

Recommendable block sizes: a case study on Finnish official variety trials of barley cultivars

Lauri Jauhiainen, Jukka Öfversten

MTT Agrifood Research Finland, Services Unit, FI-31600 Jokioinen, Finland, e-mail:lauri.jauhiainen@mtt.fi

Arjo Kangas

MTT Agrifood Research Finland, Plant Production Research, FI-61400 Ylistaro, Finland

Well-established results in the current statistical literature imply that plant breeders should use incomplete block designs wherever spatial variability exists and the number of treatments is large. But the theoretical position does not indicate the recommendable number of cultivars in an incomplete block. In this study we used data from 28 official variety trials conducted in Finland during the period 2001–2005 to study the effect of block size on the efficiency of testing pairwise yield differences of barley cultivars and cultivar rankings. In previous trials some 6–7 cultivars have usually been included in one block. Our results imply that the efficiency of testing procedures could be improved by using a block size as small as 4–5. The results further imply that if an experiment with an incomplete block design is well planned to mitigate the effects of within-block heterogeneity, the spatial mixed model techniques and the conventional analysis of variance techniques have approximately the same efficiency in testing pairwise yield differences. Thus, if appropriate blocking strategies are used in planning the trials, there is usually no need to change the conventional practice followed in statistical analysis.

Key-words: spatial variation, experimental design, block size, statistical analysis

Introduction

It is well known that soil heterogeneity complicates field experiments. The statistical issues that arise can be dealt with in two ways. On the one hand we can use experimental designs that account for spatial

effects in the soil and, on the other, we can use statistical analyses that account for spatial dependence in the data obtained. In designing experiments, both randomised complete block designs and incomplete block designs have been widely used. The data obtained have been conventionally analysed using analysis of variance techniques complying with

the experimental design but ignoring the potential spatial dependency of the residual variances. Several more complicated attempts to account for spatial variability have also been made. The earlier ones were based on nearest neighbour adjustments introduced by Papadakis (1937) and later improved e.g. by Kempton (1981). In recent years, spatial mixed models with spatial covariance structures have been proposed.

A lot of work has been done to compare the efficiency of the different methods for analysing data from variety trials. The results obtained suggest that in most cases the spatial mixed model gives the most efficient analysis (e.g. Kravchenko et al. 2006). This leads to the question of whether the choice of experimental design has any effect on the efficiency of spatial mixed model analysis. The power analyses made by Stroup (2002) give an explicit answer to this question: the design does matter. Even when the spatial structure is known and can therefore be specified exactly, incomplete block designs specifically intended to mitigate the effects of within-block heterogeneity are more powerful than the complete block designs used in conjunction with spatial mixed models. Spatial mixed model techniques used together with incomplete block designs are even more powerful.

The results of Stroup (2002) lead to the conclusion that researchers should use incomplete block designs wherever spatial variability exists and the number of treatments is large. But there still remains at least one open question: what is the recommendable number of treatments in one block? The answer to this question clearly depends on the nature of the spatial variability. In practice the variability is always unknown, but it is linked to geographic conditions such as the regional location of the experimental site and the soil characteristics of the experimental fields concerned, and also depends on the shape and layout of the field plots used in experimentation. Moreover, it depends on the response variables recorded for the analyses (Watson 2000).

In this study we used Finnish official variety trial data on barley cultivars to investigate the effect of block size in testing pairwise yield differences and cultivar rankings. The results were intended

to help in determining the recommendable block sizes of the incomplete block designs needed for forthcoming Finnish variety trials on barley cultivars. The research approach was made possible by the long and well-organised tradition of official variety testing supervised by MTT Agrifood Research Finland. Variety trials on barley have been carried out continuously for more than 30 years now. The main response variable considered has always been the grain yield. During the long testing period most of the operating procedures related to the experimentation have become well established. In particular, the shape of experimental plots and the practices for laying them out in the field have been standardised. But there have never been any well-established principles for determining appropriate block sizes for the incomplete block designs predominantly used.

Material and methods

Trial data

We used Finnish official variety trial data originating from 28 trials conducted in 2001–2005 at 8 sites in southern Finland. In each trial the cultivars to be tested were arranged in the field using incomplete block designs with four replications. The number of cultivars in one trial varied from 24 to 43 and the block size ranged from 4 to 7 plots. Within each trial the block sizes never differed more than one plot. The data are given in Table 1. Plots were 1.5–1.75 metres wide and 8–10 metres long. Trials were carried out according to well standardised and documented cultivation recommendations (Järvi et al. 1998).

Spatial models

A variogram measures the correlation between data pairs as a function of the displacement between the pairs (Brooker 2001). In our data, detailed layout

Table 1: Trials (years and names of locations), numbers of cultivars tested, used block size or range of original block sizes, average grain yields and standard error of difference (SED).

Site	year*	number of cultivars	original block size	average yield	SED
Mietoinen	2001	41	5-6	5039	146
	2002a	32	4	3260	224
	2002b	24	4	3562	120
	2003a	28	4	2711	132
	2003b	25	5	3572	171
	2004a	37	5-6	6754	308
	2004b	28	4	5690	316
	2005a	42	6	5820	155
	2005b	25	5	5423	423
Pälkäne	2001	31	5-6	5157	191
	2002	35	5	4190	271
	2003	32	4	4988	319
	2004	43	6-7	4977	352
Mikkeli	2001	31	5-6	5439	415
	2002	25	5	4846	314
	2004	32	4	4809	221
	2005	30	5	5800	377
Jokioinen	2001	32	4	6382	171
	2002	26	4-5	4549	160
	2003	25	5	5288	218
Ylistaro	2002a	33	5-6	4844	319
	2002b	24	4	3714	318
	2004a	43	6-7	5159	312
	2004b	37	5-6	652	276
Tuusula	2001	31	5-6	6725	176
Pernaja	2004	26	4-5	5954	218
Inkoo	2004	38	5-6	5327	272

*two trials conducted at Ylistaro and Mietoinen in the same year were placed at separate fields with different soil types

information was available on 9 trials. Using these trials a semivariogram $g(h)$ measuring the spatial dependency was calculated as:

$$g(h) = \frac{1}{2N(h)} \sum_{d_{ij}=h} (y_i - y_j)^2 \quad (1)$$

where y_i and y_j are observed yields from field plot i and j , respectively, d_{ij} is the distance between these plots, and $N(h)$ is the number of field plots separated by the same distance h . The semivariogram was computed in different directions to ensure that the

spatial variation was isotropic. Variogram $\gamma(h)$ was then modelled from the semivariogram values using a spherical model with a nugget effect as follows:

$$\gamma(h) = \begin{cases} c_0 + c \left\{ \frac{3h}{2r} - \frac{1}{2} \left(\frac{h}{r} \right)^3 \right\} & \text{for } h \leq r \\ c_0 + c & \text{for } h > r \end{cases} \quad (2)$$

where r , c_0 and c refer to the range, the nugget and the sill, respectively. The semivariograms and parameter estimates of the modelled variogram were calculated using VarioWin software (Pannatier, 1996). Before the analysis, the effect of cultivars

was removed from the data using the one-way ANOVA. Kriging is one technique among many for interpolating a variable from sample points. It has the advantage that the estimated points are obtained with minimum variance (Lark 2000). The estimated values of the parameters of equation (2) were therefore used to predict values at the nodes of a 30 cm grid by ordinary punctual kriging using the KRIGE2D procedure included in SAS software (SAS 1999). A yield map was created using these kriging estimates.

To indicate the effect of block size and the resulting efficiency of an experimental design we used standard errors of pairwise yield differences (SED). To estimate such standard errors we used the following spatial mixed model:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \tag{3}$$

where μ and α_i are fixed effects associated with the grand mean and i^{th} cultivar, respectively, and the residuals ε_{ij} are spatially correlated with a covariance structure defined in equation (2). SED for all pairs of cultivars were calculated during the analysis. They were found to vary with the locations of the cultivars in the field. To measure the efficiency of a particular design and analysis we therefore used the mean of all the SED values associated with this experiment.

Superimposing new blocking on the data

In this study we used field data in which each replication included all the experimental plots arranged in a row. This made it possible to superimpose new blocking strategies on the existing data. New blocking for a block size of 2 was made by dividing one replication into several blocks of 2 plots. If the number of plots was odd, we randomly included a block containing 3 consecutive plots. New blockings for block sizes of 3, 4, 5, 6–7, 8–10 and 14–20 were made in the same way.

Differences between cultivars were analysed using the following traditional mixed model for incomplete block design:

$$y_{ijk} = \mu + r_j + b_k(r_j) + \alpha_i + \varepsilon_{ijk} \tag{4}$$

where y_{ijk} is the yield of the i^{th} cultivar from the j^{th} replication and the k^{th} block, μ is the intercept, α_i is the effect of the i^{th} cultivar, r_j is the effect of the j^{th} replication, $b_k(r_j)$ is the effect of the k^{th} incomplete block nested within the j^{th} replication and ε_{ijk} is the residual error. r_j , $b_k(r_j)$ and ε_{ijk} are assumed to be random in the model used. The assumptions for the random effects were: $b_k(r_j) \sim iid N(0, \sigma_b^2)$, $r_j \sim iid N(0, \sigma_r^2)$, $\varepsilon_{ijk} \sim iid N(0, \sigma^2)$, and all the effects are independent of one other. For comparison (Fig. 1), the data were analysed using ANOVA for the randomised complete block design based on equation (4) and without the $b_k(r_j)$ effect. The parameters of the models were estimated using the restricted maximum likelihood (REML) method with the SAS system and MIXED procedure (SAS 1999).

Standard errors for estimated pairwise cultivar effects were calculated and used as in the spatial analysis. The incomplete block designs are, however, featured by the fact that the frequency of occurrence of the same pair in the same incomplete block has an effect on the standard error. For each block size

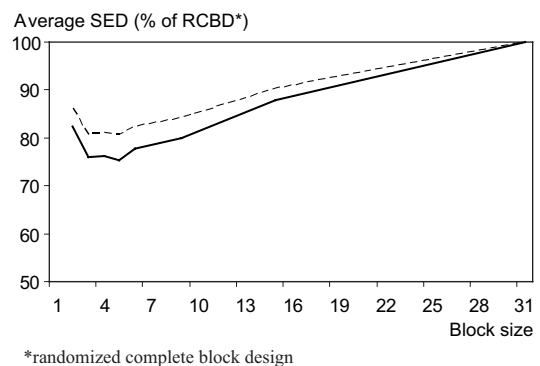


Fig. 1. Average standard error of the difference (SED) of two cultivars as a function of a block size. A SED value for randomized complete block design was fixed at 100%. Dashed line is based on all trials whereas solid line is calculated without seven trials in which the difference between analyses based on incomplete and complete block designs was not found.

we therefore drew Box and Whisker plot to describe symmetry of the distribution of standard errors. In addition, we calculated the standard deviation of the respective standard error. This was then used to identify block sizes which were more robust and less sensitive to spatial variation.

For each trial the results of the complete block analysis and the incomplete block analysis were used to rank the cultivars according to their estimated mean yields and differences between the rankings. Further, for each trial, maximum and average differences between the rankings of all pairs of cultivars were calculated. These were then used to estimate the validity and relevance of our results.

Results

The spatial variation in all trials could be considered isotropic, because no clear differences were observed in the variograms computed in different directions. Estimates for the parameters of the spatial mixed models are given in Table 2. All models had a clear nugget effect. The estimate for the range parameter was fairly similar in all trials, whereas sill and nugget effects varied considerably. The

trial at Ylistaro in 2004 had low estimates for all three parameters, while Mietoinen 2005a had the smallest estimates for sill and nugget parameters (Table 2). The spatial dependence abated clearly in all trials when the distance between field plots increased from 2 to 10 metres i.e. the number of plots between two fixed plots increased from 0 to 5. When the block size reached 20 plots, the spatial dependence between the two outermost plots was negligible in most of the trials.

Grain yield maps revealed that the areas of different yield levels did not follow the shape of replications or blocks. In many cases, such as Mietoinen in 2004 (Fig. 2), growing conditions could be seen to vary widely within one replication. In fact the yield could differ by over 500 kg/ha (10%) within a distance of five plots (e.g. Mikkeli 2005, Fig. 3).

The smallest SED values were reached when the block size was 3, 4 or 5 plots (Fig. 1). With such block sizes the SED value was on average 19% smaller than that for the randomised complete blocks design. For the historically widely used block size of 6–7 plots/block, the SED value was only slightly higher (1–2%) than for block sizes of 3, 4 or 5. The SED value increased rapidly when the block size was 2 or higher than 9. However, blocking had no effect on SED in 7 trials, and therefore the average SED value was also calculated without these trials (dashed line in Fig. 1).

Table 2: Parameter estimates for the spatial mixed model and effectiveness of the model compared to the analysis based on incomplete block designs.

Site	year	nugget	sill	range	relative SED*
Mietoinen	2004a	70 000	150 000	37	100
	2004b	111 000	207 000	53	96
	2005a	26 000	23 000	37	98
	2005b	70 000	210 000	33	93
Pälkäne	2004	140 000	110 000	43	100
Mikkeli	2004	38 000	92 000	28	98
	2005	40 000	50 000	25	100
Ylistaro	2002	100 000	300 000	38	72
	2004a	40 000	50 000	25	65

*compared to the most accurate analysis of incomplete block designs, SED=standard error of difference

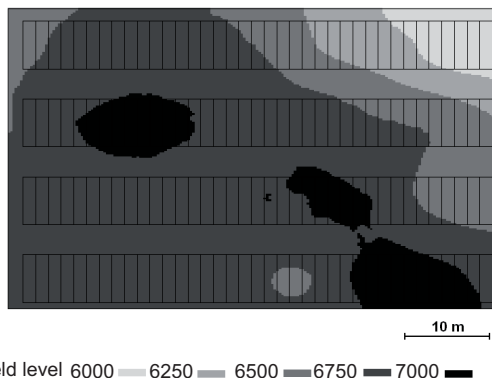


Fig. 2. Grain yield map including the outline of plots for trial at Mietoinen in 2004. Plots of one replication are in one row.

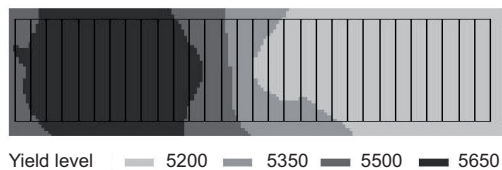


Fig. 3. Grain yield map of one replication for trial at Mikkeli in 2005. Yield level decreased almost 500 kg/ha during five plots.

Three of these 7 trials were conducted in 2002 (at Jokioinen, Mikkeli and Ylistaro), the year with the lowest grain yield level (Table 1). Experimental fields were very homogeneous in only two of these trials, but in the remaining five trials the random variation in the field was clearly above average. No other similarities between trials were found.

Block sizes of 3, 4 or 5 were the best for all sites. The advantage of incomplete block designs over the randomised complete block designs was highest at Mietoinen (28%) and lowest at Tuusula (0%, based on one trial) and Jokioinen (3%). A single trial with the highest advantage was found from Pälkäne (in 2001). When the trial was analysed using ANOVA for a randomised complete block designs the SED value of the trial was 427 kg/ha, the second highest SED in all 28 trials analysed using the same model. When the same trial was processed as an incomplete block designs with block size 5, the SED value fell to 191 kg/ha. This

was lower than the average for all 28 trials (=240 kg/ha).

The highest SED value (451 kg/ha) was reached at Ylistaro in 2002. Optimal block size for this trial was 3 plots in a block. In this case the SED value was 41% lower, at 273 kg/ha.

Both alternative strategies - the spatial mixed model technique and the incomplete block design - were found to account for the variation in the field, but the effect of the strategies varied among experimental sites (Table 2). The spatial mixed model was efficient in the trials at Ylistaro, where the difference between methods was 28–35%. However, only a slight difference was found at other sites. This could not be explained in terms of the sill, nugget and range parameters or the nugget/sill ratio.

In the analysis of randomized complete block designs SED was the same for all pairs of cultivars but in the incomplete block analysis it varied according to the superimposed blocking. The distribution of SED was not altogether symmetrical. The lowest SED values were found when the same pair of cultivars appeared several times in the same block. Fig. 4 demonstrates the shape of skewness at Mikkeli. The same shape repeated itself at all

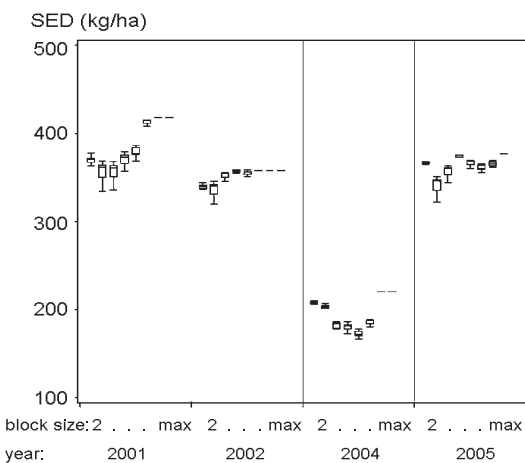


Fig. 4. Box and Whisker plots for the standard error of differences (SED). Plots are drawn for the trials set up at Mikkeli in 2001, 2002, 2004 and 2005. Each trial was analysed using block sizes of 2, 3, 4, 5, 6–7, 8–10, 14–20 and 25–32=max.

the locations. In addition, the standard deviation of SED was found to decrease as the block size increased. The sharpest fall occurred with block sizes 2 and 3 (Fig. 5). If the within field variation was small the variance component for the block effect was small, too, and the variation between the SED values was not practically important. In such cases the two alternative analyses resulted in well-matched rankings where the average variation in the ranking of a cultivar within one trial was only two positions (Fig. 6). On the other hand, if the within field variation was large the maximum change in the rankings could be more than 10 positions (Fig. 7). In a few cases the change could be more than 20 positions. For example, considering the trial Mietoinen 2003b, the complete block analysis ranked the cultivars Scarlett and Maaren at positions 6 and 23, respectively. When the same data were analysed using a superimposed incomplete block design with block size 3, Scarlett and Maaren were ranked 18 and 14, respectively. The latter analysis further adjusted the estimated difference between the mean yields of the two cultivars from 536 kg/ha to -120 kg/ha. The last result is in accordance with the results from official variety testing which propose an estimated difference of -202 kg/ha (Kangas et al. 2006).

Discussion

The precision of experimental results can be improved by keeping experimental error low. This can be done by: i) increasing the number of replicates in the experiment, ii) using uniform fields and managing them in the same way, iii) using experimental designs that control the variation within a replicate, iv) choosing an optimal field layout that reduces the variation within replicates, and v) using statistical models that take spatial variation into account. In cost-effective experiments the number of replicates in an experiment must be kept as low as possible, and so attention must be given instead to the other issues referred to above if improved results are to be obtained.

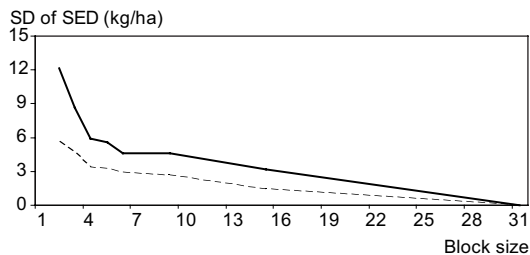


Fig. 5. Standard deviation of standard error of difference as a function of a block size. Dashed line is based on all trials whereas solid line is calculated using seven trials in which within field variation was the largest.

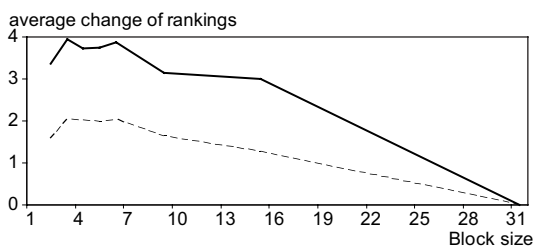


Fig. 6. Average change of rankings of cultivars as a function of a block size. Dashed line is based on all trials whereas solid line is calculated using seven trials in which within field variation was largest.

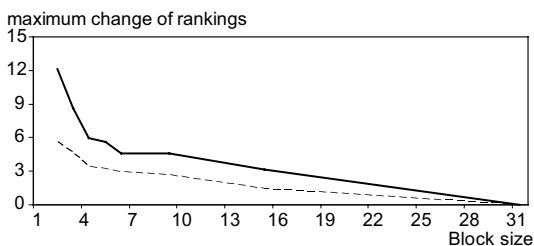


Fig. 7. Maximum change of rankings as a function of a block size. Dashed line is based on all trials whereas solid line is calculated using seven trials in which within field variation was largest.

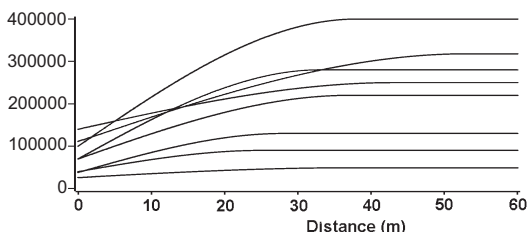


Fig. 8. Estimated spherical variogram for nine variety trials.

In this study we used the results obtained by Stroup (2002) showing the uniform superiority of incomplete block designs and spatial mixed model analysis in variety trials. These results were obtained by comparing the power of pairwise tests over a range of hypothetical spatial dependence patterns, experimental designs and yield recordings. Recommendable block sizes for practical testing situations can also be determined by applying power analysis. A possible strategy would be to use empirical data to estimate the extent of variability in practical test fields, compute the power of pairwise tests over the range of covariance structures implied by the estimated variability and then use designs that appear to be the most powerful. The data available to us, however, allowed a more direct approach. The characteristic feature of our data was that within each replication the blocks were arranged in a row. This made it possible to superimpose new blocking structures on the data and so study the effect of block size on the efficiency of testing pairwise yield differences.

The results obtained showed that the analytical efficiency of spatial mixed model techniques and incomplete block analysis with duly planned block sizes were well-matched. They also suggested that the ranking of a cultivar could vary up to 10–20 positions if an incomplete block design was used instead of a complete block design. These results are in accordance with the results obtained by Stroup (2002) whose research frame is very close to that of ours. Our results also comply with the conclusions of Lopez and Arrue (1995). Focusing on incomplete blocks of size 2 they report that the incomplete block design is, on average, 24 % more efficient than the complete block design. Grondona et al. (1996) have analysed variety trial data with nineteen different spatial models and compared the results with traditional incomplete block analysis. They report that spatial analysis is in general more efficient in reducing residual variation than incomplete block analysis although there was no one model that best fit all the trials. Our findings are in agreement with their studies, too.

We found that among the data for the different years, experimental sites and experimental fields there was a significant difference in the extent of

soil variation and the associated spatial dependence. Based on graphs such as in Fig. 8, we concluded that spatial dependence was of minor importance in some 20% of the trials. Recommendations on optimal blocking strategies were then made using data from the trials where spatial dependence was of major importance. The justification for this was that wherever there is little soil variation, experimental design is always of minor importance. In planning an experiment the extent and strength of spatial dependence are usually unclear, and so it is wise to be prepared for the worst and use designs that are robust against strong spatial variation.

It is clear that different response variables reflect the variation in the soil differently. In this paper we report results concerning grain yield only. We have, however, studied other traits determining the agricultural value of barley cultivars. The results show that in most cases the experimental designs that are efficient in testing pairwise differences of grain yields are also efficient in testing pairwise differences of many other traits. In particular, we have found that e.g. for the length of the stand the pattern of spatial dependence is very similar to that of grain yield and the same kind of incomplete block designs are therefore appropriate for studying both these traits.

The incomplete block designs are problematic because the frequency of occurrence of the same pair in the same incomplete block has an effect on the standard errors of the differences between the pairs. When the block size is small and the number of cultivars is large many pairs of cultivars never occur in the same block. Our results showed that this is a very relevant problem with block sizes 2 and 3. A straightforward precept to use incomplete block designs is therefore not justifiable. Yet, in variety testing all comparisons between the cultivars are not relevant. In many cases only differences between the old control cultivars and the new cultivars are of interest. So, the use of incomplete block designs with small block sizes like 2 or 3 can be recommended under the circumstances where the relative importance of different comparisons is known and can be taken into account in planning the experimental designs. Because trials were planned for a certain block sizes, it was not pos-

sible to control the concurrence of pairs of cultivars during the superimposition of new blocks. This increased the skewness of the distribution of SED and emphasised the significance of the original planning process. However, in the current study skewness was not a major issue all told.

In this study we concentrated only on variety trials of barley cultivars. The approach presented was possible because most of the operating procedures related to the variety trials on barley had been standardised during the long period of these trials. Official variety trials have been carried out in Finland for other cereals over many years, too. In addition to barley, we have variety trials data on spring wheat, winter wheat, oats and rye covering a period of more than 30 years. Like variety trials on barley these trials are also well-established and standardised. The same techniques applied in studying the optimal block sizes for barley trials could therefore be applied in finding optimal block sizes for variety trials on these cereal species.

Our experience shows that optimal block size is mainly related to spatial dependence and to a lesser extent to species, treatment types (e.g. variety, fertilization levels) or the number of levels of treatment. This makes it possible to generalise the results obtained in this study to other species and to other types of experiment. Traditionally, incomplete block designs have been used in MTT Agrifood Research Finland when the number of treatments is 16 or more. Similar practices are common in other research organisations. According to our results, incomplete block designs can also be recommended in experiments with 7 to 15 treatments.

During the course of this study we found that if an experiment with an incomplete block design is well planned for mitigating the effects of within-block heterogeneity, the spatial mixed model techniques and the conventional analysis of variance techniques have approximately the same efficiency in testing pairwise yield differences. In terms of the labour required the latter is less onerous. The principal conclusion from this study is therefore

as follows: Finnish official variety trials and many similar field trials can be conducted more efficiently by using smaller block sizes, and if appropriate blocking strategies are used there is no reason to change the conventional practice followed in statistical analysis.

References

- Brooker, P. 2001. Modelling spatial variability using soil profiles in the Riverland of South Australia. *Environmental International* 27: 121–126.
- Grondona, M.O., Crossa, J., Fox, P.N., & Pfeiffer, W.H., 1996. Analysis of variety yield trials using two-dimensional separable ARIMA processes. *Biometrics* 52: 763–770.
- Järvi, A., Kangas, A., Mattila, I., Mäkelä, L., Rahkonen, A., Salo, Y., Vuorinen, M., & Öfversten, J. 1998. Virallisten lajikekokeiden suoritusohjeet. [Guidelines for conducting official variety experiments]. *Maatalouden Tutkimuskeskuksen Julkaisuja. Sarja B.* 14, 71 pp.
- Kangas, A., Laine, A., Niskanen, M., Salo, Y., Vuorinen, M., Jauhiainen, L., & Nikander, H., 2006. Virallisten lajikekokeiden tulokset 1998–2005. Results of official variety trials 1998–2005. *MTT:n selvityksiä* 105, 210 pp.
- Kempton R.A., & Howes, C.W., 1981. The use of neighbouring plot values in the analysis of variety trials. *Applied Statistics* 30: 59–70.
- Kravchenko, A.N., Robertson, G.R., Snap, S.S., & Smucker, A.J.M., 2006. Using information about spatial variability to improve estimates of total soil carbon. *Agronomy Journal* 98 (3): 823–829.
- Lark, R. 2000. Designing sampling grids from imprecise information on soil variability, an approach based on the fuzzy kriging variance. *Geoderma* 98: 35–59.
- Lopez, M.V., & Arrue, J.L., 1995. Efficiency of an incomplete block design based on geostatistics for tillage experiments. *Soil Science Society of America journal* 59 (4): 1104–1111.
- Pannatier, Y. 1996. *VARIOWIN: Software for Spatial Data Analysis in 2D*, Springer-Verlag, New York, NY.
- Papadakis, J.S. 1937. Methode statistique pour des experiences sur champ. *Bulletin de l'Institut d'Amelioration des Plantes*, Thessalonike 23.
- SAS 1999. *SAS/STAT User's Guide*, Version 8, Cary, NC: SAS Institute Inc.
- Stroup, W.W. 2002. Power analysis based on spatial mixed effects analysis: a tool for comparing design and analysis strategies in the presence of spatial variability. *Journal of Agricultural Biological, and Environmental Statistics* 7, 491–511.
- Watson, S. 2000. Spatial dependence and block designs in spaced plant herbage trials. *Journal of Agricultural Science, Cambridge*, 245–258.

SELOSTUS

Suosittelava lohkon koko ohran lajikekokeissa

Lauri Jauhiainen, Jukka Öfversten ja Arjo Kangas
MTT Palveluyksikkö ja Kasvintuotannon tutkimus

Pellolla esiintyvä vaihtelu heikentää peltokokeista saatavien tulosten tarkkuutta. Tilanne on erityisen hankala vertailtavien käsittelyjen lukumäärän ollessa suuri, jolloin koekentän koko voi kasvaa niin suureksi, että kasvuolosuhteet saattavat olla kovinkin erilaiset eri koeruuduissa. Koeasetelman valinnalla on perinteisesti pyritty kontrolloimaan pellolla esiintyvää vaihtelua. Epätäydellisten lohkojen koeasetelmassa yhteen lohkoon valitaan vain osa vertailtavista käsittelyistä, kun taas täydellisten lohkojen koeasetelman mukaisesti suunnitellussa kokeessa kaikki vertailtavat käsittelyt sisältyvät jokaiseen lohkoon. Lohkomisen jälkeen vain lohkon sisäisellä kasvuolosuhteiden vaihtelulla on merkitystä, joten isoissa kokeissa epätäydellisten lohkojen koeasetelman on todettu johtavan tarkempiin tuloksiin kuin täydellisten lohkojen koeasetelman käytön. Käyttökelpoinen lohkon koko ei kuitenkaan ole tiedossa.

Lajikekokeet ovat tyypillisesti kokeita, joissa vertailtavien käsittelyjen eli lajikkeiden määrä on suuri, jopa useita kymmeniä lajikkeita. Tutkimuksessa käytetty aineisto sisälsi 28 virallista ohran lajikekoetta vuosilta 2001–2005. Tyypillisessä epätäydellisten lohkojen koeasetelman mukaisesti suunnitellussa ohran lajikekokeessa lohkot sisältävät 6–7 lajiketta, mutta tulosten perusteella vain 4–5 lajiketta kannattaisi sijoittaa samaan lohkoon. Tulosten perusteella voidaan olettaa epätäydellisten lohkojen koeasetelmista olevan hyötyä myös pienissä, jopa 7–10 käsittelyn peltokokeissa. Tulokset osoittavat lisäksi, että oikein käytettyinä epätäydellisten lohkojen koeasetelmien avulla pystytään huomioimaan pellolla esiintyvä sängen nopeakin vaihtelu niin hyvin, ettei tarvetta ole siirtyä käyttämään monimutkaisempia tilastollisia menetelmiä, joita ovat mm. koeruutujen välisen spatiaalisen riippuvuuden huomioivat sekamallit.