Payne, R. W. 2013. The design and analysis of long-term rotation experiments. *Agronomy Journal.* 105 (2), pp. 772-785.

Published July 12, 2013

Symposium: Statistical Concepts

# The Design and Analysis of Long-Term Rotation Experiments

Roger William Payne*


## ABSTRACT

**Rotation experiments are intended to compare different sequences of crop (and possibly husbandry) combinations. To avoid the conclusions being dependent on a specific sequence of years, it is advantageous to phase the start of the experiment, with new replicates of the rotations starting in successive years. Once a complete cycle has taken place, comparisons can then be made between the rotations in every subsequent year. If sufficient resources are available to have more than one replicate in each year, it will be possible to do an interim analysis with the data from a single year. Otherwise meaningful analyses will need several years' data and the assumption, e.g., that higher order interactions can be ignored or that responses over years can be modeled by low-order polynomials. Other analysis complications are that the within-year variances may be unequal and that the correlation between observations on a plot may differ according to the distance in time between them. The old-fashioned method of analysis, feasible if the data are balanced, would be to do a repeated-measurements analysis of variance. A more recent, and more satisfactory, alternative is to do a mixed model analysis by residual (restricted) maximum likelihood estimation, possibly fitting a model to the between-year correlation structure. The issues are illustrated using data from the Woburn Ley–Arable Experiment.**


Rotation experiments can play an important role in the study of alternative cropping systems, providing useful insights into the effects of the proposed new strategies in more realistic situations than a single year's trial. For example, it may take several years for the benefits or disadvantages of a new strategy of pest control to become apparent. Likewise, the yields of a particular crop cultivar may depend strongly on the previous cropping history of the field. They are also invaluable for study of the long-term effects of the systems on aspects, such as soil organic matter, that underpin agricultural sustainability.

The key aim of a rotation experiment is to compare different sequences of crop (and possibly husbandry) combinations. The separate crops in the sequence are usually called *courses*. In most situations these will occur at annual intervals, but the same principles apply with the shorter intervals that may be used, for example, in glass houses. Often not all of the crops are of practical interest. For example, a sequence may include a fallow year, where there is nothing to be measured or assessed, or it may include crops that form part of the treatment for a subsequent crop but are not themselves of any commercial interest. So *rotation experiment* here refers to experiments that aim to compare different rotations, following the example of Patterson (1964), who excluded the simpler *fixed rotation experiments* that study the effects of treatments on the crops of a single rotation. (These can involve some of the problems discussed in below but are inherently much easier to handle.)

To illustrate the concepts, Table 1 shows one block of an experiment by Glynne and Slope (1959), which was designed to assess the effects of previous cropping by bean (*Vicia faba* L.) or potato (*Solanum tuberosum* L.) on the incidence of eyespot (*Oculimacula yallundae* and *Oculimacula acuformis*) in winter wheat (*Triticum aestivum* L.). The crops in Years 1 and 2 are *treatment crops* that set the scene for Year 3, when *test crops* are grown to enable the differences between the sequences to be assessed. The two winter wheat crops in Year 2 also act as *partial test crops* in that they allow the effects from the wheat and potato crops in Year 1 to be assessed. No analysis was made of the bean or potato yields.

This particular experiment is a short-term (or fixed-cycle) rotation experiment in which the sequences run through one simultaneous cycle to compare the sequences in the final year. These can be designed and analyzed in much the same way as ordinary single-year field experiments. More interesting design and analysis problems are posed by long-term rotation experiments that are intended to run through several cycles and involve analyses of data from more than 1 yr.

## DESIGN ISSUES

In the simplest long-term experiments, the rotations are all of the same length and have the test crops at the same points in the cycle. The situation becomes much more complicated when rotations of different lengths are included or the test years do not coincide. Some of the issues are illustrated by the Woburn

VSN International, Hemel Hempstead, Hertfordshire HP1 1ES, U.K., and Department of Computational and Systems Biology, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, U.K. Received 26 Oct. 2012. Accepted 25 Apr. 2013. *Corresponding author (roger.payne@vsni.co.uk)

Published in Agron. J. 107:772–785 (2015)
doi:10.2134/agronj2012.0411



**Abbreviations:** REML, residual (restricted) maximum likelihood.

Table 1. One block from a short-term rotation experiment to study eyespot.

| Year | Type of crop | Crop† Plot 1 | Plot 2 | Plot 3 | Plot 4 |
|---|---|---|---|---|---|
| 1 | treatment | W | P | W | Be |
| 2 | treatment | W | W | P | P |
| 3 | test | W | W | W | W |

† Be, bean; P, potato; W, winter wheat.

Ley–Arable Experiment (Johnston, 1973; see Table 2). Initially, when this began in 1938, it was designed to compare the effects of four cropping systems, each lasting 3 yr, on soil fertility and the yields of two arable (test) crops in Years 4 and 5. The four rotation treatments were: a grass–clover ley, with 44% meadow fescue (*Festuca pratensis* Huds.), 44% timothy grass (*Phleum pratense* L.), and 12% white clover (*Trifolium repens* L.), given a little N fertilizer and grazed by sheep (R1); lucerne (*Medicago sativa* L.) cut for conservation (R2); potato, wheat, grass hay (50% meadow fescue and 50% timothy grass (R3); and potato, wheat, kale (*Brassica oleracea* L.) (R4).

The two test crops were potato followed by spring barley (*Hordeum vulgare* L.). It is usual to phase the start of these experiments, with new replicates of the rotations starting in successive years so that once a complete cycle has taken place, comparisons can be made between the rotations in every subsequent year. The advantages of this scheme were pointed out by Yates (1949), who noted that year-to-year variations mean that "[t]o obtain a proper measure of the effect of any treatment, therefore, it is necessary to repeat even 1-yr experiments in a number of years... The same holds for rotation experiments.... The first rule in the design of rotation experiments, therefore, is that such experiments should include all phases of the rotation." Within each year, comparisons will be made only between the plots with rotations that started at the same time. One straightforward and effective strategy is therefore to use a randomized-block design, with the rotations in each block all beginning at the same time. Year differences are then confounded with blocks, so no information is lost on the treatments. If sufficient resources are available to have more than one replicate block in each year, it will be possible to do an interim analysis with the data from a single year. Otherwise meaningful analyses will need several years' data and the assumption, for example, that some year × treatment interactions can be ignored (these degrees of freedom can then provide the residual). This is illustrated below in the analysis of the Woburn experiment, where no replicate blocks were included (i.e., there were five blocks corresponding to the 5 yr of the rotation cycles).

If there are many different rotations, the number of plots per block may become too large for the blocking to represent the fertility patterns in the field effectively. The rotations may then need to be partitioned into sets to be placed in separate blocks. To enable this to be done effectively, Patterson (1964) introduced the concept of *comparable rotations*: two rotations are defined to be comparable if they have at least 1 yr when they both grow the same test crop. Ideally, therefore, only rotations that are not comparable should be placed in different blocks, or if that is not feasible, it should be those that are least comparable that are allocated to different blocks. Comparability is the key issue to consider if the rotations are of different lengths or if the test years do not coincide. It may then be necessary to start the plots within a block in different years. The designer should construct a table showing the crop scheduled to be grown on each plot in each year and should check that there are years when sufficient plots are growing the same test crop in each block for meaningful analyses to be done.

Additional, auxiliary treatment factors can be incorporated in a similar way as in designs for ordinary single-year experiments. Initially, in the Woburn experiment, each plot was split into two subplots to study the effect of applying farmyard manure at 38 Mg ha$^{-1}$) between Years 3 and 4 of each cycle, i.e., just before the first test crop. The design, therefore, was a split-plot design with blocks in different years, rotations as the whole-plot factor, and farmyard manure as the subplot factor.

The Woburn experiment had a further complication: to study the effect of changing the treatment crops, each block contained a further four plots on which the sets of three treatment crops were changed in a sequence that alternated between the arable treatments (R3 and R4) and the ley treatments (R1 and R2). As a result, the complete system required 20 yr to complete one cycle (Johnston, 1973). An unrandomized plan showing a full cycle is shown in Table 2, which is based on Cochran (1939, Table XII). This emphasizes that is vital to get the design right or else considerable time and effort will be wasted. Analysis techniques may, of course, improve during the course of a 20-yr experiment, but the designer should simulate a specimen data set and check that the current best analysis strategy can deliver the required results.

Nonstatistical aspects must also be considered. For example, it is important to ensure that the plots are sufficiently large to avoid treatment effects spreading to adjacent plots, and cultivation techniques should aim to avoid the movement of soil across plot boundaries. Also, if the experiment is to continue through several series of rotations, it may be beneficial to be able to split the plots later in order to apply additional treatment factors. For a more detailed discussion of all these issues, see Dyke (1974, Ch. 7).

## ANALYSIS

The analysis of a long-term rotation experiment may use similar methods to those involved in ordinary field experiments, but there are some special issues to consider:

1. The results will be recorded from several different years, and these may show different amounts of random variation.
2. The same plot may be observed in several years, and, unless these are well separated, the results may show a nonuniform correlation structure, where correlations decline with increasing distance in time.
3. The effect of a crop may depend on where it is within the rotation cycle.
4. There may be no replication, other than over years.
5. Treatment effects may build up (or decline) during the period of the experiment.
6. Basal treatments (fertilizers, cultivation practice, pesticides, etc.) or even the precise makeup of the rotations themselves may have changed during the experiment to keep it relevant with current farming practices.

Table 2. Cropping sequences in the first 20 yr of the Woburn Ley–Arable Experiment. Test crops are in italics.

| | | Crop† | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Block | Plot | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1 | 1 | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* |
| | 2 | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* |
| | 3 | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* | *B* |
| | 4 | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* | *B* |
| | 5 | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K | *P* | *B* |
| | 6 | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H | *P* | *B* |
| | 7 | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* |
| | 8 | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* |
| 2 | 9 | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* |
| | 10 | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* |
| | 11 | B | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* |
| | 12 | B | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* |
| | 13 | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | H | *P* |
| | 14 | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | K | *P* |
| | 15 | B | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* |
| | 16 | B | P | W | K | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* |
| 3 | 17 | P | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ |
| | 18 | P | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ |
| | 19 | P | B | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H |
| | 20 | P | B | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K |
| | 21 | P | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K |
| | 22 | P | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H |
| | 23 | P | B | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ |
| | 24 | P | B | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ |
| 4 | 25 | B | P | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ |
| | 26 | B | P | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ |
| | 27 | B | P | B | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W |
| | 28 | B | P | B | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W |
| | 29 | B | P | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W |
| | 30 | B | P | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W |
| | 31 | B | P | B | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K | *P* | *B* | $L_1$ | $L_2$ |
| | 32 | B | P | B | P | W | K | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ |
| 5 | 33 | B | K | P | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | $L_1$ |
| | 34 | B | H | P | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | $Lu_1$ |
| | 35 | B | K | P | B | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P | W | H | *P* | *B* | P |
| | 36 | B | H | P | B | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P | W | K | *P* | *B* | P |
| | 37 | B | K | P | B | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | H | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P |
| | 38 | B | H | P | B | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | K | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P |
| | 39 | B | K | P | B | P | W | H | *P* | *B* | $L_1$ | $L_2$ | $L_3$ | *P* | *B* | P | W | K | *P* | *B* | $Lu_1$ |
| | 40 | B | H | P | B | P | W | K | *P* | *B* | $Lu_1$ | $Lu_2$ | $Lu_3$ | *P* | *B* | P | W | H | *P* | *B* | $L_1$ |

† B, barley (pre-experiment cropping); H, hay; K, kale; $L_1$, $L_2$ and $L_3$: first, second, and third year's ley, respectively; $Lu_1$, $Lu_2$ and $Lu_3$: first, second, and third year's lucerne, respectively; P, potato; W, winter wheat.

The first issue occurs in many meta analyses, whether combining data from several years or from several sites. The traditional way to handle this, in ordinary analysis of variance, would be to analyze the years separately, test for homogeneity of variance, e.g., by using the test of Bartlett (1937), and then, if necessary, weight the data from each year by the reciprocal of that year's residual variance. If, as in many rotation experiments, there is no within-year replication, this will not be possible. Fortunately, though, the more recent residual or restricted maximum likelihood (REML) method for the analysis of linear mixed models allows different residual variances for the years to be estimated during the combined analysis (see Patterson and Thompson, 1971; Gilmour et al., 1995). If there are additional random terms, for example whole plots in a split-plot design, their variance components may also differ from year to year. This too can be handled in a REML analysis (for an example, see Payne et al., 2012, Ch. 2).

The second issue might traditionally be handled by repeated-measures analysis of variance, which mitigates the effects of the nonuniform correlations by adjusting the numbers of degrees of freedom of the affected sums of squares (e.g., see Winer, 1962, p. 523, 594–599; Payne, 2012, Section 8.2.1). This is feasible if the design is balanced, i.e., if the same plots have been measured in each of the years for which data are available, as shown, e.g., by Christie et al. (2001), Liebman et al. (2008), and Barton et al. (2009). In long-term rotation experiments, however, different plots will usually have been measured in different sets of years, and the use of repeated-measures ANOVA becomes a difficult (if not impossible) task. Fortunately, here too, the new REML methodology provides a solution, with the ability to fit models to the correlations (e.g., see Gilmour et al., 1997); Galwey, 2006, Section 9.7; Littell et al., 2006, Ch. 5; Payne, 2006; Payne et al., 2012, Ch. 4).

For examples, see Singh et al. (1997), Singh and Jones (2002), Richter and Kroschewski (2006), and Machado et al. (2008).

The third issue is relevant if the same test crop is grown several times in a particular rotation. It can be resolved by doing a separate analysis for each instance of the test crop in the rotation cycle, as would be necessary if they were actually different crops. (The only sensible way to perform an analysis combining data from several different test crops would be to assign some measure such as economic value to each one, but these could be rather arbitrary and unlikely to remain constant through the whole experiment.) An alternative would be to include a factor for occurrence-within-cycle within a combined analysis.

Issues 4 and 5 can be more difficult to resolve. If there are many auxiliary treatment factors, it may be acceptable to use some of the higher order interactions among these and the rotation factor as the residual, i.e., to use the traditional approach, e.g., of treating second-order (and higher) interactions as the residual and then feeling justified if the analysis detects no significant first-order interactions. (This reasoning tends to be rather circular but can often be justified by experience from previous similar experiments.) A more easily justifiable variant of this approach was suggested by Patterson (1959) for an experiment that studied fertilizer response as one of the treatments. The fertilizer was applied at five different levels. On the assumption that the relationship between yield and fertilizer can be represented adequately by a second-order polynomial, the interactions between rotations and the cubic and quartic polynomials could therefore be used for the residual. (Under these circumstances, of course, it is arguable that it might have been safer to have had genuine replication and fewer levels of fertilizer, but this is the same issue that arises in any study of fertilizer response.) An alternative approach would be to model the year × treatment interaction. Again, all the standard methods are available. For example, echoing Patterson's ideas for the fertilizer-response curves, one might fit interactions between the treatments and polynomial effects of year. The linear year effects and their interactions would assess whether the effects are increasing or declining in any uniform way with time (Issue 5), but again the success of this strategy is dependent on the assumption of a low-order polynomial response. If this is not feasible, an alternative would be to fit spline functions over years. For example, Verbyla et al. (1999) described how to fit random smoothing splines in REML. Polynomial models, though, may be easier to explain.

Changes in basal treatments (Issue 6) may be handled by including additional factors in the analysis to indicate the underlying basal conditions applying on each year–plot observation. Ideally, they should not have affected differences between rotations or any of the other treatments; provided the changes have not been too frequent, it should be possible to check this by fitting the relevant interactions. It may be more difficult, however, to accommodate changes in the makeup of the rotations themselves within a single analysis. For example, the test crops may not have remained the same or the whole purpose of the experiment may have changed. In that case, the analysis will need to start afresh once a complete cycle of the new rotations has taken place.

A substantial change like this took place on the Woburn Ley–Arable Experiment in the 1970s, which led to the redefined rotations Lc3, Lc8, Ln3, Ln8, AB, and AF. The first change, in 1973, was to introduce leys lasting 8 yr to compare with the 3-yr leys. These were included because in the Woburn Organic Manuring experiment, on a very similar soil, 6-yr leys had given small measurable increases in soil organic matter (Mattingly et al., 1973). The 8-yr leys were assigned to the plots with the alternating rotations, which had provided very little information in the analysis of the data so far recorded. On these plots, therefore the alternating cycles of the previous rotations were replaced by 8 yr of ley treatments. The test crops would thus coincide with the corresponding 3-yr-ley plots on every second cycle. However, there had been a sufficient number of plots in the alternating rotations to allow for duplication of the 8-yr-ley treatment; the introduction of half of these was phased, thus allowing comparison with the 3-yr ley treatments (and with the plots in continuous arable treatments) in every cycle. The ley treatments themselves were also changed in order to compare an all-grass ley given fertilizer N (rotations Ln3 and Ln8, on plots that had been in the grazed ley, R1) and a grass–clover ley without fertilizer N (rotations Lc3 and Lc8, on plots that had been in lucerne–clover, R2). Then, in 1978, the 3-yr arable rotations were changed to become either spring barley–spring barley–bean (rotation AB following rotation R3) or bare fallow–bare fallow–bean (rotation AF following rotation R4). This change was intended to see whether there was less risk of take-all (*Gaeumannomyces graminis*) in wheat and barley following the latter rotation than the former. The first test crop, winter wheat, was grown in 1981, and further data are available for this crop for 1982 to 2000. The second test crop was spring barley during 1982 to 1991 and then winter rye during 1992 to 2001. An N fertilizer treatment was also introduced on the second test crop in 1973 and on the first test crop in 1976 by further dividing the plots for four different N levels: 0, 70, 140, and 210 kg ha$^{-1}$. (By then the farmyard manure treatment had been discontinued, and no residual effects of this treatment were being detected.) For expository purposes, however, to simplify the discussion and conclusions, the subplot structure is ignored in the analyses below.

Table 3 shows the yields of the first test crop, winter wheat, from 1981 to 2000, which will be used to illustrate the analysis. Here a different, Roman, numbering is used for the blocks to distinguish them from those in the unrandomized plan shown in Table 2. The random model simply contains years (confounded with blocks), while the fixed model contains rotations, N rate, polynomial effects of years, and their interactions. Initially, fourth-order polynomials are fitted, while we investigate whether there is evidence of unequal variation in the different years. So the higher order polynomials are assumed to be absent, and their degrees of freedom used to estimate the variances between and within years. Thus the variances are essentially being estimated by the deviations from the fitted fourth-order polynomials. The analysis—and conclusions—thus depend on the appropriateness of this assumption. However, some justification for the approach is given by the fact that the analyses find no evidence for the inclusion of either the cubic or the quartic terms in the model, i.e., a (relatively simple) quadratic relationship seems to hold.

Table 3. Yields of the first test crop, winter wheat grain, by year and block in the Ley–Arable Experiment at Woburn, 1981–2000.

| Rotation† | N rate | 1981 (III) | 1982 (V) | 1983 (IV) | 1984 (II) | 1985 (I) | 1986 (III) | 1987 (V) | 1988 (IV) | 1989 (II) | 1990 (I) | 1991 (III) | 1992 (V) | 1993 (IV) | 1994 (II) | 1995 (I) | 1996 (III) | 1997 (V) | 1998 (IV) | 1999 (II) | 2000 (I) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Test crop yield | | | | | | | | | | | | | | | | | | | |
| | kg ha$^{-1}$ | Mg ha$^{-1}$ at 85% dry matter | | | | | | | | | | | | | | | | | | | |
| AB | 0 | 3.84 | 4.47 | 4.11 | 3.66 | 2.39 | 4.17 | 4.39 | 2.98 | 1.16 | 1.47 | 4.48 | 6.31 | 3.11 | 0.93 | 1.30 | 1.19 | 1.58 | 3.21 | 0.00 | 1.45 |
| | 70 | 6.59 | 6.38 | 6.28 | 6.56 | 5.90 | 6.91 | 6.18 | 6.28 | 3.94 | 4.94 | 8.56 | 7.84 | 5.92 | 3.94 | 4.21 | 7.24 | 5.73 | 6.70 | 1.97 | 4.54 |
| | 140 | 7.49 | 7.82 | 8.70 | 7.74 | 7.76 | 7.21 | 6.75 | 6.77 | 4.58 | 5.83 | 9.94 | 7.21 | 5.89 | 4.04 | 4.35 | 7.80 | 7.37 | 9.35 | 3.44 | 4.52 |
| | 210 | 7.39 | 8.13 | 8.17 | 9.41 | 8.62 | 8.53 | 7.84 | 6.20 | 4.74 | 6.33 | 10.23 | 6.81 | 6.63 | 3.51 | 4.35 | 8.43 | 7.88 | 10.26 | 2.28 | 5.53 |
| AF | 0 | 3.06 | 4.30 | 3.76 | 4.28 | 2.03 | 4.08 | 3.02 | 3.09 | 2.80 | 1.38 | 3.46 | 3.82 | 2.86 | 1.80 | 1.27 | 0.65 | 2.40 | 2.52 | 0.52 | 0.96 |
| | 70 | 6.32 | 6.82 | 6.79 | 8.94 | 5.46 | 5.08 | 5.56 | 6.60 | 4.92 | 5.72 | 8.00 | 8.05 | 5.79 | 5.32 | 3.82 | 6.60 | 6.52 | 6.35 | 6.55 | 4.87 |
| | 140 | 7.61 | 8.16 | 8.50 | 9.12 | 7.72 | 6.32 | 6.60 | 6.63 | 5.17 | 6.30 | 9.75 | 8.21 | 6.72 | 8.08 | 4.60 | 7.69 | 9.25 | 8.80 | 7.53 | 6.28 |
| | 210 | 7.78 | 8.52 | 9.43 | 9.35 | 9.20 | 7.88 | 6.43 | 6.61 | 5.82 | 5.18 | 10.57 | 7.59 | 7.37 | 8.55 | 4.96 | 7.79 | 9.24 | 9.72 | 8.48 | 7.39 |
| Ln3 | 0 | 5.82 | 5.37 | 3.86 | 4.92 | 4.24 | 3.36 | 4.41 | 4.01 | 4.04 | 1.73 | 6.75 | 2.73 | 3.13 | 4.76 | 2.17 | 3.82 | 1.74 | 3.77 | 1.69 | 3.40 |
| | 70 | 7.52 | 7.91 | 7.05 | 7.66 | 7.26 | 5.65 | 6.55 | 6.77 | 5.94 | 4.94 | 8.85 | 6.47 | 5.40 | 6.16 | 5.01 | 7.19 | 3.83 | 7.13 | 6.58 | 7.06 |
| | 140 | 8.12 | 7.53 | 8.06 | 9.75 | 8.26 | 6.62 | 7.59 | 7.12 | 6.10 | 5.43 | 9.96 | 7.49 | 6.60 | 7.35 | 5.39 | 7.15 | 5.15 | 8.67 | 7.58 | 8.64 |
| | 210 | 7.40 | 8.46 | 8.28 | 10.35 | 9.69 | 6.05 | 7.13 | 6.14 | 6.04 | 6.17 | 10.41 | 7.26 | 6.52 | 7.14 | 5.79 | 8.41 | 5.02 | 9.62 | 7.83 | 8.71 |
| Ln8 | 0 | 4.71 | 5.55 | 4.45 | 5.46 | 4.07 | 4.68 | 4.80 | 4.34 | 3.77 | 2.62 | 5.94 | 4.19 | 3.42 | 3.64 | 2.52 | 6.37 | 2.53 | 4.97 | 3.42 | 3.42 |
| | 70 | 6.52 | 8.04 | 7.23 | 8.68 | 6.98 | 6.55 | 6.74 | 6.73 | 5.58 | 5.79 | 8.83 | 7.17 | 5.16 | 5.14 | 5.71 | 8.23 | 6.20 | 7.77 | 6.59 | 6.58 |
| | 140 | 8.03 | 8.27 | 7.48 | 9.20 | 8.39 | 7.20 | 7.86 | 7.46 | 5.56 | 5.08 | 9.64 | 7.54 | 6.47 | 7.00 | 5.36 | 8.77 | 6.93 | 9.21 | 8.26 | 7.22 |
| | 210 | 7.83 | 7.31 | 6.93 | 10.33 | 8.55 | 6.84 | 7.00 | 7.23 | 4.91 | 5.25 | 9.75 | 6.67 | 6.55 | 7.16 | 6.53 | 8.46 | 7.25 | 9.24 | 6.51 | 7.49 |
| Lc3 | 0 | 5.35 | 5.16 | 6.36 | 7.18 | 4.97 | 6.14 | 5.51 | 5.68 | 5.45 | 3.59 | 6.47 | 6.33 | 5.58 | 5.06 | 2.57 | 5.23 | 4.40 | 4.78 | 4.42 | 5.05 |
| | 70 | 6.70 | 7.81 | 9.67 | 11.06 | 7.64 | 7.15 | 7.24 | 7.39 | 6.28 | 6.06 | 9.37 | 7.48 | 7.01 | 6.00 | 5.70 | 7.76 | 7.70 | 7.48 | 7.27 | 8.24 |
| | 140 | 7.69 | 8.38 | 9.34 | 10.52 | 9.57 | 6.89 | 7.74 | 7.54 | 6.12 | 7.20 | 10.46 | 6.13 | 7.69 | 6.28 | 6.46 | 8.19 | 8.01 | 8.50 | 8.65 | 8.96 |
| | 210 | 7.53 | 7.40 | 8.40 | 9.88 | 8.84 | 6.20 | 7.61 | 7.51 | 5.81 | 6.42 | 10.48 | 4.79 | 7.91 | 7.50 | 5.78 | 8.67 | 8.30 | 8.75 | 9.54 | 10.33 |
| Lc8‡ | 0 | 6.47 | 6.56 | 7.39 | 7.51 | 4.44 | 6.09 | 5.26 | 5.26 | 4.91 | 3.31 | 6.08 | 7.11 | 6.08 | 3.46 | 3.52 | 5.73 | 4.10 | 4.11 | 1.79 | 4.31 |
| | 70 | 7.84 | 8.36 | 9.64 | 9.66 | 8.08 | 7.31 | 7.48 | 7.87 | 6.69 | 6.51 | 8.81 | 6.65 | 7.03 | 6.48 | 6.60 | 7.97 | 6.78 | 7.55 | 4.65 | 7.47 |
| | 140 | 7.98 | 8.60 | 8.80 | 11.04 | 8.76 | 6.85 | 8.31 | 6.94 | 6.39 | 6.65 | 9.63 | 6.45 | 7.20 | 6.07 | 6.36 | 8.48 | 7.36 | 9.01 | 5.54 | 8.95 |
| | 210 | 7.68 | 8.41 | 8.66 | 9.36 | 10.19 | 6.75 | 8.13 | 7.06 | 5.06 | 6.99 | 10.10 | 6.14 | 7.69 | 7.53 | 6.14 | 8.28 | 7.43 | 8.98 | 4.95 | 9.65 |

† AB, arable rotation with spring barley, spring barley, bean between each (winter wheat) test crop; AF, arable rotation with bare fallow, bare fallow, bean between each test crop; Ln3, 3-yr all-grass ley between each test crop; Ln8, 8-yr all-grass ley between each test crop (Ln8); Lc3, 3-yr grass–clover ley between each test crop; Lc8, 8-yr grass–clover ley between each test crop.

‡ Lc8 results are from first-cycle plots in 1981–1985 and 1991–1995 and second-cycle plots in 1986–1990 and 1996–2000.

Table 4. GenStat analysis of the data in Table 3: First analysis, do we need different variances in each year?

| Accumulated summary of REML random models | | | |
|---|---|---|---|
| | Deviance | Akaike Information Criterion | Random df |
| Constant variance | 2160.92 | 2164.92 | 2 |
| Different variance in each year | 2087.06 | 2129.06 | 21 |

Note: omits constant, –log(det($X'X$), that depends only on the fixed model.
Conclusion: Yes, we do we need different variances.

Table 5. GenStat analysis of the data in Table 3: Second analysis: can we simplify the fixed model?

| Tests for fixed effects | | | | | |
|---|---|---|---|---|---|
| Fixed term | Wald statistic | numerator df | *F* statistic | denominator df | *F* probability |
| Sequentially adding terms to fixed model | | | | | |
| Rotation | 294.57 | 5 | 58.91 | 7.0 | <0.001 |
| N | 1999.46 | 3 | 666.49 | 7.0 | <0.001 |
| LinYear | 2.88 | 1 | 2.88 | 14.9 | 0.110 |
| QuadYear | 0.85 | 1 | 0.85 | 14.8 | 0.371 |
| CubYear | 0.09 | 1 | 0.09 | 14.7 | 0.770 |
| QuartYear | 0.71 | 1 | 0.71 | 14.8 | 0.412 |
| Rotation.N | 128.11 | 15 | 8.54 | 7.0 | 0.004 |
| Rotation.LinYear | 36.65 | 5 | 7.33 | 5.0 | 0.024 |
| N.LinYear | 37.99 | 3 | 12.66 | 5.0 | 0.009 |
| Rotation.QuadYear | 14.41 | 5 | 2.88 | 2.7 | 0.224 |
| N.QuadYear | 26.60 | 3 | 8.87 | 2.7 | 0.064 |
| Rotation.CubYear | 14.38 | 5 | 2.88 | 4.3 | 0.154 |
| N.CubYear | 2.24 | 3 | 0.75 | 4.3 | 0.575 |
| Rotation.QuartYear | 12.78 | 5 | 2.56 | 8.1 | 0.113 |
| N.QuartYear | 10.44 | 3 | 3.48 | 8.1 | 0.069 |
| Rotation.N.LinYear | 9.72 | 15 | 0.65 | 5.0 | 0.764 |
| Rotation.N.QuadYear | 7.95 | 15 | 0.53 | 2.7 | 0.822 |
| Rotation.N.CubYear | 12.09 | 15 | 0.81 | 4.3 | 0.663 |
| Rotation.N.QuartYear | 19.16 | 15 | 1.28 | 8.1 | 0.374 |
| Dropping individual terms from full fixed model | | | | | |
| Rotation.N.LinYear | 8.53 | 15 | 0.57 | 5.0 | 0.818 |
| Rotation.N.QuadYear | 8.25 | 15 | 0.55 | 2.7 | 0.810 |
| Rotation.N.CubYear | 10.41 | 15 | 0.69 | 4.3 | 0.732 |
| Rotation.N.QuartYear | 19.16 | 15 | 1.28 | 8.1 | 0.374 |
| Conclusion: Drop the cubic and quartic polynomials. | | | | | |

Table 6. GenStat analysis of the data in Table 3: Third analysis, any further simplification of the fixed model?

| Tests for fixed effects | | | | | |
|---|---|---|---|---|---|
| Fixed term | Wald statistic | numerator df | *F* statistic | denominator df | *F* probability |
| Sequentially adding terms to fixed model | | | | | |
| Rotation | 276.73 | 5 | 55.35 | 54.5 | <0.001 |
| N | 1936.96 | 3 | 645.65 | 54.5 | <0.001 |
| LinYear | 3.07 | 1 | 3.07 | 17.1 | 0.097 |
| QuadYear | 0.90 | 1 | 0.90 | 16.9 | 0.355 |
| Rotation.N | 113.16 | 15 | 7.54 | 54.5 | <0.001 |
| Rotation.LinYear | 28.52 | 5 | 5.70 | 80.2 | <0.001 |
| N.LinYear | 29.44 | 3 | 9.81 | 80.2 | <0.001 |
| Rotation.QuadYear | 13.02 | 5 | 2.60 | 95.9 | 0.030 |
| N.QuadYear | 18.41 | 3 | 6.14 | 95.9 | <0.001 |
| Rotation.N.LinYear | 12.37 | 15 | 0.82 | 80.2 | 0.648 |
| Rotation.N.QuadYear | 9.38 | 15 | 0.63 | 95.9 | 0.847 |
| Dropping individual terms from full fixed model | | | | | |
| Rotation.N.LinYear | 7.43 | 15 | 0.50 | 80.2 | 0.936 |
| Rotation.N.QuadYear | 9.38 | 15 | 0.63 | 95.9 | 0.847 |
| Conclusion: Drop Rotation.N.LinYear and Rotation.N.QuadYear. | | | | | |

Table 7. Fourth analysis, final model.

Estimated variance components

| Random term | Component | SE |
|---|---|---|
| Year | 0.9993 | 0.3538 |

Residual model for each experiment

| Experiment factor:Year | Term | Model (order) | Parameter | Estimate | SE |
|---|---|---|---|---|---|
| 1981 | residual | identity | variance | 0.128 | 0.052 |
| 1982 | residual | identity | variance | 0.161 | 0.058 |
| 1983 | residual | identity | variance | 0.680 | 0.210 |
| 1984 | residual | identity | variance | 1.108 | 0.335 |
| 1985 | residual | identity | variance | 1.353 | 0.407 |
| 1986 | residual | identity | variance | 0.560 | 0.174 |
| 1987 | residual | identity | variance | 0.213 | 0.073 |
| 1988 | residual | identity | variance | 0.159 | 0.058 |
| 1989 | residual | identity | variance | 0.615 | 0.193 |
| 1990 | residual | identity | variance | 0.343 | 0.111 |
| 1991 | residual | identity | variance | 0.686 | 0.215 |
| 1992 | residual | identity | variance | 1.934 | 0.580 |
| 1993 | residual | identity | variance | 0.310 | 0.103 |
| 1994 | residual | identity | variance | 1.315 | 0.399 |
| 1995 | residual | identity | variance | 0.322 | 0.107 |
| 1996 | residual | identity | variance | 0.743 | 0.233 |
| 1997 | residual | identity | variance | 1.539 | 0.479 |
| 1998 | residual | identity | variance | 0.951 | 0.321 |
| 1999 | residual | identity | variance | 2.390 | 0.747 |
| 2000 | residual | identity | variance | 0.728 | 0.284 |

Tests for fixed effects

| Fixed term | Wald statistic | numerator df | *F* statistic | denominator df | *F* probability |
|---|---|---|---|---|---|
| Sequentially adding terms to fixed model | | | | | |
| Rotation | 292.22 | 5 | 58.44 | 189.4 | <0.001 |
| N | 1977.82 | 3 | 659.27 | 189.4 | <0.001 |
| LinYear | 3.05 | 1 | 3.05 | 17.1 | 0.099 |
| QuadYear | 0.90 | 1 | 0.90 | 16.9 | 0.356 |
| Rotation.N | 114.81 | 15 | 7.65 | 189.4 | <0.001 |
| Rotation.LinYear | 31.34 | 5 | 6.27 | 107.1 | <0.001 |
| N.LinYear | 35.08 | 3 | 11.69 | 107.1 | <0.001 |
| Rotation.QuadYear | 15.45 | 5 | 3.09 | 109.2 | 0.012 |
| N.QuadYear | 20.83 | 3 | 6.94 | 109.2 | <0.001 |
| Dropping individual terms from full fixed model | | | | | |
| Rotation.N | 114.81 | 15 | 7.65 | 189.4 | <0.001 |
| Rotation.LinYear | 28.94 | 5 | 5.79 | 107.1 | <0.001 |
| N.LinYear | 51.10 | 3 | 17.03 | 107.1 | <0.001 |
| Rotation.QuadYear | 15.45 | 5 | 3.09 | 109.2 | 0.012 |
| N.QuadYear | 20.83 | 3 | 6.94 | 109.2 | <0.001 |

Tables 4 through 8 show the output from a sequence of analyses by GenStat, release 15 (VSN International, GenStat.co.uk). Observations on the same subplots occurred only every 5 yr, and so it seems reasonable to assume a uniform (or constant) correlation structure for the repeated observations from each subplot. The first analysis (Table 4) compares the models with constant and nonconstant within-year variances. In linear mixed models, it is customary to assess the random model by examining its deviance, which is defined as −2 times the log-likelihood for the model. In this case, the second random model, with a different variance in each year, is a generalization of the first one, and so the difference between their deviances can be treated as a chi-square statistic. If neither random model is a generalization of the other, Akaike or Bayesian information criteria are generally used to assess which one to select; for a practical example, see Kehel et al. (2010). The difference in the deviances, 73.86 on 19 degrees of freedom, shows that there is strong evidence that the variances within years are not constant.

Once the appropriate random model has been decided, the treatment model can be assessed to see whether there are any unnecessary fixed terms. The standard way to do this is to examine their Wald statistics. These would have exact chi-square distributions if the variance parameters were known, but, because those must be estimated, the statistics are only asymptotically distributed as chi-square. In practical terms, the chi-square values will be reliable if the residual degrees of freedom for the fixed term is large compared with its own degrees of freedom. Alternatively, statistical software systems such as ASReml, GenStat, and SAS use the method of Kenward and Roger (1997) to obtain an estimate of the number of residual degrees of freedom relevant to each term, so that an *F* statistic can be used instead. The *F* statistic is

Table 8. Predicted mean yield of winter wheat under four N application rates from the GenStat analysis of the data in Table 3.

| Year | Rotation† | Predicted mean yield‡ | | | |
|---|---|---|---|---|---|
| | | 0 kg ha$^{-1}$ | 70 kg ha$^{-1}$ | 140 kg ha$^{-1}$ | 210 kg ha$^{-1}$ |
| | | Mg ha$^{-1}$ at 85% dry matter | | | |
| 1981 | AB | 4.517 | 7.038 | 8.054 | 8.190 |
| | AF | 4.153 | 7.057 | 8.305 | 8.442 |
| | Lc3 | 6.002 | 7.763 | 8.520 | 8.123 |
| | Lc8 | 6.782 | 8.610 | 8.936 | 8.741 |
| | Ln3 | 5.772 | 7.961 | 8.787 | 8.626 |
| | Ln8 | 5.584 | 7.633 | 8.546 | 8.185 |
| 1982 | AB | 4.400 | 6.945 | 7.866 | 7.999 |
| | AF | 3.986 | 6.914 | 8.068 | 8.201 |
| | Lc3 | 5.919 | 7.704 | 8.366 | 7.966 |
| | Lc8 | 6.599 | 8.451 | 8.681 | 8.483 |
| | Ln3 | 5.489 | 7.702 | 8.433 | 8.268 |
| | Ln8 | 5.403 | 7.476 | 8.294 | 7.930 |
| 1983 | AB | 4.275 | 6.848 | 7.687 | 7.821 |
| | AF | 3.823 | 6.778 | 7.850 | 7.984 |
| | Lc3 | 5.836 | 7.649 | 8.229 | 7.829 |
| | Lc8 | 6.421 | 8.301 | 8.449 | 8.252 |
| | Ln3 | 5.223 | 7.464 | 8.113 | 7.949 |
| | Ln8 | 5.229 | 7.330 | 8.067 | 7.703 |
| 1984 | AB | 4.141 | 6.747 | 7.516 | 7.655 |
| | AF | 3.661 | 6.650 | 7.651 | 7.791 |
| | Lc3 | 5.752 | 7.598 | 8.108 | 7.715 |
| | Lc8 | 6.249 | 8.161 | 8.240 | 8.049 |
| | Ln3 | 4.973 | 7.247 | 7.826 | 7.668 |
| | Ln8 | 5.063 | 7.196 | 7.863 | 7.505 |
| 1985 | AB | 3.999 | 6.641 | 7.353 | 7.503 |
| | AF | 3.502 | 6.527 | 7.472 | 7.623 |
| | Lc3 | 5.668 | 7.551 | 8.004 | 7.621 |
| | Lc8 | 6.081 | 8.030 | 8.052 | 7.872 |
| | Ln3 | 4.740 | 7.050 | 7.573 | 7.425 |
| | Ln8 | 4.904 | 7.074 | 7.684 | 7.337 |
| 1986 | AB | 3.847 | 6.531 | 7.199 | 7.364 |
| | AF | 3.345 | 6.412 | 7.313 | 7.478 |
| | Lc3 | 5.584 | 7.508 | 7.917 | 7.549 |
| | Lc8 | 5.919 | 7.910 | 7.887 | 7.722 |
| | Ln3 | 4.524 | 6.875 | 7.353 | 7.221 |
| | Ln8 | 4.752 | 6.964 | 7.529 | 7.197 |
| 1987 | AB | 3.687 | 6.417 | 7.053 | 7.238 |
| | AF | 3.191 | 6.303 | 7.172 | 7.358 |
| | Lc3 | 5.500 | 7.469 | 7.847 | 7.499 |
| | Lc8 | 5.763 | 7.799 | 7.745 | 7.599 |
| | Ln3 | 4.323 | 6.721 | 7.167 | 7.055 |
| | Ln8 | 4.607 | 6.865 | 7.398 | 7.086 |
| 1988 | AB | 3.519 | 6.298 | 6.915 | 7.125 |
| | AF | 3.039 | 6.202 | 7.051 | 7.262 |
| | Lc3 | 5.415 | 7.435 | 7.793 | 7.470 |
| | Lc8 | 5.611 | 7.697 | 7.624 | 7.503 |
| | Ln3 | 4.140 | 6.587 | 7.015 | 6.927 |
| | Ln8 | 4.470 | 6.777 | 7.292 | 7.004 |
| 1989 | AB | 3.341 | 6.175 | 6.786 | 7.025 |
| | AF | 2.890 | 6.107 | 6.950 | 7.189 |
| | Lc3 | 5.330 | 7.404 | 7.756 | 7.462 |
| | Lc8 | 5.465 | 7.606 | 7.526 | 7.435 |
| | Ln3 | 3.973 | 6.475 | 6.896 | 6.838 |
| | Ln8 | 4.340 | 6.702 | 7.210 | 6.952 |
| 1990 | AB | 3.155 | 6.048 | 6.665 | 6.938 |
| | AF | 2.742 | 6.018 | 6.868 | 7.142 |
| | Lc3 | 5.245 | 7.378 | 7.736 | 7.476 |
| | Lc8 | 5.324 | 7.523 | 7.450 | 7.393 |
| | Ln3 | 3.823 | 6.383 | 6.811 | 6.786 |
| | Ln8 | 4.217 | 6.638 | 7.152 | 6.928 |
| 1991 | AB | 2.960 | 5.917 | 6.552 | 6.864 |
| | AF | 2.598 | 5.937 | 6.806 | 7.118 |
| | Lc3 | 5.159 | 7.355 | 7.733 | 7.512 |
| | Lc8 | 5.188 | 7.451 | 7.397 | 7.378 |
| | Ln3 | 3.689 | 6.313 | 6.759 | 6.773 |
| | Ln8 | 4.101 | 6.585 | 7.119 | 6.933 |

(continued)

Table 8. Continued.

| Year | Rotation† | Predicted mean yield‡ | | | |
|---|---|---|---|---|---|
| | | 0 kg ha$^{-1}$ | 70 kg ha$^{-1}$ | 140 kg ha$^{-1}$ | 210 kg ha$^{-1}$ |
| 1992 | AB | 2.757 | 5.781 | 6.448 | 6.804 |
| | AF | 2.455 | 5.862 | 6.762 | 7.118 |
| | Lc3 | 5.073 | 7.337 | 7.746 | 7.568 |
| | Lc8 | 5.058 | 7.388 | 7.365 | 7.390 |
| | Ln3 | 3.571 | 6.263 | 6.741 | 6.799 |
| | Ln8 | 3.992 | 6.544 | 7.109 | 6.967 |
| 1993 | AB | 2.545 | 5.640 | 6.352 | 6.756 |
| | AF | 2.316 | 5.794 | 6.739 | 7.142 |
| | Lc3 | 4.987 | 7.323 | 7.776 | 7.647 |
| | Lc8 | 4.933 | 7.335 | 7.357 | 7.429 |
| | Ln3 | 3.471 | 6.234 | 6.757 | 6.862 |
| | Ln8 | 3.891 | 6.515 | 7.124 | 7.030 |
| 1994 | AB | 2.324 | 5.496 | 6.264 | 6.721 |
| | AF | 2.178 | 5.733 | 6.735 | 7.191 |
| | Lc3 | 4.901 | 7.313 | 7.823 | 7.746 |
| | Lc8 | 4.813 | 7.292 | 7.370 | 7.496 |
| | Ln3 | 3.386 | 6.227 | 6.806 | 6.964 |
| | Ln8 | 3.797 | 6.498 | 7.164 | 7.122 |
| 1995 | AB | 2.094 | 5.347 | 6.185 | 6.699 |
| | AF | 2.043 | 5.679 | 6.750 | 7.264 |
| | Lc3 | 4.814 | 7.307 | 7.886 | 7.867 |
| | Lc8 | 4.698 | 7.258 | 7.406 | 7.589 |
| | Ln3 | 3.319 | 6.240 | 6.888 | 7.105 |
| | Ln8 | 3.711 | 6.492 | 7.227 | 7.244 |
| 1996 | AB | 1.856 | 5.194 | 6.114 | 6.690 |
| | AF | 1.910 | 5.631 | 6.784 | 7.361 |
| | Lc3 | 4.727 | 7.305 | 7.967 | 8.010 |
| | Lc8 | 4.589 | 7.234 | 7.464 | 7.709 |
| | Ln3 | 3.268 | 6.274 | 7.005 | 7.283 |
| | Ln8 | 3.631 | 6.497 | 7.315 | 7.394 |
| 1997 | AB | 1.609 | 5.036 | 6.051 | 6.694 |
| | AF | 1.780 | 5.591 | 6.838 | 7.482 |
| | Lc3 | 4.639 | 7.307 | 8.064 | 8.174 |
| | Lc8 | 4.485 | 7.219 | 7.544 | 7.856 |
| | Ln3 | 3.233 | 6.329 | 7.154 | 7.500 |
| | Ln8 | 3.559 | 6.515 | 7.427 | 7.573 |
| 1998 | AB | 1.353 | 4.875 | 5.997 | 6.712 |
| | AF | 1.652 | 5.557 | 6.912 | 7.627 |
| | Lc3 | 4.552 | 7.314 | 8.178 | 8.359 |
| | Lc8 | 4.386 | 7.215 | 7.647 | 8.031 |
| | Ln3 | 3.215 | 6.405 | 7.338 | 7.755 |
| | Ln8 | 3.494 | 6.544 | 7.564 | 7.781 |
| 1999 | AB | 1.088 | 4.709 | 5.951 | 6.742 |
| | AF | 1.526 | 5.530 | 7.005 | 7.796 |
| | Lc3 | 4.464 | 7.324 | 8.308 | 8.566 |
| | Lc8 | 4.292 | 7.219 | 7.772 | 8.232 |
| | Ln3 | 3.214 | 6.502 | 7.555 | 8.048 |
| | Ln8 | 3.436 | 6.584 | 7.724 | 8.018 |
| 2000 | AB | 0.815 | 4.538 | 5.913 | 6.785 |
| | AF | 1.403 | 5.509 | 7.117 | 7.989 |
| | Lc3 | 4.376 | 7.339 | 8.455 | 8.794 |
| | Lc8 | 4.204 | 7.234 | 7.919 | 8.460 |
| | Ln3 | 3.229 | 6.620 | 7.805 | 8.380 |
| | Ln8 | 3.386 | 6.637 | 7.909 | 8.284 |

† AB, arable rotation with spring barley, spring barley, bean between each (winter wheat) test crop; AF, arable rotation with bare fallow, bare fallow, bean between each test crop; Lc3, 3-yr grass–clover ley between each test crop; Lc8, 8-yr grass–clover ley between each test crop; Ln3, 3-yr all-grass ley between each test crop; Ln8, 8-yr all-grass ley between each test crop (Ln8).

‡ Standard errors of differences: average, 0.4946; maximum, 0.9009; minimum, 0.0431.

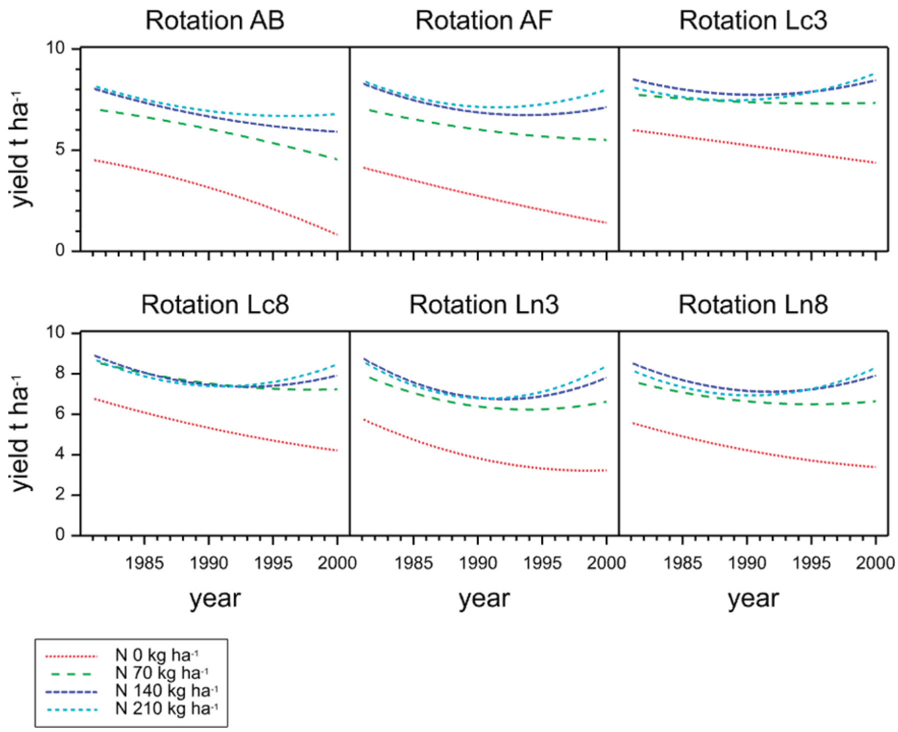


Fig. 1. Predicted mean trends of winter wheat yields under four N application rates from the Woburn Ley–Arable Experiment from 1981 to 2000. Rotations included: arable rotation with spring barley, spring barley, bean between each (winter wheat) test crop (AB); arable rotation with bare fallow, bare fallow, bean between each test crop (AF); 3-yr grass–clover ley between each test crop (Lc3); 8-yr grass–clover ley between each test crop (Lc8); 3-yr all-grass ley between each test crop (Ln3); and 8-yr all-grass ley between each test crop (Ln8).

equal to the Wald statistic divided by the number of degrees of freedom of the term (i.e., the number of degrees of freedom for the numerator of the $F$ test). It is usable, however, only if the number of residual degrees of freedom of the term (i.e., the degrees of freedom for the denominator of the $F$ test) is also known. Simulations performed by Kenward and Roger (1997) showed that the $F$ statistics are not subject to the biases of the chi-square statistics and can thus be used with confidence.

The second analysis, in Table 5, shows the $F$ (and Wald) tests for the fixed effects with the unequal variance model. There is no evidence that there are either cubic or quartic effects of year, so the third analysis (Table 6) fits a simpler model, containing only linear and quadratic effects of year. It is clear now that neither of the three-way interactions is needed, and so the fourth (and final) analysis can be produced with those interactions omitted. The variance components for the years range from 0.128 to 2.390, reinforcing the earlier conclusion that it would not be appropriate to assume a constant variance.

The final model (Table 7) shows a complicated picture, with linear and quadratic trends over years that depend on both rotation and N treatments. It is not very easy to pick up the pattern of responses from the predicted means (Table 8), so they are also plotted in Fig. 1.

The curves for the ley rotations are higher than those for the arable rotations, showing that they have been more effective in maintaining yields; however, some N fertilizer is also necessary because the curves for 0 N are well below those of the other N levels.

## DISCUSSION

The design and analysis of long-term rotation experiments present many interesting statistical challenges, which can be rather daunting if we consider the amount of time and effort that is at stake. A clear awareness of the special issues discussed above should avoid pitfalls, however, and lead to clear and useful conclusions. On design, it is important to allow comparisons to be made between the rotations in several years. The key issues are to decide how to allocate the rotation treatment and how to allow for several occurrences of the rotations, starting in successive years. We have shown above that one effective method is to start occurrences of the rotations in successive years, each in a separate block so that we have a randomized-block design, with year-of-starting-point as the block factor and rotation as the plots factor. Other, auxiliary treatments can be applied, as in conventional single-year experiments, for example, by splitting the plots into subplots to give a split-plot design with the auxiliary treatment factor(s) as the split-plot factor(s).

This may lead to a rather large experiment, with insufficient resources to provide within-year treatment replication, so strategies are needed in the analysis to provide degrees of freedom for the residual. Again, these can be drawn from the armory of conventional analysis techniques, for example, by allocating higher order interactions to the residual or by modeling responses by low-order polynomials or splines. Other aspects to consider include the fact that variability is unlikely to be constant across years, and there may be nonuniform correlation structures if there is an insufficient distance in time between the observations on each plot. These problems mean that conventional analysis of variance is unlikely to be suitable;

however, appropriate analyses can be performed using the more recent REML methodology (Patterson and Thompson, 1971; Gilmour et al., 1995, 1997; Littell et al., 2006; Payne et al., 2012). The GenStat commands that performed the analyses above are listed in Appendix 1. Equivalent R commands, using ASReml-R, are in Appendix 2. A SAS program for the final analysis is in Appendix 3. There are some differences in the tests for fixed effects, which are believed to arise from the use of average Fisher information by ASReml and GenStat in the Kenward and Roger (1997) calculations (see Gilmour et al., 1995); however, the analyses lead to the same conclusions. Note though, that the more complicated analysis, taking account of the subplot structure of the experiment, requires a rather more complicated response model over years.

More details about GenStat and ASReml-R can be found at www.vsni.co.uk.

## ACKNOWLEDGMENTS

I would like to thank Paul Poulton of Rothamsted for invaluable advice about the design and history of the Woburn Ley–Arable Experiment and for checking the biological details in the paper (any errors, though, are my own). I thank Sue Welham of VSN International for the R and ASReml-R script, and Nick Galwey and James Roger of London School of Hygiene and Tropical Medicine for the SAS program. I am also grateful to VSN International and Rothamsted for their support for the research. Finally, a special acknowledgement is due to Desmond Patterson, who passed away on 12 Mar. 2013. He and Robin Thompson were the original authors of the REML methodology on which so many of our analyses now depend.

## REFERENCES

Bartlett, M.S. 1937. Some examples of statistical methods of research in agriculture and applied biology. J. R. Stat. Soc. Suppl. 4:137–183.

Barton, L., G.G.Y. Wan, R.P. Buck, and T.D. Colmer. 2009. Nitrogen increases evapotranspiration and growth of a warm-season turfgrass. Agron. J. 101:17–24. doi:10.2134/agronj2008.0078

Christie, P., D.L. Easson, J.R. Picton, and S.C.P. Love. 2001. Agronomic value of alkaline-stabilized sewage biosolids for spring barley. Agron. J. 93:144–151. doi:10.2134/agronj2001.931144x

Cochran, W.G. 1939. Long-term agricultural experiments. J. R. Stat. Soc. (Suppl.) 6:104–148. doi:10.2307/2983686

Dyke, G.V. 1974. Comparative experiments with field crops. Butterworths, London.

Galwey, N.W. 2006. Introduction to mixed modelling: Beyond regression and analysis of variance. John Wiley & Sons, Chichester, UK.

Gilmour, A.R., B.R. Cullis, and A.P. Verbyla. 1997. Accounting for natural extraneous variation in the analysis of field experiments. J. Agric. Biol. Environ. Stat. 2:269–273. doi:10.2307/1400446

Gilmour, A.R., R. Thompson, and B.R. Cullis. 1995. Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. Biometrics 51:1440–1450. doi:10.2307/2533274

Glynne, M.D., and D. Slope. 1959. Effects of previous wheat crops, seed-rate and nitrogen on eyespot, take-all, weeds and yields of two varieties of winter wheat: Field experiment 1954–56. Ann. Appl. Biol. 47:187–199. doi:10.1111/j.1744-7348.1959.tb02536.x

Johnston, A.E. 1973. The effects of ley and arable cropping systems on the amounts of soil organic matter in the Rothamsted and Woburn Ley–Arable experiments. In: Rothamsted Experimental Station Report for 1972. Part 2. Rothamsted Exp. Stn., Harpenden, UK. p. 131–159.

Kehel, Z., D.Z. Habash, S.A. Gezan, S.J. Welham, and M.M. Nachit. 2010. Estimation of spatial trend and automatic model selection in augmented designs. Agron. J. 102:1542–1552. doi:10.2134/agronj2010.0175

Kenward, M.G., and J.H. Roger. 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997. doi:10.2307/2533558

Liebman, M., L.R. Gibson, D.N. Sundberg, A.H. Heggenstaller, P.R. Westerman, C.A. Chase, et al. 2008. Agronomic and economic performance characteristics of conventional and low-external-input cropping systems in the central Corn Belt. Agron. J. 100:600–610. doi:10.2134/agronj2007.0222

Littell, R.C., G.A. Milliken, W.W. Stroup, R.D. Wolfinger, and O. Schabenberger. 2006. SAS for mixed models. 2nd ed. SAS Inst., Cary, NC.

Machado, S., S. Petrie, K. Rhinhart, and R.E. Ramig. 2008. Tillage effects on water use and grain yield of winter wheat and green pea in rotation. Agron. J. 100:154–162. doi:10.2134/agrojnl2006.0218

Mattingly, G.E.G., M. Chater, and P.R. Poulton. 1973. The Woburn Organic Manuring Experiment 1964–72. In: Rothamsted Experimental Station Report for 1973. Part 2. Rothamsted Exp. Stn., Harpenden, UK. p. 98–151.

Patterson, H.D. 1959. The analysis of a non-replicated experiment involving a single four-course rotation of crops. Biometrics 15:30–59.

Patterson, H.D. 1964. Theory of cyclic rotation experiments. J. R. Stat. Soc. B 26:1–45.

Patterson, H.D., and R. Thompson. 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554. doi:10.1093/biomet/58.3.545

Payne, R.W. 2006. New and traditional methods for the analysis of unreplicated experiments. Crop Sci. 46:2476–2481. doi:10.2135/cropsci2006.04.0273

Payne, R.W., editor. 2012. The guide to the GenStat command language (Release 15). Part 2. Statistics. VSN Int., Hemel Hempstead, UK.

Payne, R., S. Welham, and S. Harding. 2012. A Guide to REML in GenStat. 15th ed. VSN Int., Hemel Hempstead, UK.

Richter, Ch., and B. Kroschewski. 2006. Analysis of a long-term experiment with repeated measurement models. J. Agron. Crop Sci. 192:55–71. doi:10.1111/j.1439-037X.2006.00167.x

Singh, M., S. Christiansen, and B.K. Chakraborty. 1997. An assessment of the effect of covariances of plot errors over time on the precision of means of rotations with wheat. Exp. Agric. 33:469–475. doi:10.1017/S0014479797004018

Singh, M., and M.J. Jones. 2002. Modeling yield sustainability for different rotations in long-term barley trials. J. Agric. Biol. Environ. Stat. 7:525–535. doi:10.1198/108571102744

Verbyla, A.P., B.R. Cullis, M.G. Kenward, and S.J. Welham. 1999. The analysis of designed experiments and longitudinal data using smoothing splines (with discussion). Appl. Stat. 48:269–311.

Winer, B.J. 1962. Statistical principles in experimental design. McGraw-Hill, New York.

Yates, F. 1949. The design of rotation experiments. Commun. 46. Commonwealth Bur. Soil Sci., Harpenden, UK.

(Appendices 1, 2, and 3 follow)

## APPENDIX I

### GenStat Commands to Analyze the Woburn Ley–Arable Experiment

```
" suppress messages and echoing of command lines "
SET                 [INPRINT=*; DIAGNOSTIC=warnings]
IMPORT              [PRINT=*] 'Table 3.xls'
" calculate orthogonal polynomial contrasts over years "
CALCULATE           X = Year
ORTHPOLYNOMIAL      [MAXDEGREE=4] X; POLYNOMIAL=YearPol
CALCULATE           LinYear,QuadYear,CubYear,QuartYear = YearPol[]
POINTER             [VALUES=LinYear,QuadYear,CubYear,QuartYear] YearPol
CAPTION             'First analysis: do we need different variances in each year?';\
                    STYLE=meta
VCOMPONENTS         [FIXED=Rotation*N*YearPol[]] Year
REML                [PRINT=*] Yield
VRACCUMULATE        [PRINT=*; METHOD=restart] 'Constant variance'
VCOMPONENTS         [FIXED=Rotation*N*YearPol[]; EXPERIMENTS=Year] Year
REML                [PRINT=*] Yield
VRACCUMULATE        [PRINT=deviance,dfrandom,aic]\
                    'Different variance in each year'
CAPTION             'Conclusion: yes we do we need different variances.';\
                    STYLE=stress
CAPTION             'Second analysis: can we simplify the fixed model?';\
                    STYLE=meta
VDISPLAY            [PRINT=Wald]
CAPTION             'Conclusion: drop the cubic and quartic polynomials.';\
                    STYLE=stress
CAPTION             'Third analysis: any further simplification of the fixed model?';\
                    STYLE=meta
VCOMPONENTS         [FIXED=Rotation*N*YearPol[1,2]; EXPERIMENTS=Year] Year
REML                [PRINT=Wald] Yield
CAPTION             'Conclusion: drop Rotation.N.LinYear and Rotation.N.QuadYear.';\
                    STYLE=stress
CAPTION             'Fourth analysis: final model.'; STYLE=meta
VCOMPONENTS         [FIXED=Rotation*N*YearPol[1,2]\
                    – Rotation.N.YearPol[1,2]; EXPERIMENTS=Year] Year
REML                [PRINT=components,Wald] Yield
" form predicted means assuming quadratic year trends "
VARIATE             [VALUES=1981...2000] xlin
ORTHPOLYNOMIAL      [MAXDEGREE=2] xlin; POLYNOMIAL=xpred
VPREDICT            [PRINT=*;PREDICTIONS=predictedmeans; SED=sed]\
                    LinYear,QuadYear,Rotation,N;LEVELS=xpred[],*,*;\
                    PARALLEL=*,LinYear,*,*;NEWFACTOR=*,year,*,*
FACTOR              [LEVELS=!(1981...2000); LABELS=\
                    !t('81','82','83','84','85','86','87','88','89','90',\
                    '91','92','93','94','95','96','97','98','99','00');\
                    MODIFY=yes] year
CALCULATE           averagesed = MEAN(sed)
&                   maxsed = MAX(sed)
&                   minsed = MIN(sed)
CAPTION             'Predicted means'; STYLE=minor
PRINT               [IPRINT=*] predictedmeans
CAPTION             'Standard errors of differences';STYLE=minor
PRINT               !t('average:','maximum:','minimum:'),\
                    !(averagesed,maxsed,minsed); JUST=left,right
" plot predicted means "
PEN                 2...5; SYMBOL=0;LINESTYLE=2,3,5,7; THICKNESS=2
PEN                 –1,–2; THICKNESS=2
DTABLE              [METHOD=line;XFREPRESENTATION=label] predictedmeans;\
                    XFACTOR=year; GROUPS=N; TRELLIS=Rotation; PEN=!(2...5);\
                    TITLE=' '; YTITLE='yield t ha?^{–1}'
```

## APPENDIX 2

### R and ASReml-R Commands to Analyze the Woburn Ley–Arable Experiment

```
# load asreml library
library(asreml)

# read data
Table 3 <- asreml.read.table("Table 3.txt",header = T)
summary(Table 3)

# get factor versions of year and n
Table 3$Year <- as.factor(Table 3$year)
Table 3$N <- as.factor(Table 3$n)
head(Table 3)

# 1: full fixed model with constant variance
model1 <- asreml(fixed=yield ~ Rotation*N*pol(year,4), random=~Year,
data=Table 3)
summary(model1)
wald(model1,denDF="algebraic")
# calculate AIC for this model (on deviance scale – smaller = better)
aic1 <- –2*(model1$loglik – length(model1$gammas))
aic1

# 2: full fixed model with separate variances across years
model2 <- asreml(fixed=yield ~ Rotation*N*pol(year,4), random=~Year,
rcov=~at(Year):id(units), data=Table 3)
summary(model2)
wald(model2)
# calculate AIC
aic2 <- –2*(model2$loglik – length(model2$gammas))
aic2

# compare AIC across models 1 and 2: model 2 better (smaller AIC)
aic1 – aic2

# construct individual vectors to separate out polynomial orders
matpol <- poly(Table 3$year,degree=4)
Table 3$linyear <- matpol[1:480,1]
Table 3$quadyear <- matpol[1:480,2]
Table 3$cubyear <- matpol[1:480,3]
Table 3$quaryear <- matpol[1:480,4]

# 3: model 2 with polynomial components separated
model3 <- asreml(fixed=yield ~
Rotation*N*(linyear+quadyear+cubyear+quaryear), random=~Year,
rcov=~at(Year):id(units), data=Table 3)
summary(model3)
wald(model3,denDF="default")

# 4: drop cubic and quartic polynomial components
model4 <- asreml(fixed=yield ~ Rotation*N*(linyear+quadyear), random=~Year,
rcov=~at(Year):id(units), data=Table 3)
summary(model4)
wald(model4,denDF="default")

# 5: drop 3-way interaction and return to pol function (easier prediction)
model5 <- asreml(fixed=yield ~ Rotation*N*pol(year,2) –
(Rotation:N:pol(year,2)), random=~Year, rcov=~at(Year):id(units),
data=Table 3)
summary(model5)
wald(model5,denDF="default")

# get predictions from final model
model5.pv <-
predict(model5,classify=c("Rotation:N:year"),levels=list(Rotation=1:6,N=1:4
,year=1981:2000))
model5.pv$predictions

# extract results
model5.pred <- model5.pv$predictions$pvals$predicted.value
```

```
model5.pR <- model5.pv$predictions$pvals$Rotation
model5.pN <- model5.pv$predictions$pvals$N
model5.py <- rep(1981:2000, times=24)

# make data frame containing predictions
model5.predict <-
data.frame(pred=model5.pred,Rotation=model5.pR,N=model5.pN,year=model5.py)
model5.predict

# plot predictions
require(lattice)
xyplot(pred ~ year | Rotation, data=model5.predict, groups=N, auto.key=T)
# save to pdf file
pdf(file="xyplot.pdf")
xyplot(pred ~ year | Rotation, data=model5.predict, groups=N, auto.key=T)
dev.off()
```

## APPENDIX 3

### SAS Commands to Analyze the Woburn Ley–Arable Experiment

```
PROC IMPORT OUT = rotation
        DATAFILE = "&pathname.\long-term rotation\Table 3.xlsx";
        SHEET = "Sheet1";
RUN;
* Center covariates;
/*
DATA rotation; SET rotation;
        Year_num = Year_num – 1991;
        N = N – 105;
RUN;
*/
* Origin at Year = 1980;
DATA rotation; SET rotation;
        Year_num = Year_num – 1980;
RUN;
* Obtain the orthogonal polynomial, and merge it with the rest of the data;
PROC IML;
USE rotation;
READ ALL VAR {Year_num} INTO y;
yp = ORPOL(y,4);
cname = {"YearPol0" "YearPol1" "YearPol2" "YearPol3" "YearPol4"};
CREATE yp_data FROM yp [ COLNAME = cname ];
APPEND FROM yp;
RUN;
QUIT;
DATA yp_data2; SET yp_data;
row_no = _N_;
RUN;
DATA rotation2; SET rotation;
row_no = _N_;
ysq = Year_num ** 2;
RUN;
PROC SQL;
        CREATE TABLE rotation3 AS
        SELECT a.*, b.*
        FROM rotation2 AS a, yp_data2 AS b
        WHERE a.row_no eq b.row_no;
QUIT;
* Fit model with heterogeneity of residual variance
* among years, with YearPol3 and YearPol4 omitted,
* with 3-way interactions also omitted
* and with terms in same order as in GenStat;
ODS RTF FILE = "sasrtf het year, no cub quart or 3-way.rtf";
PROC MIXED ASYCOV DATA = rotation3 ;
        CLASS Year Rotation N Plot;
        MODEL Yield = Rotation N YearPol1 YearPol2 Rotation*N
                Rotation*YearPol1 N*YearPol1
                Rotation*YearPol2 N*YearPol2
                / DDFM = KENWARDROGER HType=1 3;
        RANDOM Intercept / subject=Year;
parms 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1;
repeated intercept / subject = Plot*year type=vc group=year;
RUN;
ODS RTF CLOSE;
```