THE USE OF NEAREST NEIGHBOUR (NN) MODELS IN FIELD TRIALS

David Baird
Agricultural Research Division
Ministry of Agriculture and Fisheries
Lincoln

ABSTRACT

Wilkinson's recent method of NN analysis for field experiments and the topic of NN balanced designs are described and discussed, indicating the types of situations where this method may be useful. An example is presented that shows the improvements that may be obtained by this method. It is concluded that having the opportunity to apply an NN analysis provides an "insurance" against some of the factors that may lead to the standard statistical analyses being inefficient.

Additional Key Words: ***Randomised block design, trend effects, estimation of treatment effects, efficiency, bias.***

INTRODUCTION

Many of the field trials laid down in New Zealand are in the form of randomised block (RB) designs. Some of the reasons why this design is used so extensively are:

1. That the estimates of treatment effects and their standard errors are computationally simple to obtain.
2. That the design is easily understood and managed, simple and flexible.
3. That the blocking allows for more precise treatment estimates than a completely randomised design where the site is not uniform.
4. That the design has been shown by Fisher to give unbiassed estimates of variance components.

In using the RB design, several assumptions are made which may not be supported in the field. The departures from these assumptions will affect 4. above in particular. It is assumed that the blocks are homogeneous areas of experimental material. Although the RB design is fairly robust to departures from this assumption, if the area is far from being uniform then the stepwise detrending provided by blocking may not be sufficient to reduce the variance to an acceptable level, especially if the block size is large. The experimental errors are assumed to be independently distributed from a normal distribution and have common variance. If this is not the case, for example when the errors are correlated, then one can end up with poor estimates of treatment effects even though the treatment estimates are unbiassed over all the possible randomisations. Also there is a sizeable loss in the sensitivity of the tests of significance. This is frequently observed in field experiments where crop yields on neighbouring plots are positively correlated, i.e. strong or weak patches occur.

These deficiencies in the RB analysis should be overcome by taking into account any underlying patterns in the errors by the use of an appropriate model and so obtaining the best possible treatment estimates and tests of significance. Fisher himself comments "One can arrive at spurious or misleading inferences by failing to take into proper account all the relevant evidence". (Wilkinson *et al.*, 1983).

The Nearest Neighbour Models attempt to take into account certain forms of non uniformity within a trial and so arrive at more accurate and precise estimates of the treatment effects. They deal with the problem of neighbouring residuals being correlated which is often caused by trends only being partially removed by the blocking used in the experiment. To do this they provide a continuous detrending of plot yields in contrast to the stepwise detrending of blocks.

USE OF NN MODELS

In a field trial a variety of factors can lead to non uniformity of yields within a block after the treatment effects have been accounted for. NN Models can remove the effects of the following factors:

1. Trends across the site, for example fertility trends, trends caused by shelter or changes in soil depth or properties.
2. Localized differences not removed by blocking, for example the patchy occurrence of pests, diseases or weeds.

The need to use a NN Model can be indicated by higher variances than expected, residuals being correlated or showing some trend or pattern, yields of some areas being quite different from the overall run of results.

The use of NN Models is most effective when there are at least 3 replications and the block size is large (or equivalently there is a large number of treatments), viz. more than 40 plots in the total experiment. Also there should be a reasonable correlation between neighbouring residuals (i.e. for 3 or 4 replicates $r>0.4$, for 5 or more replicates $r>0.3$). If a NN Model is used when there is little correlation between neighbouring residuals, then a loss in accuracy can occur. In practice it would be advisable to use a NN Model only when there is some justification for its use, since the NN Models:

1. Have not yet been shown to be completely unbiassed.
2. Have the cost of a loss in error degrees of freedom which may not be met by an improvement in the reduction of variance.

THE NN MODELS

The first NN Model was suggested by Papadakis (1937) and consisted of a detrending of a plot yield by using the neighbours on either side to estimate the trend by linear interpolation. The analysis took the form of an analysis of covariance, where the covariate was formed, after the analysis of the RB design, from the mean of the residuals of the neighbouring plots. This method of analysis was extended by Bartlett (1978) from the simple form above to an interative form which used the adjusted treatment effects to calculate new residuals which were then used in a new analysis of covariance, this process being repeated until convergence was achieved i.e. the successive adjustments diminish to zero.

Wilkinson in his consideration of the simple form of Papadakis' method found it to be inefficient as it used the poor treatment estimates from the RB analysis to obtain the residuals and the residual mean square was biassed upwards. The iterated form of Papadakis is somewhat inappropriate as the residual mean square was biassed downwards. He has proposed a new method of analysis which he claims to be more efficient in removing trends from the data.

The underlying relationship in Wilkinson's NN model is that of a smooth trend plus error, identified by forming small incomplete blocks "after the event". The size of the small blocks formed depends on the number of neighbours used in the adjustment of plot values. Using first Nearest Neighbours corresponds to a block size of 3, and second Nearest Neighbours to a block size of 5. By moving the block positions, all the groupings of Nearest Neighbours are obtained. The analysis is performed in a way similar to that for an incomplete block design with recovery of interblock information. Thus with Wilkinson's method one ends up with a range of standard errors of differences (SED's), as comparisons within a block are more precise than comparisons between blocks. Hence the precision of a comparison between two treatments depends on how often the treatments occur as nearest neighbours.

The above models can be applied to a line of plots (a one dimensional model) or a rectangular array of plots (a two dimensional model). The two dimensional model is more complex than the one dimensional model as simultaneous adjustments in perpendicular directions must be made. Often an adjustment in only one direction of the rectangular array is needed. If the plot dimensions are far from square (e.g. when the plots are drill rows), then the residuals of the plots with the longer edges in common are most likely to show the greatest relationship. Thus a one dimensional adjustment across the narrower dimension of the plots will often be sufficient to remove any trends present.

DESIGN IMPLICATIONS

The range of SED's in Wilkinson's method is a problem in that the results are much more difficult to present and, if the treatments have a factorial structure, then the analysis of the main effects and interactions will not be entirely orthogonal. This, however, does not present a problem if the range is small. Another difficulty is that the adjustments of the edge plots for trend effects is poor as they have only one nearest neighbour.

To overcome the first problem Wilkinson has suggested that certain restrictions be placed on the randomisation of the RB design. The restrictions are:

1. No treatment should occur as it own neighbour.
2. That every treatment should have each other treatment as a neighbour equally often or, if this is not possible, that the neighbours up to n (n = 1 or 2) plots away must be all different from one another.

These two criteria for design make both Wilkinson's and Papadakis' analyses more efficient (Williams, 1952) and reduce the range of the SED's in Wilkinson's analysis.

The second problem can be overcome by the use of border plots around the edge of the design to give extra neighbours to the edge plots. The positioning of border plots depends on whether a one or two dimensional model is to be used, as plots only need to be placed on the edges perpendicular to the direction of adjustment. Thus, for a line of plots, only two border plots are needed, and for say, a 5 by 8 rectangular array of plots, 10 or 16 border plots are needed for a one dimensional adjustment (depending on the direction), or 30 border plots for a two dimensional adjustment.

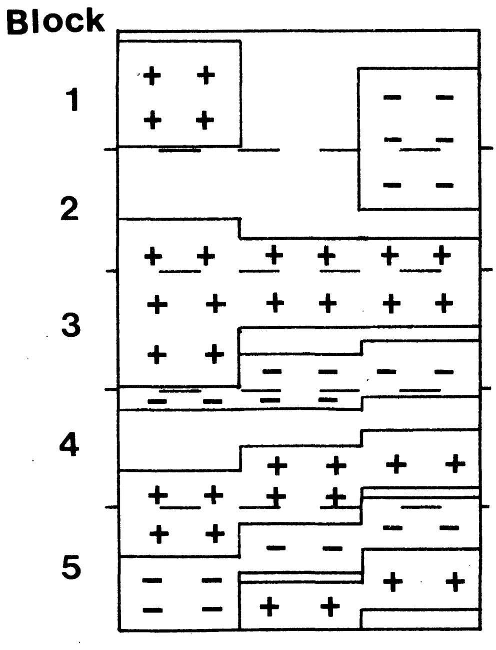

Figure 1: The residuals from an RB analysis of trial WM 8/64 which are greater than half a standard deviation from zero.

AN APPLICATION OF WILKINSON'S NN METHOD

A field trial (WM 8/64) for winter-sown wheat was laid down in 1982 at Winchmore Irrigation Research Station on border dyked land. The trial was laid down in 5 blocks with 27 treatments in each. The treatments had a 3 x 3 x 3 factorial structure in a split split plot design:

Level of Irrigation (Main plots)
1. Irrigated at 10% soil moisture.
2. Irrigated at 15% soil moisture.
3. Irrigated at 20% soil moisture.

Wheat Cultivar (Sub plots)
1. Karamu.
2. Oroua.
3. Rongotea.

Nitrogen Application (Sub sub plots)
1. No Nitrogen.
2. 50 kg/ha at tillering.
3. 50 kg/ha at booting.

The experiment was analysed in the normal way but a much higher sub sub plot coefficient of variation (CV = 15%) than usual (5 – 10%) was obtained. Also the treatment effects were very different in some cases from what was expected from previous work with the same cultivars at Winchmore. The treatment effects that had been anticipated were:

1. That irrigation at 15% and 20% soil moisture would give higher yields than that at 10% for all cultivars and nitrogen applications.
2. That all cultivars would be nitrogen responsive on this particular site.
3. That the form and the level of the irrigation response would be similar for applications of nitrogen at tillering or booting.

As can be seen in Figure 3, the treatment effects in some cases were out of line with these expectations.

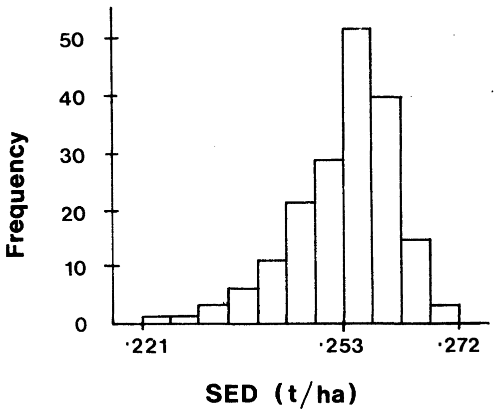

Figure 2: The histogram of Standard Errors of Differences from Wilkinson's NN analysis of WM 8/64.

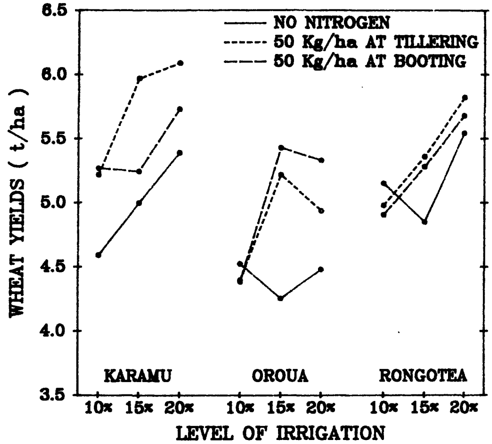

Figure 3: The crude treatment effects of WM 8/64.

The experimentalist involved pinpointed the shifting of the topsoil in the border dyking of the site as the cause of the irregularities in the yield. The trial was reanalysed as an RB and the residuals were scanned by flagging those more extreme than half their standard deviation. The areas of positive residuals were marked + and the areas of negative residuals –. As can be seen in Figure 1, a strong pattern emerged where the positive and negative residuals were highly aggregated. The correlation coefficient between residuals and the means of their neighbouring residuals was 0.81. Thus this experiment was a good candidate for analysis by a NN Model. Both Papadakis (Single and Iterated forms) and Wilkinson's analysis were applied but only the results from Wilkinson's analysis will be reported here. Wilkinson's NN analysis was performed with a Fortran program written and made available by Wilkinson and his colleagues and the Papadakis analysis was performed using a Genstat macro of R.A. Kempton. To simplify the comparisons between methods a mean SED will be presented but because the experiment was designed as a split split plot there are actually a variety of SED's, with the one being used depending on the comparison being made. The individual SED's in general follow the same pattern as their mean. The mean SED of the normal analysis was 0.434. It was comforting to note that the Iterated Papadakis and Wilkinson analyses gave almost identical estimates of treatment effects and also that the mean SED of the Wilkinson's analysis (0.253) lay between that of the Single form (0.268) and Iterated form (0.230) of Papadakis' analysis, as expected from Wilkinson's work.

The results of Wilkinson's analysis can be seen in Figure 4 and the range of SED's obtained in Figure 2. As this experiment was not designed with the design criteria of Wilkinson, the range of SED's in this case is not as small as could have been obtained otherwise. The form of the treatment effects is now in accordance with the expected effects as outlined previously.

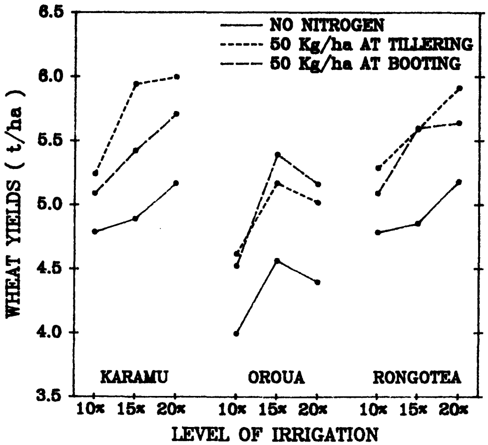

Figure 4: The estimated treatment effects from Wilkinson's NN analysis of WM 8/64.

DISCUSSION

There is some criticism of these NN methods in that they are used after the event and use the data iteratively. There is a fear that, in the great amount of "data massaging" done, a treatment estimate that is correct, but is not in accord with the ideas of the researcher, stands a 50% chance of being adjusted more into line with his thinking. Clearly using a decision rule to apply NN analysis if $r > 0.4$ and abiding by the results counters such criticism. These methods have been shown to be approximately unbiassed and work is still continuing to prove that they are theoretically as valid as the RB analysis. However, as in the RB analysis, these methods may be inefficient or lead to poor treatment estimates if the underlying model is incorrect.

The inclusion of border plots suggested by Wilkinson can markedly increase the size of the experiment (depending on the layout), especially for a two dimensional adjustment, by up to 100%. The return from the extra information contained in these results is doubtful, since the yields from the border plots are only used to remove trend effects and are not used in the estimation of treatment effects. Some statisticians insist that it would be a better use of extra plots to include another replicate in the experiment, rather than only to place these extra plots on the borders for adjustment of trend effects but this depends on site variability characteristics. The stronger the trends are within a site, the more efficient border plots become.

CONCLUSIONS

In the past, many of the analyses performed on RB designs have been inefficient as the underlying assumptions about spatial variation have not been correct. In some cases it is very difficult to be sure the blocks chosen before starting an experiment will be uniform. Because of inefficient blocking and a variety of other factors, one is going to have some experiments that turn out to be more variable than expected. In the past these experiments have returned little conclusive information and so much of the effort put into them has been wasted through inefficient statistical analyses. Now one has the option of performing a more efficient NN analysis (with the aid of a powerful computer) on the experiment. However, care should be taken to examine the appropriateness of applying a NN model to the experimental data as not all factors that cause unexplained variation will fit a NN model. Thus in some cases, precision that has been lost through the use of RB analysis can regained.

In the case of deciding to design an experiment with Wilkinson's criteria for NN balance, one should have a strong reason to believe that there are trends present within the site. If after being analysed, the trial does not show any unusual variation, the precision may be less than of a RB design. Although the normal RB analysis could still be applied, it would be inappropriate, as, due to the restricted randomisation, there will be a more precise estimate of treatment effects but a higher error variance and so a biassed F test. A more cautious approach in this situation would be to use a standard design with a higher degree of blocking (such as a latin square or an incomplete block design) and to use NN analysis if the level of blocking is still not sufficient to reduce the variance to an acceptable level.

NN analysis can be applied to many other types of design, including latin squares, lattice designs, and incomplete block designs, and may give improvements on these also if the blocking in the design does not give reasonably uniform blocks.

Thus, NN analysis provides a further tool which, when used judiciously, may remove the effects of some confounding factors, and so some experimental effort that would have been wasted can now be salvaged.

ACKNOWLEDGEMENTS

I am grateful to Mr C.B. Dyson for introducing me to this area of work and for his encouragement and help throughout. Thanks also to Mr E.G. Drewitt for the use of his trial data and Mr D.J. Saville for helpful discussions.

REFERENCES

Bartlett, M.S. 1978. Nearest neighbour models in the analysis of field experiments. *Journal of the Royal Statistical Society B 40:* 147-174.

Freeman, G.H. 1979. Some two-dimensional designs balanced for nearest neighbours. *Journal of the Royal Statistical Society B 41:* 88-95.

Papadakis, J.S. 1937. Methode statistique pour des experiences sur champ. *Bull. Inst. Amel. Plantes a Salonique No. 23.*

Wilkinson, G.N., Eckert, S.R., Hancock, T.W., Mayo, O. 1983. Nearest neighbour analysis of field experiments. *Journal of the Royal Statistical Society B 45:* 151-211.

Williams, R.M. 1952. Experimental designs for serially correlated observations. *Biometrika 39:* 151-167.