computer cluster. Larger SBCB designs were not derived using simulated annealing because the computational method was not capable of converging on perfectly balanced designs due to the complexity of the multiple treatment arrangements.

Slight variations in the relative strength of the two objective functions (balance of spatial distance vs. different locations of treatments in blocks) resulted in different SBCB designs, of which one was chosen that best provided unique treatment allocation in blocks. Designs are listed in Table 3 for up to 10 treatments and 8 blocks, and in Table 4 for 12 to 15 treatments and up to 8 blocks.

Perfect spatial balance ($\sigma^2_{hp}=0$) was generally achieved when theoretically possible for up to 15×15 designs. This resulted in spatially-balanced Latin Square designs when the number of treatments and replications were equal. Such perfectly balanced square designs (i.e., all $\binom{t}{2}$ treatment contrasts have equal average distance of comparison) can be obtained when (Gomes et al., 2004):

$$t \bmod 3 \neq 1 (\text{i.e.,} 2\times2, 3\times3, 5\times5, 6\times6, 8\times8, 9\times9, \text{etc.}) \quad (7)$$

This fact allowed for an independent evaluation of the simulated annealing effort, and provided indication of its limits for large designs (greater than 15). It is noted that a perfectly balanced square design implies a Latin Square arrangement, but that the reverse does not hold and most Latin Square designs are in fact spatially unbalanced. Perfect spatial balance for non-square designs was generally achieved when theoretically feasible. Spatial balance in all other designs was optimized.

These standard layouts can be used in experimental design by randomly allocating treatments to the letter indicators (Tables 3 and 4). Blocks in the layout may also be interchanged as this does not affect spatial balance. The random assignment of treatments eliminates user bias and provides a large number of possible design outcomes, although not as numerous as in traditional RCB designs. Split-plot designs can also be based on SBCB designs through a multi-stage procedure where spatially-balanced main plots are first identified and spatially-balanced split plots are subsequently defined within the main plots using Tables 3 or 4. Although spatial balance in the designs was developed for treatments that are laid out adjacently, the results also generally provide good designs when the treatments are implemented in other arrangements (e.g., blocks of 8 treatments laid out as 2×4).

Most SBCB designs can be analyzed using ANOVA methods that account for block effects, similar to conventional RCB designs. Square SBCB designs are a special case of Latin Square designs and may be analyzed as such. The random initialization and search methods in the simulated annealing method, combined with the random allocation of treatments to indicators provide assurance that basic assumptions underlying ANOVA are adhered to.

5. Conclusions

A review of experimental procedures reported in recent volumes of the Agronomy Journal indicates that the vast majority (96.7%) of field experiments conducted by agronomists are implemented through RCB designs. The use of blocking addresses the concerns about spatial autocorrelation in fields, but such designs do not explicitly deal with other issues related to spatial balance or trends. Most agronomists do not make efforts to address such concerns and apparently consider the conventional RCB designs useful, convenient, tried and proven.

This research focused on developing designs that are inherently robust to concerns with both field trends and spatial autocorrelation, but in the context of the popular complete block designs. The designs provide spatial balance among treatment contrasts and distribute the treatments among locations in blocks in different replications. They are based on common assumptions of spatial variability structure and do not require detailed quantification of the variability structure. The SBCB designs therefore provide a simple way to ensure that the experiment is not adversely affected by spatial variability, without requiring additional field data, complex experimental design procedures, or alternative data analysis. The random nature of the simulated annealing method that was used to develop the designs, as well as the randomized allocation of treatments ensures the validity of analysis assumptions and protects against user bias. Moreover, SBCB designs can readily be implemented by field professionals for use in experimentation.

Although we have focused on field experiments, these designs also have application to other types of experiments where trends and autocorrelation are of concern, including greenhouse trials. The principles of spatially-balanced designs also apply to a number of laboratory methods, including multi-well titer plates used for chemical and biological analyses, and microarray slides used in genomics research, all of which are known to have problems with non-random spatial patterns.

Acknowledgements

This work was supported by the Intelligent Information Systems Institute (IISI), AFOSR grant F49620-01-1-0076, and the USDA-CSREES funded Computational Agriculture Initiative, Special Grant 2003-6193. We also acknowledge the contributions of Andres Berger with the review of the Agronomy Journal articles.

References

Student, 1938. Comparison between balanced and random arrangements of field plots. Biometrika 29, 363–379.

Anagnostopoulos, A., van Hentenryck, Michel L., Vergados, Y.P., 2003. A simulated annealing approach to the traveling tournament problem. Proc. 5th Intern. Workshop on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems (CP-AI-OR 2003), Montreal, Canada, May 2003.

Atkinson, A.C., Bailey, R.A., 2001. One hundred years of the design of experiments on and off the pages of Biometrika. Biometrika 88, 53–97.

Brownie, C., Gumperts, M.L., 1997. Validity of spatial analyses for large field trials. J. Agric. Biol. Environ. Stat. 2, 1–23.

Cheng, C.S., Steinberg, D.M., 1991. Trend-robust two-factorial level designs. Biometrika 78, 325–336.

David, M., 1977. Geostatistical Ore Reserve Estimation. Elsevier Publ. Co., New York.

Davidoff, B., Lewis, J.W., Selim, H.M., 1986. A method to verify the presence of a trend in studying spatial variability of soil temperature. Soil Sci. Soc. Am. J. 50, 1122–1127.

Fagroud, M., van Meirvenne, M., 2002. Accounting for soil spatial autocorrelation in the design of experimental trials. Soil Sci. Soc. Am. J. 66, 1134–1142.

Gill, P.S., Sukla, G.K., 1985. Efficiency of nearest neighbor balanced block designs for correlated observations. Biometrika 72, 539–544.

Gomes, C., Sellmann, M., van Es, C., van Es, H., 2004. The challenge of generating spatially balanced scientific experiment designs. In: Régin, J.-C., Rueher, M. (Eds.), Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems, First International Conference, CPAIOR 2004, Nice, France, April 20–22, 2004, Proceedings. Lecture Notes in Computer Science 3011 Springer 2004. ISBN: 3-540-21836-X.

Jeffreys, H., 1939. Random and systematic arrangements. Biometrika 31, 1–8.

Katsvairo, T., Cox, W.J., van Es, H.M., Glos, M.A., 2003. Spatial yield responses of two corn hybrids to two N levels. Agron. J. 95, 1012–1022.

Kirk, H.J., Haynes, F.L., Monroe, R.J., 1980. Application of trend analysis to horticultural field trials. J. Am. Hortic. Soc. 105, 189–193.

Lopez, M.V., Arrue, J.L., 1995. Efficiency of an incomplete block design based on geostatistics for tillage experiments. Soil Sci. Soc. Am. J. 59, 1104–1111.

Magri, A., van Es, H.M., Glos, M., Cox, W.J., 2005. Integrated assessment of the use of crop, soil, and remote sensing information for precision management of maize. Precision Agric. 6, 87–110.

Martin, R.J., 1986. On the design of experiments under spatial correlation. Biometrika 73, 247–277.

Martin, R.J., Eccleston, J.A., Chan, B.S.P., 2004. Efficient factorial experiments when the data are spatially correlated. J. Stat. Plan. Inference 126, 377–395.

McBratney, A.B., Webster, R., 1981. Detection of ridge and furrow patterns by spectral analysis. Int. Stat. Rev. 49, 45–52.

Nielsen, D.R., Biggar, J.W., Erh, K.T., 1973. Spatial variability of field-measured soil–water properties. Hilgardia 42, 215–259.

Papadakis, J.S., 1937. Methode statistique pour des experiences sur champs. Bull. Inst. Amel. Plantes a Salonique 23.

Russo, D., Bresler, E., 1981. Effect of field variability in soil hydraulic properties on solutions of unsaturated water and salt flows. Soil Sci. Soc. Am. J. 45, 675–681.

Tamura, R.N., Nelson, L.A., Naderman, G.C., 1988. An investigation on the validity and usefulness of trend analysis for field plot design. Agron. J. 80, 712–718.

van Es, H.M., 2002. Sources of soil variability. In: Dane, J., Topp, C. (Eds.), Methods of Soil Analysis, Part 4: Physical Properties. Soil Sci. Soc. Am., Madison WI.

van Es, H.M., van Es, C.L., 1993. Spatial nature of randomization and its effects on the outcome of field experiments. Agron. J. 85, 420–428.

van Es, H.M., van Es, C.L., Cassel, D.K., 1989. Application of regionalized variable theory to large-plot field experiments. Soil Sci. Soc. Am. J. 53, 1178–1183.

Vieira, S.R., Nielsen, D.R., Biggar, J.W., 1981. Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 45, 1040–1048.

Watson, S., 2000. Spatial dependence and block designs in spaced plant herbage trials. J. Agric. Sci. 134, 245–258.

Wilkinson, G.N., Eckert, S.R., Hancock, T.W., Mayo, O., 1983. Nearest neighbour (NN) analysis of field experiments. J. R. Stat. Soc., B 45, 151–211.

Yates, F., 1936. A new method for arranging variety trials involving a large number of varieties. J. Agric. Sci. 26, 424–455.

Yates, F., 1939. The comparative advantages of systematic and randomized arrangements in the design of agricultural and biological experiments. Biometrika 30, 440–446.